Experiencing Brain-Computer Interface Control

Bram van de Laar

EXPERIENCING BRAIN-COMPUTER INTERFACE CONTROL

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. H. Brinksma,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op woensdag 12 oktober om 16:45 uur 2016

door

Bram Lucas Adriaan van de Laar
geboren op 13 augustus 1984
te Woerden

Promotor: Prof. dr. ir. A. Nijholt en
Copromotor: Dr. M. Poel
hebben dit proefschrift goedgekeurd.

© Bram van de Laar, 2016
ISBN nummer: 978-90-365-4192-3
DOI: 10.3990/1.9789036541923

De promotiecommissie bestaat uit:

Voorzitter: Prof. dr. P.M.G. Apers

Promotor: Prof. dr. ir. A. Nijholt en

Copromotor: Dr. M. Poel

Prof. dr. D. K. J. Heylen

Prof. dr. D. H. Coyle

Prof. dr. M. W. Tangermann

Prof. dr. F. van der Velde

Prof. dr. J. B. F. van Erp

CTIT Ph.D.-thesis Series No. 16-423
Centre for Telematics and Information Technology, University of Twente
P.O. Box 217, NL – 7500 AE Enschede

SIKS Dissertation Series No. 2016-45
The research reported in this thesis has been carried out under the
auspices of SIKS, the Dutch Research School for Information and
Knowledge Systems.

ABSTRACT

Brain-Computer Interfaces (BCIs) are systems that extract information from the user's brain activity and employ it in some way in an interactive system. While historically BCIs were mainly catered towards paralyzed or otherwise physically handicapped users, the last couple of years applications with a focus on entertainment meant for healthy users gained a lot of momentum in research. While from a disabled user's perspective functionality and accuracy have been key to get a working system, especially for healthy users the user's experience (UX) can be considered even more important. A vast amount of effort has been put into increasing the accuracy of various types of BCIs, less research has been done on the impact this accuracy has on the UX.

This thesis is structured in the following way: I Introduction, II Studies and III conclusion. Within part II, 4 systems and experiments are conducted and described. With these experiments we try to show what the influence of accuracy on the UX is.

In the first chapter of Part II we outline an experiment in which a large group of users played a simulation of various levels of control (similar to the way a BCI works). This online experiment consisted of a rather simple game in which users controlled a fictitious hamster with a fixed amount of control. After every level of control the user answered a short questionnaire on UX. We found that for the lower half of the control scale, a linear relationship exists between control and the fun a user experiences. The interesting finding was that for higher levels of control, fun reaches an optimum and even tends to decrease above a certain level of control. The non perfect control could make the game more interesting to people, an indication that we can use a BCI to make a game more challenging and interesting.

In the second chapter we describe an experiment which compares the modalities of using actual movements and imagined movements (or Motor Imagery, MI) in an Event Related (De)Synchronization (ERD/ERS) based BCI. Within the group of 20 participants the average accuracy was higher for actual movement, but most participants found imagined movement to be more challenging and fun. MI is a popular paradigm within the field of BCI, especially physically handicapped or paralyzed users can still use their brain activity to control such a system. For healthy users however, we can still use the signals from actual movements as well. Which and why signals fare better is discussed in this chapter.

Building on the first and second chapters in which we simulated a BCI and in which the input was solely from a BCI; in the third chap-

ter we looked at whether a game with combined keyboard, mouse and BCI input would be favoured against the classical game with just keyboard and mouse. In a large study with 48 participants we found that the participants enjoyed playing the game with BCI control as long as the one without the use of a BCI, while their level of perceived control was significantly lower in the game with the use of BCI. This chapter shows us that by implementing the BCI input in such a way that it's not detrimental to the UX (i.e. not frustrating or boring) the advantages (i.e. interesting new technology, a different modality to master besides hand-eye coordination) can overcome the disadvantages.

In the fourth chapter, we present BrainBrush, a system that takes input from three different modalities: brain activity (by means of the P300 in the EEG), eye blinks (also from the EEG) and head movement (by means of a gyroscope). The experiment that was carried out is of a qualitative nature but also used a well-validated questionnaire for the usability of the system. After two design iterations the final results show that a system including multiple modalities can be engineered which can measure itself with conventional (i.e. not using brain activity) systems. What we see is that these physiological signals can result in lower control for some, with careful design the system can still be useful and provide a good experience.

From the results of all these studies we can draw the conclusion that using a BCI as an input channel can make a game or system more challenging and interesting, although it's far from perfect as a controller. However, care has to be taken when implementing a BCI-based input. By evaluating the UX in user studies we are able to see whether the BCI just frustrates the user or adds an extra dimension.

SAMENVATTING

Brein-Computer Interfaces (BCI's) zijn systemen die gebruik maken van de hersenactiviteit van de gebruiker van het systeem. De laatste jaren zijn er op het gebied van de neurowetenschappen, elektrotechniek en informatica veel technieken, apparaten en methodes ontwikkeld die dit mogelijk maken. Sommige BCI's gebruiken enkel hersenactiviteit en geen enkele andere vorm van invoer, deze methode vloeide voornamelijk voort uit de belangrijkste aanjager voor het gebruik van BCI's, patiënten met Amyotrofe Laterale Sclerose (ALS), nekwerfelfractuur of anderszins verlamde of anderszins in hun beweging en uitdrukingskracht beperkte personen. De laatste jaren wordt er echter ook steeds meer gekeken naar toepassingen voor gezonde personen en dit proefschrift is daar een voorbeeld van. De belangrijkste vraag waar ik in dit proefschrift antwoord op probeer te geven is wat de invloed van de precisie van de BCI is op hoe de gebruiker het

systeem ervaart.

Deze thesis is opgebouwd uit de volgende delen: I Introductie, II Experimenten and III Conclusie. Deel II bestaat uit de beschrijving van vier verschillende BCI systemen waar elk een experiment mee is uitgevoerd om de relatie tussen de precisie van de BCI en de gebruikerservaring (UX) te onderzoeken.

In het eerste hoofdstuk van Deel II wordt een experiment beschreven met een grote groep gebruikers die een spelletje speelden met een simulatie van een BCI waarin verschillende niveau's van controle wordt gemoduleerd. Dit spelletje kon destijds online gespeeld worden en bestond uit verschillende doolhoven waaruit de hamster die door de proefpersonen bestuurd werd moest ontsnappen. Na elke ronde in het spel kreeg de gebruiker een korte vragenlijst over de afgelopen rond voorgeschoteld. Gebruikers speelden meerdere rondes en zo kregen we genoeg data om een en ander te onderzoeken. Zo zagen we bijvoorbeeld dat voor de lage waardes wat betreft controle er een lineaire relatie was tussen de hoeveelheid control en hoe leuk de proefpersonen het spelletje vonden. Interessanter nog was dat juist voor de hogere waardes van controle het plezier aftopt en een maximum bereikt. Een manier om dit te verklaren is dat de niet perfecte controle over de hamster een extra uitdaging creëert die spel interessanter maakt.

Het tweede hoofdstuk beschrijft een experiment waarin twee modaliteiten om een BCI mee aan te sturen (op basis van ERD: Event Gerelateerde Desynchronisatie in het EEG) gebruikt worden: echte bewegingen en ingebeelde bewegingen (Motor Imagery, MI). In de groep van 20 proefpersonen was de gemiddelde precisie (wat overeenkomt met controle) hoger voor de herkenning van echte bewegingen uit het EEG dan voor ingebeelde bewegingen. Maar de meeste proefpersonen vonden de ingebeelde bewegingen uitdagender en leuker. De MI modaliteit wordt veel gebruikt voor mensen die een handicap hebben zoals eerder genoemd. Voor gezonde gebruikers kunnen natuurlijk ook de echte bewegingen gebruikt worden. Een uitgebreide discussie van de voor- en nadelen staat in dit hoofdstuk.

Vorbordurent op het eerste en tweede hoofdstuk in Deel II waarbij er eigenlijk maar sprake is van 1 modaliteit, kijken we in het derde hoofdstuk naar een experiment waarin 3 modaliteiten (toetsenbord, muis en BCI) geprefereerd wordt boven de ordinaire combinatie zonder BCI. In een groot experiment met maar liefst 48 proefpersonen zagen we dat de proefpersonen net zo lang bleven spelen met de BCI als toevoeging als zonder BCI. Ze gaven tegelijkertijd aan dat hun controle in het spel een stuk lager was in het geval van de BCI. Wat dit hoofdstuk in feite aantoont is dat het implementeren van een BCI in een systeem kan gebeuren zonder dat het negatieve effecten heeft op de gebruikerservaring door bijvoorbeeld frustratie of saaiheid vanwege een slechtwerkende BCI. De voordelen, zoals het gebruiken van

een vernieuwende technologie en het bieden van een nieuwe uitdaging, kunnen dan naar de voorgrond geschoven worden.

In het vierde hoofdstuk presenteren we BrainBrush, een systeem dat gebruik maakt van 3 niet-alledaagse modaliteiten, namelijk: hersenactiviteit, oogknipperingen en hoofdbewegingen. Met dit systeem is een voornamelijk kwalitatieve analyse uitgevoerd. Naast deze analyse is er ook een uitstekend gevalideerde vragenlijst voor de bruikbaarheid van systemen gebruikt die erg algemeen toepasbaar is. Na een dubbele iteratie van het ontwerp van het systeem laat de score uit deze vragenlijst zien dat BrainBrush, ondanks de van nature onnauwkeurige fysiologische signalen, qua bruikbaarheid het helemaal niet slecht doet en zeer bruikbaar is.

Het resultaat van al deze experimenten is dat BCI een spel of een systeem uitdagender en/of interessanter kan maken voor gebruikers, ondanks het onnauwkeurige karakter van de signalen. Er moet hoe dan ook goed opgelet worden bij het ontwerpen van zulke systemen. Voorts dienen er duchtige analyses van de gebruikerservaring gedaan te worden, zowel tijdens als na het ontwerpproces om te voorkomen dat een onnauwkeurige BCI voor frustratie zorgt.

PUBLICATIONS

0.1 PUBLICATIONS CONTRIBUTING TO THIS THESIS:

- Van de Laar, B.L.A. and Gürkök, H. and Plass-Oude Bos, D. and Poel, M. and Nijholt, A. (2013) Experiencing BCI Control in a Popular Computer Game In: IEEE Transactions on Computational Intelligence and AI in Games, 5 (2). pp. 176-184. ISSN 1943-068X, ISI Impact 1.69
- Van de Laar, B.L.A. and Reuderink, B. and Plass-Oude Bos, D. and Poel, M. and Nijholt, A. (2013) How much control is enough? Influence of unreliable input on user experience In: Cybernetics, IEEE Transactions on. Special Issue on Modern Control. vol.43, no.6pp. 1584-1592, Dec. 2013 doi: 10.1109/T-CYB.2013.2282279, ISI Impact 3.236
- Van de Laar, B.L.A. and Plass-Oude Bos, D. and Reuderink, B. and Heylen, D.K.J. (2009) Actual and Imagined Movement in BCI Gaming. In: Proceedings of the International Conference on Artificial Intelligence and Simulation of Behaviour (AISB 2009), 06-09 Apr 2009, Edinburgh, Scotland. SSAISB, Brighton. ISBN 1902956818
- Van de Laar, B.L.A. and Gürkök, H. and Plass-Oude Bos, D. and Nijboer, F. and Nijholt, A. (2012) Brain-Computer Interfaces and User Experience Evaluation. In: Towards Practical Brain-Computer Interfaces: Bridging the Gap from Research to Real-World Applications. Biological and Medical Physics, Biomedical Engineering. Springer Verlag, Berlin Heidelberg, pp. 223-237. ISSN 1618-7210 ISBN 978-3-642-29746-5
- van de Laar, B.L.A. and Gürkök, H. and Plass-Oude Bos, D. and Nijboer, F. and Nijholt, A. (2011) Perspectives on User Experience Evaluation of Brain-Computer Interfaces. In: Universal Access in Human-Computer Interaction. Users Diversity. 6th International Conference, UAHCI 2011, Held as Part of HCI International 2011, 9 - 14 July, Orlando. pp. 600-609. Lecture Notes in Computer Science 6766. Springer Verlag. ISSN 0302-9743 ISBN 978-3-642-21662-6
- Plass-Oude Bos, D. and Gürkök, H. and van de Laar, B.L.A. and Nijboer, F. and Nijholt, A. (2011) User Experience Evaluation in BCI: Mind the Gap! International Journal of Bioelectromagnetism, 13 (1). pp. 48-49. ISSN 1456-7857

- Gürkök, H. and Plass-Oude Bos, D. and van de Laar, B.L.A. and Nijboer, F. and Nijholt, A. (2011) User Experience Evaluation in BCI: Filling the Gap. *International Journal of Bioelectromagnetism*, 13 (1). pp. 54-55. ISSN 1456-7857
- Van de Laar, B.L.A. and Nijboer, F. and Gürkök, H. and Plass-Oude Bos, D. and Nijholt, A. (2011) User Experience Evaluation in BCI: Bridge the Gap. *International journal of bioelectromagnetism*, 13 (3). pp. 157-158. ISSN 1456-7857
- Van de Laar, B.L.A. and Reuderink, B. and Plass-Oude Bos, D. and Heylen, D.K.J. (2010) Evaluating user experience of actual and imagined movement in BCI gaming. *International journal of Gaming and Computer Mediated Simulations*, 2 (4). pp. 33-47. ISSN 1942-3888
- Van de Laar, B.L.A. and Brugman, I. and Nijboer, F. and Poel, M. and Nijholt, A. (2013) BrainBrush, a Multimodal Application for Creative Expressivity. In: *Sixth International Conference on Advances in Computer-Human Interactions (ACHI 2013)*, 24 Feb-1 March 2013, Nice, France. pp. 62-67. IARIA XPS Press. ISBN 978-1-61208-250-9

0.2 PUBLICATIONS NOT DIRECTLY CONTRIBUTING TO THIS THESIS:

- Plass-Oude Bos, D. and Duvinage, M. and Oktay, O. and Delgado Saa, J. and Guruler, H. and Istanbulu, A. and van Vliet, M. and van de Laar, B.L.A. and Poel, M. and Roijendijk, L. and Tonin, L. and Bahramisharif, A. and Reuderink, B. (2011) Looking around with your brain in a virtual world. In: *IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB 2011)*, 11-15 Apr 2011, Paris, France. pp. 1-8. IEEE Computational Intelligence Society. ISBN 978-1-4244-9890-1
- Plass-Oude Bos, D. and van de Laar, B.L.A. and Duvinage, M. and Oktay, O. and Delgado Saa, J. and van Vliet, M. and Poel, M. and Roijendijk, L. and Bahramisharif, A. and Reuderink, B. (2011) Wild Photoshoot: Applying Overt and Covert Attention. In: *Society of Applied Neuroscience Meeting 2011*, 5-8 May 2011, Thessaloniki, Greece. pp. 26-27. Laboratory of Medical Informatics, Medical School, Aristotle University of Thessaloniki. ISBN 978-9-60243-679-9
- Plass-Oude Bos, D. and Reuderink, B. and van de Laar, B.L.A. and Gürkök, H. and Mühl, C. and Poel, M. and Nijholt, A. and

- Heylen, D.K.J. (2010) Brain-Computer Interfacing and Games. In: Brain-Computer Interfaces. Applying our Minds to Human-Computer Interaction. Human-Computer Interaction Series. Springer Verlag, London, pp. 149-178. ISSN 1571-5035 ISBN 978-1-84996-271-1
- Plass-Oude Bos, D. and Reuderink, B. and van de Laar, B.L.A. and Gürkök, H. and Mühl, C. and Poel, M. and Heylen, D.K.J. and Nijholt, A. (2010) Human-Computer Interaction for BCI Games: Usability and User Experience. In: Proceedings of the International Conference on CYBERWORLDS 2010, 20-22 Oct 2010, Singapore. pp. 277-281. IEEE Computer Society. ISBN 978-0-7695-4215-7
  - Van de Laar, B.L.A. and Nijholt, A. and Zwiers, J. (2010) Monitoring User's Brain Activity for a Virtual Coach. In: 9th International Conference on Entertainment Computing (ICEC 2010), 8-11 September 2010, Seoul, Korea. pp. 511-513. Lecture Notes in Computer Science 6243. Springer Verlag. ISSN 0302-9743 ISBN 978-3-642-15398-3
  - Reidsma, D. and van Welbergen, H. and Paul, R.C. and van de Laar, B.L.A. and Nijholt, A. (2010) Developing Educational and Entertaining Virtual Humans using Elckerlyc. In: 9th International Conference on Entertainment Computing (ICEC 2010), 8-11 September 2010, Seoul, Korea. pp. 514-517. Lecture Notes in Computer Science 6243. Springer Verlag. ISSN 0302-9743 ISBN 978-3-642-15398-3 2009

CONTENTS

0.1	Publications contributing to this thesis:	ix
0.2	Publications not directly contributing to this thesis:	x
I	PART I INTRODUCTION	1
1	INTRODUCTION	3
1.1	General Introduction	3
1.2	Unreliable input	3
1.2.1	Physiological Modalities	4
1.2.2	BCI input	5
1.3	User experience and BCI modalities.	7
1.4	Passive BCIs and UX	8
1.4.1	BCI Games evaluated on UX	8
1.4.2	Immersion and Presence	9
1.5	Motivation and Research question	10
1.6	Overview of the thesis	11
1.7	Contributing Articles	12
2	BRAIN-COMPUTER INTERFACES AND USER EXPERIENCE EVALUATION	15
2.1	Abstract	16
2.2	Introduction	16
2.3	Current State of User Experience Evaluation of BCI	17
2.3.1	User Experience Affects BCI	17
2.3.2	BCI Affects User Experience	19
2.4	Applying HCI User Experience Evaluation to BCIs	20
2.4.1	Observational analysis	20
2.4.2	Neurophysiological measurement	21
2.4.3	Interviewing and questionnaires	22
2.4.4	Other methods	23
2.5	Case studies	24
2.5.1	Case study: Mind the Sheep!	24
2.5.2	Case study: Hamster Lab	26
2.6	Discussion and Conclusions	28
II	PART II STUDIES	31
3	HOW MUCH CONTROL IS ENOUGH? INFLUENCE OF UNRELIABLE INPUT ON USER EXPERIENCE	33
3.1	Abstract	34
3.2	Introduction	34
3.2.1	Unreliable input	35
3.2.2	BCI based input	36

3.3	Related Work	38
3.3.1	How to measure control?	38
3.3.2	Dealing with Errors in BCIs	40
3.3.3	Perception of BCI Control	40
3.3.4	Perception of Control	41
3.4	Methods	42
3.4.1	Experimental Design	42
3.4.2	Game	43
3.4.3	Levels of control	44
3.4.4	Participants	46
3.4.5	Questionnaire	46
3.5	Results	46
3.5.1	Perceived control	47
3.5.2	Fun - Control	48
3.6	Discussion	51
4	EVALUATING USER EXPERIENCE OF ACTUAL AND IMAGINED MOVEMENT IN BCI GAMING	55
4.1	Abstract	56
4.2	Introduction	56
4.3	Related Work	57
4.4	Methods	58
4.4.1	Experiment Setup	58
4.4.2	BrainBasher	59
4.4.3	The BrainBasher BCI	63
4.4.4	BCI performance evaluation	64
4.4.5	Questionnaire design	66
4.5	Results	66
4.6	Discussion	71
5	EXPERIENCING BCI CONTROL; TO BCI OR NOT TO BCI?	75
5.1	abstract	76
5.2	Introduction	76
5.3	Related Work	77
5.3.1	BCI Games	77
5.3.2	User experience	78
5.3.3	Measuring Immersion	78
5.4	Methods	79
5.4.1	World of Warcraft	79
5.4.2	α WoW	80
5.4.3	Experimental setup	80
5.4.4	Participants and instructions	84
5.4.5	Validation of the control signal	86
5.5	Results	86
5.5.1	Item Analysis	87
5.5.2	Duration Estimation	87
5.5.3	Analysis of Variance	89

5.5.4	Second experiment: accuracy	89
5.6	Discussion	90
6	BRAINBRUSH, A MULTIMODAL APPLICATION FOR CREATIVE EXPRESSIVITY	95
6.1	abstract	96
6.2	Introduction	96
6.3	Methods	99
6.3.1	Development of BrainBrush	99
6.3.2	Experimental Design	103
6.4	Results	103
6.5	Discussion and Conclusion	106
III	PART III CONCLUSION	109
7	GENERAL DISCUSSION	111
7.1	Discussion of chapter 3	111
7.1.1	Dataplots	111
7.1.2	Curve fit and sample size	112
7.2	Discussion of chapter 4	114
7.3	Discussion of chapter 6	116
8	CONCLUSIONS	117
8.1	Answering the Research Question	117
8.1.1	Control	117
8.1.2	UX evaluation of a BCI game with direct control	118
8.1.3	UX evaluation of an additional BCI control in a popular game	118
8.1.4	BCI control in a multimodal application	119
8.2	Overall Conclusions	119
8.2.1	Passive BCIs	119
8.2.2	BCI for direct control	120
8.3	Limitations and Future work	120
8.3.1	The games	121
8.3.2	BCI hardware	121
8.3.3	Participants	122
IV	APPENDICES	125
A	APPENDICES	127
A.1	Questionnaire used in Chapter 3	127
A.2	Questionnaire used in Chapter 5	128
A.3	Questionnaire used in Chapter 6	130
	BIBLIOGRAPHY	131

LIST OF FIGURES

Figure 1	A classification of current user experience evaluation methods used in human-computer interaction for entertainment technologies (adapted from [82]).	21
Figure 2	A screenshot from the game Mind the Sheep! depicting the game world with ten sheep, three dogs and the pen.	24
Figure 3	A screenshot from the game Hamster Lab depicting the game world with the hamster in the first level of the game.	27
Figure 4	Runs, rounds and levels.	43
Figure 5	A screenshot of the game that was used for the experiment.	43
Figure 6	The relation between the MI and the accuracy for a five-class confusion matrix with the same probability for all correctly detected events, and the same probability of each pair of mistaking one event for another.	45
Figure 7	The median values of Perceived control are plotted for all conditions (varying amounts of Control(MUI)) in the experiment. A linear regression through all data is plotted as well ($R^2 = 0.499$).	49
Figure 8	The median values of Frustration are plotted for all conditions (varying amounts of Control(MUI)). A linear regression through all data is plotted as well ($R^2 = 0.133$).	49
Figure 9	Medians of Fun vs. Control and the third order curvefit through all data points.	50
Figure 10	Overview of the experiment setup	60
Figure 11	The symbols for left and right hand movement.	61
Figure 12	A game session.	62
Figure 13	BrainBasher System View	62
Figure 14	Photograph of a subject playing BrainBasher.	63
Figure 15	BrainBasher BCI Pipeline	65
Figure 16	Photograph of the Emotiv EPOC headset that was tilted forward about 25°	82

Figure 17	The pipeline including the signal processing steps as implemented in Alpha WoW. For a detailed description see section 5.4.3.1.	82
Figure 18	Examples of EEG data during Eyes closed, 'Relaxation' and 'Stressed' states.	82
Figure 19	Overview of the experimental setup	83
Figure 20	Screenshot of WoW with the add-on enabled. The orange bar (top left) provides feedback on the user's current alpha level. Bigger means less alpha.	84
Figure 21	Screenshot of the ball/basket game.	87
Figure 22	User interface in Painting mode with the 'New Painting' menu option at the top and the 'Undo' and 'Redo' menu options at the bottom	101
Figure 23	User interface during brush selection showing the 11 available brushes and the eraser	102
Figure 24	User interface during color selection showing the 12 available colors	102
Figure 25	BrainBrush system overview	102
Figure 26	Ordering of pleasantness of modalities: Head movement was most often ranked most pleasant, eyeblinks most often second and the P3Speller (BCI) most often last.	105
Figure 27	2nd to 5th order curve fits, relating to the curve fit analysis in Chapter 3.	112

LIST OF TABLES

Table 1	Constructs and items	47
Table 2	Number of samples per condition.	48
Table 3	1 st through 4 th order models with respective R^2 , AIC and the change in AIC from the lower order polynomial.	50
Table 4	Constructed Scales including alpha and variance explained by 1st principal component.	68
Table 5	Paired <i>t</i> -Tests Scales, comparing imagined and actual movement	69

Table 6	ITR for actual and imagined movement for all subjects (in bits/minute), missing data is marked with x. 70
Table 7	Grand averages of participants. Durations per condition (min.) and number of over- and underestimations of duration. 88
Table 8	Averages of experienced and inexperienced participants. Durations per condition (min.), number of over and underestimations of duration and average amount of over and underestimations. 89
Table 9	Multivariate tests. Effect, Wilks' Lambda, F-measure, hypothesis df, Error df and p-value. 'Experience' is between-subjects, 'Condition' is within-subjects 90
Table 10	Within-subjects and between-subjects univariate tests. Source of the effect, Effect, Mean Squares, F-measure, hypothesis df, Error df and p-value 90
Table 11	Subject accuracies, average, standard deviation and Confidence Interval. 91
Table 12	P300 classification accuracy during copy spelling and free painting 104
Table 13	Summary of interview results 105
Table 14	F-test of curve fit. 113
Table 15	Bartlett's test of sphericity and Kaiser-Meyer-Olkin test for the questionnaires scales used in Chapter 4 116

LISTINGS

ACRONYMS

Part I

PART I INTRODUCTION

This part contains a general introduction to Brain-Computer Interfaces and the various aspects of user experience evaluation that play an important role in this thesis. Finally we formulate our motivation and the research questions addressed in this thesis.

INTRODUCTION

1.1 GENERAL INTRODUCTION

Brain-computer interfaces (BCIs) aim to provide a reliable control signal for assistive technology for people. BCIs make use of the brain activity of the user to control anything from a wheelchair to a video game. With the merge of the fields of human-computer interaction (HCI) and BCI new applications are being developed for entertainment and education which may be of interest to users with and without disabilities. BCIs will be integrated into existing interactive applications. The aim of such applications is to create positive experiences that enrich our lives rather than only providing reliable control. It is suggested at several keynote presentations at large BCI conferences that reliability is the most important issue to be addressed to achieve technology transfer to the market and the society. However, perfectly reliable systems are not necessarily usable. Even reliable assistive technologies may get abandoned by users when usability is not warranted [122]. Making interactive systems usable is the core expertise of the field of HCI. The process of designing interactive systems in the field of HCI consists of analysis of requirements, design and implementation of the system and user evaluation. To evaluate such systems, the user experience (UX) needs a more important role in BCI studies. Researchers should not only focus on the reliability of the control signal, so that we can better understand how such a system can satisfy the needs of the user. In the following sections we will introduce various concepts important to the field of other input modalities such as BCIs, user experience (UX) evaluation and related concepts. Hereafter we motivate our research and introduce the main research question.

1.2 UNRELIABLE INPUT

Recent developments in interfaces between humans and machines show that there is a need for less artificial means of interaction between the two. The most prominent examples of the moment are the Nintendo Wii and the Microsoft Kinect, both gesture interfaces. But speech, eye gaze, and other physiological measures are also promising a more intuitive way of inter-

action. By allowing the user to apply knowledge from previous interactions, for example from interacting with the real world or from interacting with comparable systems, the interface is easy to learn, easy to remember, and easy to use, which are key aspects for usable systems [89, 14, 93]. Brain activity as input modality also has potential in this area, as it can provide some insight in the intention of the user, without depending on external expression. Unfortunately, most of the current systems are still in the phase of proving that using brain activity for control is even possible, and are therefore not making full use of the intuitiveness this input could provide. The biggest issue right now is the reliability with which we can recognize and analyze the signals we need.

1.2.1 *Physiological Modalities*

One thing physiology-based inputs have in common is that the interpretation of the input is often problematic. This is mainly because of the noisiness and ambiguity of the input, but also because of the problem of intentionality, see [31, 114, 63] for example. This noise and ambiguity will make an unreliable input channel. We try to clarify this through three examples:

1.2.1.1 *Keyboard*

In case of input through a keyboard, the keys which are pressed by a user are always recognized as what was typed in. There is (in case of a wired keyboard) no measurable noise between the keyboard and the computer and no problematic interpretation of the keystrokes. The only possibility for an error to occur is by uncontrolled motor activity of the user. A recognition or classification rate is not applicable in this case.

1.2.1.2 *Vision-based*

In case of a gaze, gesture or other vision-based interface: Interfaces based on gaze, gestures et cetera depend on visually capturing (part of) the user. Either eyes or limbs point in a certain direction or make a certain movement which is classified by the computer and translated into an action within the system. Users can make errors by not accurately carrying out what they intended to do. As the system depends on the captured image, different kinds of artefacts and noise can hinder classification of the correct action. Pixel noise, different lighting, occlusion, movement of the person, movement of the camera, vibrations

et cetera all may have a negative impact on the recognition rate. Input can change over time, for example as the user becomes more fatigued and is less expressive, or as the sun sets and the lighting changes. Besides, there can also be a large variability between users, such as between children and elderly, or men and women. As an example of ambiguity: when somebody waves their hand, it could mean ‘good bye’, ‘hello’, or even ‘no’. A final problem is intentionality (also referred to as the Midas touch problem [64]). Not all actions will have a purposeful intention related to it. What if the user was not waving at the system, but waving to get rid of a mosquito passing by? Or in the case of an eye tracker, the user will already look at the system simply to take in information. In that case, not every eye gaze is meant as an input command. Especially when the system is always on, there will be times when the user is not purposefully interacting with the system.

1.2.1.3 *Brain-based*

Brain-Computer Interfaces (BCI) are dependent on the brain activity generated by the user performing a task. This can either be the user attending a visual, auditory or tactile stimulus, or by generating brain activity by actively performing a mental task (e.g. imaginary motor movement, performing mental calculus). Errors can occur when the user is not attending the stimulus or not (correctly) performing the mental task. But because of the highly varying nature of the resulting measured signal, the system will make errors in recognizing the users intention to attend/not attend or to perform the mental task or not. Recognition rates vary greatly between users and within users over time, largely independent of the type of BCI [139]. Related to the Midas touch problem, in BCIs it is difficult to detect whether the user is performing a task or not (idling) [152].

In the last two cases, there is an inherent noise (with different causes) on the input channel to the system. This noise makes recognizing the users intentions non-trivial. We therefore define unreliable input as the system not being able to reliably recognize the users intention, caused by inherent noise on the input channel.

1.2.2 *BCI input*

In the case of brain-computer interfaces, each of these problems can be even more difficult to solve as there is no way to obtain a ground truth. Systems based on visual input can be evalu-

ated using expert assessments on recordings of for example gaze and movements. First, BCIs require robust and noise-free brain activity recordings which is problematic at user's homes. For the general population, undergoing surgery to get electrodes implanted is also not a viable option, as the surgery, but also long-term implantation of these electrodes, are still too risky. Finally, there is the issue of response time. If a system is to be used for direct control, the response time should be minimal. This means that systems that depend on indirect measures such as increased blood flow in more active brain areas will be considered too slow for this purpose. What the general home-user then is left with is electroencephalography (EEG). Unfortunately, this measurement is highly sensitive to noise, both from the environment and from the user's body. It has a good temporal resolution, but because it uses electrodes on the outside of the head, it is difficult to exactly measure the signal of interest. Instead a smeared out, attenuated signal from numerous interfering sources is measured. Even if all this would have been perfect, the brain has multiple inter-cell interactions that not only use one type of interaction (0 or 1) but multiple interactions and many different neurotransmitters. As an example, certain areas of the brain are involved in making gestures. These areas may also activate, however, when the user is imagining to make that gesture, or when the user is looking at somebody else making that gesture [35], [85]. The problem of intentionality also still remains, as the user's brain may be responding to something that is not at all related to this particular part of the interaction with the system. As a result, the interpretation of such physiological input modalities may never be perfect, and at the moment, brain input seems to be the least perfect of all.

In most studies concerning user control, the input itself is considered to be near perfect, although mistakes may still be made because the user is distracted, unskilled, or is unsure about what to do. The general solutions that are provided to solve problems caused by imperfect control are therefore generally in the range of: make sure that the system is responsive (small delays for feedback or updated system status), that the feedback is easily understood, and include an undo button [93]. There are very little guidelines for interfaces where the control input itself may be a critical issue. How many mistakes can be made before a system becomes unusable or unacceptable? One could argue that it is not the actual control that matters, only the user's perception, but how do these two relate? And do we even need to aim for perfect control? Especially in the case of entertainment applications, some imperfections may add to the challenge of the task up until the point the user gets frustrated. Keeping in

mind a not perfect control when designing a game can keep the user in a state of flow while they deal with the problem and learn how to cope with the imperfect control [99].

1.3 USER EXPERIENCE AND BCI MODALITIES.

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 electroencephalogram (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 [97].

Games are usually the first applications to adopt new paradigms, driven by the gamers' continuing search for novelty and challenges [98]. 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 [46, 77]. 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 [103]. This was done by comparing keyboard control with BCI control for a simple game called BrainBasher, and evaluating the user experience with the Game Experience Questionnaire [61]. Both this game and this questionnaire were also used for the study described in chapter 3, comparing actual and imagined movement.

Before BCI technology can be accepted, adopted and utilized 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 the phenomenon that some people are unable to use a BCI at all [120]. 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 [75].

While most research into using movement for BCI has focused on imagined movement, some clinical research shows that actual movement in fact elicits a more pronounced and therefore better discernable signal in the motor cortex [86].

Actual movement can also be used with other interfaces than a BCI. Interfaces such as motion tracking systems, for example, which are probably more reliable at this moment. One big potential advantage of a BCI however is that the measured brain signals are always preceding actual muscle activity at the limbs, and can be measured before the muscles activate. This advantage is amplified by the onset of a potential in preparation of a movement, the so called Bereitschaftspotential, or Readiness Potential [71, 124]. A very useful aspect of the RP is the lateralized readiness potential (LRP), where the preparation of left-sided movement is reflected in a potential occurring at the right motor cortex, and vice-versa. Krauledat et al. show that this lateralized readiness potential can be used to classify actual movement even before the movement itself is carried out [72]. This could give a gamer an advantage over other interfaces especially in fast paced, highly reactive games.

But using BCI can also provide other benefits. While measuring brain activity for detection of movement, whether actual or imagined, other information can be derived from the brain as well, such as the user's mental and emotional state. This could be used to make smarter applications which are more aware of the user.

1.4 PASSIVE BCIS AND UX

A good game provides an immersive experience to the user, giving them the feeling they are *in* the game. But BCIs are inherently different from the classical input devices such as the mouse, keyboard and joystick. A BCI poses the user with an unreliable input channel to control the game. Does this hinder immersion and feelings of presence in the game? Can a BCI be of value in a popular modern game? Is it fun to play with BCI control?

1.4.1 BCI Games evaluated on UX

BCI games have been developed by several research groups, often as proof-of-concepts or to evaluate the use of mental task in an online application. One classic example of a BCI game which also inspired us to some extent is Brainbal, developed by Hjelm

et al. [56, 55]. The concept of the game is to relax more than your competitor. Thus two players are competing to be more relaxed, which makes it a paradoxical and fun game. In this study the ratio between frontal alpha and beta waves is taken as a measure of relaxedness.

Pineda et al. [107] developed a BCI for a 3D shooter game. Forward and backward movements were controlled by the keyboard, turning left and right by alpha levels over the motor cortex. As this study focussed on the ability of participants to learn to control their alpha/mu levels over the course of several weeks the study only included 4 participants. They found that control over mu activity was easily obtained and maintained.

Instead of focussing only on bit rates and accuracies Plass-Oude Bos et al. applied a user-centered approach and found that the ease of executing a certain mental task is an important factor as well [113]. Gürkök et al. found that certain mental tasks for BCI games while being appropriate might not be the best option as the major modality for interaction. [50].

1.4.2 *Immersion and Presence*

According to Witmer et al. [145] both involvement and immersion are necessary for experiencing presence. Witmer et al. introduced a presence questionnaire addressing these factors for specific use in virtual environments. Nowadays, games can be at least as realistic as virtual environments. A succesful game drags the user along with its immersive graphics, compelling characters and creative narrative. Users are then able to feel themselves present in the environment or world that is created by the game.

The sense of presence can be interrupted by distracting factors such as audiovisual stimuli that are not congruent with the virtual world or interfaces that are unnatural or faulty. At the intersection of these games which aim at providing the ideal circumstances for a feeling of high presence, evaluation of the user experience is key to make games with BCIs succesful. Van de Laar et al. [134] made an attempt to evaluate a BCI game with a questionnaire based on the GEQ by IJsselsteijn et al. [61] with specific items on the role of the BCI. Van de Laar et al. [135] gave an overview of the several methods available to researchers to evaluate BCI games and when to use them. Administering a questionnaire is the best method for quantifying results. To be able to answer the *why* question, interviewing would be a better method.

1.5 MOTIVATION AND RESEARCH QUESTION

In the earliest stages of a new technology (such as BCIs) showing that the principle works, making a proof-of-concept and improving performance are the most vital of all. This can be captured under the category 'Functionality'. Functionality is something that fills a need. If a certain problem exists that cannot be solved with existing tools or apparatuses, a new one has to be invented. To apply this to the case of BCIs: studies showed that (amongst others) ALS (Amyotrophic Lateral Sclerosis) or SCI (Spinal Cord Injury) patients who suffer from locked-in syndrome are still able to cognitively function. With the use of a BCI they are again able to communicate with the outside world. Once this basic need has been filled, the next step is to look at the usability of the tool, does it do what it has to do efficiently and effectively? This also contains higher level concepts like learnability and usefulness. In the case of a BCI: can we improve the speed and accuracy at which the user can communicate? Until recently these were the foci of research within the field of BCIs. Inventing and developing proof-of-concepts of BCIs with novel mental tasks to efficiently and effectively communicate through brain activity and to improve the speed and accuracy of doing so. For users who cannot communicate otherwise filling the basic need of communicating is the most important of all, how the software that comes with the BCI works or how long it takes the technician to set up the hardware is of secondary importance. However, with the advent of cheaper EEG hardware and affordable and energy efficient computer power even in mobile devices, the user group has expanded tremendously. Now also early adoptors of new technology, gamers, actually everybody interested to try can put on an EEG device and use a BCI. Although this provides near endless possibilities for new applications, this also poses a lot of interesting questions as not only the functionality and the usability are important (otherwise nobody without the need for a BCI would even consider it), but also the user experience (UX) comes into play. McNamara et al. show us a nice overview of the three different concepts we discussed here [87]. User experience contains higher level concepts as immersion (the user is involved and/or lost track of time), fun, engagement, presence (in case of a game, users experience being 'in' the virtual world) et cetera. Even though usability and user experience evaluation is not common in current BCI studies, the user's experience may influence objective performance measures, such as BCI classifier accuracies, and has a big impact on whether users are actually willing to use a specific system. But because it is nigh impossible to evaluate

all factors and variations in software that influence software we choose to investigate what the influence of the accuracy of the BCI itself is on the user experience of the whole application. The most obvious application for healthy users are video games as they provide a platform for new technologies to be exploited. This leads us to the main research question of this thesis:

How does the accuracy of a BCI influence the UX?

We will try to get an answer to this question by doing several studies in all of which we used some kind of (simulated) BCI control in a game and looked at the effect on the UX. In the next section we will give an overview of the thesis.

1.6 OVERVIEW OF THE THESIS

This thesis is about the influence of BCI performance on the user experience. Chapter 2 gives an introduction on Brain-Computer Interfaces and how to design and implement more user friendly BCIs. Various User Experience and Usability aspects are explained to establish a firm fundament for the following chapters. In Chapter 3 we will look at how much control is enough to experience a fun game. In an experiment we modulated the amount of control users had through their keyboard to control a character through a semi-open maze. After every round with a set amount of control they filled in a questionnaire indicating their user experience. The results indicate that generally speaking more control resulted in more fun, however at a certain point this effect plateaus or even turns around. From that point on, performance seems not to be the most important factor anymore and other factors gain a bigger influence on the user's amount of fun experienced in the game. In Chapter 4 we will look at different modalities that influence user experience. Users played a game with two different modalities, imagining a right or left hand movement and actually performing a movement with their left or right hand. While actually performing the movement was easier, required less effort and yielded a higher accuracy, imagining was more challenging and fun to do. This shows us that accuracy is not the only important factor. For Chapter 5 we conducted an experimental study with a game which does not have BCI as its main modality for interaction but as an added modality reflecting the users mental state in the game to control the shape of the avatar. Although the amount of control in this BCI can be quite high in a clinical setting, a large part of the 42 users reported that the BCI control was not very accurate for them. Nevertheless, most of them saw great interest in the mapping of the BCI to the game

and indicated that they had the idea to gain more control over time. Objective measures showed that they did not spend less time playing the game and the difference in the experienced fun while playing the game was also not significant. The results therefore indicate that by replacing part of the interaction by a far from perfect BCI control can gain interest from users and does not impart frustration to such an extent that they want to spend less time playing with it. Chapter 6 contains an experimental study we did with a more qualitative nature. We developed a canvas-based painting application with 3 different modalities, eye blinks, head movement and a brain activity based modality. The main objective was to find out whether it was possible to design an application with more than one novel modality in the most intuitive manner possible. As the results are hard to quantify (there is no clear objective in a creative painting application) we are particularly interested in the subjective assessment of the participants in this chapter. Chapter 7 contains a general discussion on some of the previous chapters, where we saw some further explanation or elaboration on certain topics fit. These added discussions will be about the methods we applied or interpretation of results which came about after the actual publication of the articles in their respective journals. We felt the need to add this chapter, as scientific studies and their respective articles in our opinion never finish: there always is something more to ask, question, discuss or interpret.

1.7 CONTRIBUTING ARTICLES

This chapter is based on the following articles:

- Van de Laar, B.L.A. and Gürkök, H. and Plass-Oude Bos, D. and Poel, M. and Nijholt, A. (2013) Experiencing BCI Control in a Popular Computer Game In: Transactions on Computational Intelligence and AI in Games. Accepted.
- Van de Laar, B.L.A. and Reuderink, B. and Plass-Oude Bos, D. and Poel, M. and Nijholt, A. (2013) How much control is enough? Influence of unreliable input on user experience In: Systems Man and Cybernetics Part B. Accepted.
- Van de Laar, B.L.A. and Plass-Oude Bos, D. and Reuderink, B. and Heylen, D.K.J. (2009) Actual and Imagined Movement in BCI Gaming. In: Proceedings of the International Conference on Artificial Intelligence and Simulation of Behaviour (AISB 2009), 06-09 Apr 2009, Edinburgh, Scotland. SSAISB, Brighton. ISBN 1902956818

- Van de Laar, B.L.A. and Gürkök, H. and Plass-Oude Bos, D. and Nijboer, F. and Nijholt, A. (2012) Brain-Computer Interfaces and User Experience Evaluation. In: Towards Practical Brain-Computer Interfaces: Bridging the Gap from Research to Real-World Applications. Biological and Medical Physics, Biomedical Engineering. Springer Verlag, Berlin Heidelberg, pp. 223-237. ISSN 1618-7210 ISBN 978-3-642-29746-5

BRAIN-COMPUTER INTERFACES AND USER
EXPERIENCE EVALUATION

Bram van de Laar,
Hayrettin Gürkök,
Danny Plass-Oude Bos,
Femke Nijboer,
Anton Nijholt.

Towards Practical Brain-Computer Interfaces: Bridging the Gap
from Research to Real-World Applications. Biological and
Medical Physics, Biomedical Engineering. Springer Verlag,
Berlin Heidelberg, pp. 223-237.

2.1 ABSTRACT

The research on brain-computer interfaces (BCIs) is pushing hard to bring technologies out of the lab, into society and onto the market. The newly developing merge of the field of BCI with human-computer interaction (HCI) is paving the way for new applications such as BCI-controlled games. The evaluation or success of BCI technologies is often based on how accurate the control of a user is over the technology. However, while this is still key to its usability, other factors that influence the user experience (UX) can make or break a technology. In this paper we first review studies that investigated user experience with BCIs. Second, we will discuss how methods from the field of HCI can contribute to the evaluation of BCIs. From experience drawn from two case studies we provide recommendations for evaluating BCIs.

2.2 INTRODUCTION

Brain-computer interfaces (BCIs) aim to provide a reliable control signal for assistive technology for disabled persons. With the merge of the fields of human-computer interaction (HCI) and BCI new applications are being developed for entertainment and education which may be interesting for users with and without disabilities. BCIs will be integrated into existing interactive applications. The aim of such applications is to create positive experiences that enrich our lives rather than only providing reliable control. Recently, it was suggested at several keynote presentations at large BCI conferences that reliability is the most important issue to be addressed to achieve technology transfer to the market and the society. However, perfectly reliable systems are not necessarily usable. Even reliable assistive technologies may get abandoned by users when usability is not warranted [122]. Making interactive systems usable is the core expertise of the field of HCI. The process of designing interactive systems in the field of HCI consists of analysis of requirements, design and implementation of the system and user evaluation. To evaluate such systems, the user experience (UX) needs a more important role in BCI studies. Researchers should not only focus on the reliability of the control signal, so that we can better understand how such a system can satisfy the needs of the user.

At this point we should make clear that the concept of usability is not the same as user experience, although they are related. The most widely accepted model of measuring user-oriented

quality assessment of interactive systems consists of three elements: functionality, usability and the user experience [87]. Functionality is about what one can do with the system, id est, what role does it fulfill? Technical aspects such as performance, maintainability, reliability and durability are important. Usability contains higher level concepts such as satisfaction, efficiency, effectiveness, learnability and usefulness. These can partly result from the functionality but are mainly defined by the interaction of the user with the system. Hence, these concepts cannot be tested without real users. User experience is about what the user feels and experiences using the system. User experience therefore contains higher level concepts as immersion (the user is involved and/or lost track of time), fun, engagement, presence (in case of a game, users experience being ‘in’ the virtual world) et cetera. Even though usability and user experience evaluation is not common in current BCI studies, the user’s experience may influence objective performance measures, such as BCI classifier accuracies, and has a big impact on whether users are actually willing to use a specific system.

In this paper, we review studies that investigate user experience in BCI research and the benefits of including such evaluations. Then, we will argue how the use of various techniques from the field of HCI can be advantageous for evaluating BCIs. In the last part of this paper we will elaborate on some case studies and provide recommendations for evaluating user experience with BCIs.

2.3 CURRENT STATE OF USER EXPERIENCE EVALUATION OF BCI

2.3.1 *User Experience Affects BCI*

User-centered approaches can increase usability and user acceptance, which is why some BCI groups involve users in the design process. They assess user needs, develop user requirements, and evaluate the usability [153, 80, 59, 105]. What is often ignored, however, is the importance of assessing the UX and user acceptance in a structured way *during* or *directly after* interaction with the system. The BCI studies that do include UX evaluations indicate three main reasons: its potential to increase user acceptance, to improve performance of the system, and to increase enjoyment. Each of these are discussed in more detail next.

In a study by Münßinger et al., the mood and motivation of users of a BCI painting application was evaluated using a vi-

sual analogue scale (VAS) [92]. Patients with amyotrophic lateral sclerosis (ALS) were more motivated to train with the application than healthy users. While the healthy users also had other options for creative expression, this BCI application provided a unique opportunity to the paralyzed patients. Several BCI studies suggest a relation between motivation and BCI task performance [78, 95], and small but significant effects have been found [69] using an adapted version of the Current Motivation Questionnaire. This questionnaire assesses the current motivation in learning and performance situations [117]. Similarly, the users' belief of how accurately they can control a BCI has an influence on their actual performance. Barbero and Grosse-Wentrup observed that participants who normally perform around chance level, perform better when they think they are doing better than they actually are (positive bias). Capable participants, however, performed worse when given inaccurate feedback, whether the bias was positive or negative [7].

Motivation may be only one of the performance-related factors that are influenced by the UX. By evaluating and improving the UX, other relations between the user and BCI recognition performance could be exploited to improve performance measures. There could also be mechanisms with indirect influences. For example, a system that is perceived as more beautiful is also perceived as more usable [132]. This perception could influence motivation which in turn could influence performance. Similarly, a more positive experience may cause users to be more indulgent towards minor usability problems, increasing the user acceptance [100].

Most current BCI applications still serve only as a proof of concept [90], which may be why the entertainment value is often not evaluated. An exception is the BCI game BrainBasher, which was evaluated for the influence of different graphical interfaces and different user tasks [104, 134]. The Game Experience Questionnaire was used to assess immersion, tension, competence, flow, negative affect, positive affect, and challenge [62]. In the first study, the UX and performance were determined for a clinical setup with minimal information on the screen. This was compared to a game-like setup of exactly the same task. The game version resulted in higher immersion. The second study compared the UX for imaginary and actual movement. Imagined movement was perceived as more challenging, but when using actual movement the participants stayed more alert.

While more research is still needed, the few studies so far suggest that UX can affect a BCI system in important ways. Therefore it is vital that the UX of BCI systems is properly evaluated.

2.3.2 BCI Affects User Experience

UX can influence the performance of BCIs, but BCIs can affect the UX as well, in two ways: (1) through the effects of using this particular input modality; and (2) by using information about the user's mental state to adapt the interface or the interaction itself, with as goal to improve usability and UX. Here are some examples to illustrate this.

Using BCI for input can in itself influence the UX. Friedman et al. investigated whether the use of imaginary movement to walk in a virtual world would increase the sense of being present there, using the Slater-Usuh-Steed presence questionnaire combined with a non-structured interview [42, 125]. In a follow-up experiment, Groenegress et al. compared the presence experienced with a P300 interface to eye gaze and wand navigation [47]. Both experiments concluded that the BCI did not have a positive influence on presence. In a study by Vilimek and Zander [140], an eye gaze system was augmented with a BCI to simulate the mouse click. The resulting workload of the BCI method was compared to the standard method of using dwell times for activation, using the NASA TLX [54]. There was no significant difference between the workload for either activation method, so the BCI did not result in a higher cognitive demand. A more recent study by Hakvoort et al. compared a BCI selection method with a non-BCI selection method, made equivalent in terms of time and effort necessary for selection [53]. The comparison was based on affect, evaluated with the self-assessment manikin [15], and on immersion, which was determined with the questionnaire developed by Jennett, et al. [66]. In this case, the BCI did turn out to be more immersive and to result in a more positive experience.

With the help of BCI, users can also be supported in the tasks they are trying to accomplish, which in turn should increase user satisfaction. For example, error-related brain activity can be detected and used to fix user or system errors for improved error handling [150]. The amount of information presented on screen can be adjusted according to the user's workload [126]. BCI could also be used to create or maintain specific user experiences. As an example, brain activity indicators of stress or boredom can be used to keep the user in the optimal state of flow, where the challenge of the task is matched to the skill of the user [29, 36].

But the influence of BCI on UX may extend even further. Obink, et al. investigated the influence of using a BCI on social interaction in a cooperative game [101]. The social interaction

was assessed in terms of the amount of speech, number of utterances, and gestures. Additionally, a custom questionnaire at the end of the experiment was provided to evaluate the participants' self-reported, subjective experience. Because of the higher difficulty of the BCI-based selection, compared to point-and-click with a traditional mouse, there were more utterances and empathic gestures (see also section 4).

All in all, whether BCI is used to affect the UX purposefully or whether this happens by simply using this input modality, in both cases it is important to evaluate and be aware of the effects. The next section will show different methods to do this, and discuss the implications for using them to evaluate BCIs specifically.

2.4 APPLYING HCI USER EXPERIENCE EVALUATION TO BCIS

Although evaluating the usability and UX of BCI systems is not common practice, in HCI research and development, especially for entertainment technologies which simply aim to improve the well-being of users, UX is a major concern. Therefore, the HCI community designs for UX and develops methods to evaluate it. Current methods for evaluating UX in entertainment technologies can be classified into quadrants of a plane which has an objective versus subjective axis and a qualitative versus quantitative axis [82] (see Fig. 1). The objective methods are based on overt and covert user responses during interaction while the subjective methods rely on user expressions after the interaction. The quantitative methods employ statistical analysis on collected data whereas the qualitative methods are interpretations of user responses by researchers. Below, we describe the methods corresponding to the quadrants formed by these two axes and discuss their contribution in evaluating BCI systems.

2.4.1 *Observational analysis*

Observational analysis is a qualitative-objective method which relies on overt user response. The classical way of observing overt user behaviour is through audiovisual recorders which provide qualitative data for gestures, facial expressions and verbalisations. There are some difficulties associated with annotating and analysing such rich data though. Firstly, while analysing the data, the researchers should acknowledge their biases, address inter-rater reliability and not read inferences where none are present. Secondly, there is an enormous time commitment associated with observational analysis. The ratio of analysis time

Figure 1: A classification of current user experience evaluation methods used in human-computer interaction for entertainment technologies (adapted from [82]).

to data sequence time ranges from 5:1 to 100:1 [83]. Thirdly, the operation of audiovisual recorders impose restrictions such as a noise-free environment during audio recording or consistent illumination during video capturing. Some restrictions are also imposed by brain activity recording devices. For example, the electroencephalogram (EEG, measuring electrical brain activity) is affected by the user's movement [39], so users are usually asked to keep their bodies and faces motionless. Thus, overt behaviour of users of BCIs will be minimal and observational analysis may not obtain sufficient data to analyze UX. Moreover, severely disabled people, such as patients with locked-in syndrome (LiS) who lose all their muscle control except for vertical eye movements [12] and who constitute a non-negligible user group for BCIs, are not able to show any overt behaviour at all. Consequently, in clinical experiments observational analysis is not a strong method for evaluating UX, although for studies in natural environments it might prove useful.

2.4.2 Neurophysiological measurement

Task performance metrics have been suggested as quantitative-objective measures of UX but these are not necessarily the indicators of UX. Especially in entertainment applications, there might not be a clear task or users might prefer navigating in the virtual environment without any urge to complete tasks. More recently, use of neurophysiological signals was proposed to model the emotional state of users in play technologies [83].

Examples of psychophysiological signals are EEG, galvanic skin response (GSR, measuring skin conductivity) and electrocardiogram (ECG, measuring electrical heart activity). Measured emotions capture usability and playability through metrics relevant to play experience so they provide objective data. They account for user emotion and they are represented continuously over a session. While interacting with a BCI, at least one neurophysiological signal, the EEG, can already be recorded as it is used as an input signal. It is a golden opportunity to extract UX-related features from the brain signals using the same signals. Several problematic issues can be identified when recording psychophysiological signals. First of all, the research on using neurophysiological sensors to measure UX is in its infancy. The neurophysiological correlates of UX or its components are not well-defined which makes this method rather a questionable one. Secondly, the sensors attached to the user might induce discomfort to the user, restrict movements or influence the experience. So, the researchers should limit the number of sensors applied on the user. Thirdly, while measuring the UX through the same neurophysiological sensor that is used for controlling the application, UX-related responses should be differentiated from task-related activity.

2.4.3 *Interviewing and questionnaires*

Interviews and questionnaires provide subjective data for assessing UX. They take place after interacting with a system thus are unobtrusive but then not able to extract instantaneous experiences during interaction. One way to converge capturing short-term UX might be to conduct questionnaires and interviews incrementally, id est, in multiple sessions, rather than conducting a single questionnaire/interview after the interaction has taken place. For disabled users, especially those with LiS, using subjective methods might not seem to be the easiest way to assess UX as these people might not be able to talk or write. However, if the interviews and questionnaires are prepared in such a way that they can be answered using a small number of choices, such as yes, no and maybe, then they can be completed by these users as well.

Interviewing is a qualitative-subjective technique. During interviews, researchers should be careful to pose the right questions during the interview, if necessary, by monitoring the interaction and detecting unexpected events. The interviewers should remain neutral and refrain from asking leading questions. An example demonstrating the use of interviews in BCI UX eval-

uation is the study by Gürkk et al. [52]. In their study, the authors conducted interviews with participants to find out the reasons why people switched between BCI and speech control in a multimodal game.

Questionnaires are designed to provide quantitative-subjective data. Users rate the items in a questionnaire on a Likert-scale or a Visual Analogue scale, which yields a number of how much they agreed with a statement. Development of UX questionnaires for entertainment applications has received attention from researchers, especially those who are interested in games. The recently developed Game Engagement Questionnaire [16] includes items related to absorption, flow, presence and immersion. There are also questionnaires focusing exclusively on the components that contribute to UX such as presence [133] and immersion [66].

2.4.4 *Other methods*

Another concept that is often related to UX is the usability of the interface. Many heuristics have been proposed for evaluating the usability of video games [102]. However heuristic evaluation does not involve actual users and the usability of an interface alone does not represent the UX. Before questionnaires are used to evaluate BCIs, they may require adaptation taking into account that state-of-the-art BCI applications are relatively simple thus modest in providing rich UX. BCI recognition performance should also be taken into account, as a relatively low performance might influence the UX.

Analysing logged software data is also considered as a quantitative-objective method for UX evaluation in some studies. Logs are not direct correlates of UX but they might be helpful in understanding the course of interaction, identifying problems or certain preferences, and thus in designing for better UX. For example, by analysing the frequency of key presses in a game, one can derive a cluster of events to which the player was more reactive and can use this new information to design better interaction.

The important factors in selecting the right UX assessment method for BCIs can be listed as the ease of deployment and analysis for the researcher, the comfort of deployment on the user, the strength and reliability in representing the actual UX, and the width of the user spectrum. As seen within this section, all the methods partially fulfill these criteria. Nevertheless, questionnaires stand as strong candidates as they are easy and comfortable to apply, suitable for extracting statistical analyses quickly,

Figure 2: A screenshot from the game Mind the Sheep! depicting the game world with ten sheep, three dogs and the pen.

strong and reliable when validated and applicable to the majority of the BCI users.

2.5 CASE STUDIES

In this section we will elaborate on two case studies in which we applied various methods of UX evaluation and will try to explain why we chose a certain method and how it answers our research questions.

2.5.1 Case study: Mind the Sheep!

We did a series of UX evaluation studies using the multimodal game we developed, called Mind the Sheep!. The game world (see Figure 2) is a meadow on which a number of (white) sheep move autonomously and the (black) shepherd dogs can be commanded by the player. When a dog approaches some sheep, the sheep will tend to flock and move away from the dog. This way the sheep are herded in a desired direction. The goal of the game is to gather the sheep in a pen as quickly as possible.

The game can be played using different modalities in different ways. In the BCI controlled version of the game, to command a dog, the player positions the cursor at the point to which the dog is supposed to move. The player holds the mouse button pressed to provide the command to select the dog. Meanwhile, the dog images are replaced by circles flickering at different fre-

quencies and the player concentrates on the circle replacing the dog they want to select (so as to obtain an SSVEP). The stimulation persists and EEG data is accumulated as long as the mouse button is held. When the player releases the mouse button, the signal is analysed and a dog is selected based on this analysis. The selected dog immediately moves to the location where the cursor was located at the time of mouse button release.

In the first study we describe here, we compared BCI control to simple mouse control to study the social interaction between players in the cooperative multimodal version of the game [101]. To control the dogs using the mouse, the player first clicks on the dog they want to select and then on the location they want the dog to move to. In the cooperative multiplayer version of the game, co-located players work together to pen the sheep so they need to interact while playing to develop a strategy. However, interaction means such as speech and bodily movements might impair the accuracy of the BCI due to the noise they impose on the EEG. So there is a trade-off between maintaining a strategy during the game and maintaining a certain accuracy level. We did an experiment with 10 pairs playing the game with both controllers. We performed an observational analysis of the audio-visual data recorded during the play. Though non-significantly, during mouse control the participants produced more speech and instrumental gestures, which are overt messaging channels. This implies that they interacted more freely with mouse control. On the other hand, during BCI control participants produced, again non-significantly, more utterances and empathic gestures, which are emotion signalling channels. This finding suggests that the participants were affected more by the events during the BCI game; perhaps they were surprised, or things went wrong more often.

In another study with the single-player version of *Mind the Sheep!*, we evaluated UX in terms of immersion and affect through questionnaires [53]. Again, we compared BCI control to mouse control but this time the way the mouse was used was different. Now, the player had to hold the mouse button pressed when they wanted to make a selection. The dogs were highlighted one at a time with an increasing highlight period. When the player released the mouse button, the currently highlighted dog was selected. To make an accurate selection, the player needs to react in the time when the dog they want to select is highlighted. This way, mouse control becomes similar to BCI control so that they both offer some challenge to the player. In our experiment we let 17 participants play the game with BCI and mouse control and after each game we evaluated UX using the self assessment manikin [15] for affect and the immersion questionnaire

[66]. Evaluation results showed that BCI control was found to be more immersive ($p=0.031$) and positively affective ($p=0.044$) than mouse control. Furthermore, analysis of the logged game data revealed that participants appeared to have more patience with BCI control than mouse control, which could have been caused by the curiosity of participants for BCI control or by their self overestimation during mouse control.

2.5.2 Case study: *Hamster Lab*

The aim of the study was to investigate the effect of level of control on user experience. We conducted an experiment, with 200 participants, in which they played a game with a varying amount of control: the game is called *Hamster Lab* [?]. The user controls a hamster situated in maze like game levels, set in a laboratory setting, see Figure 3. The user controls the hamster by pressing the arrow keys on a keyboard. This way, the user has five possible options for control: up, right, down, left and do not move. The 15 controlled conditions in the game are specified with a certain amount of control. From perfect control, where every press of the button is directly translated to the corresponding action in the game, to 20% control, id est, chance level, where there is an even chance for a certain action to be translated to any of the five possible actions. Using this manipulation of control we can simulate the feeling of unreliable input such as is the case with a BCI. Thus, whereas the relation between control and user experience can only be investigated through correlational analysis in conventional BCI experiments, through this simulation we can study the effects using controlled conditions. *Hamster Lab* is an online game suitable for playing in a web browser. This made it easier to gather participants for the study we will next describe.

This study was conducted to find the relationship between fun and control. After the user commanded the hamster to the exits of the four mazes a short questionnaire was presented to the user. The questionnaire was kept to a bare minimum of what we wanted to know. There were 9 questions in total: 6 were Visual Analogue Scale (VAS) items and 3 questioned for basic demographics (age, gender and a field where user could give feedback. As it was an online game probably not all participants were motivated to answer the number of questions one would normally ask after an experiment. Furthermore, we wanted the participants to play multiple sessions to gather more data. Based on IP addresses 200 unique participants started a round. 351 rounds in total were started. 212 (60,4%) of these

Figure 3: A screenshot from the game Hamster Lab depicting the game world with the hamster in the first level of the game.

runs were continued through the four levels and filled in the questionnaire completely. By most participants the short questionnaire was appreciated (12 people played 5 rounds of 4 levels). Though for 39,6% of the runs, even 9 questions were too much for the participant to bother, as some entries in our database showed with comments like "Why these questions, I want to play!" and items that remained unanswered. This number includes the runs that were started and terminated halfway, for example participants that did not like the game and closed their browser window.

The six VAS items provided us with information on the UX of the participants on the following concepts: fun, frustration, control, dominance and empowerment. First we had to assert that our way of influencing the amount of control was also perceived by the user as such. Regression analysis on the amount of control and the perceived control showed a very significant linear trend, hence, the higher the amount of control 'given' to the user, the higher the amount of control they perceived. For the relationship between fun and control we hypothesized that fun is positively influenced by control (users are able to do the things they intend) but an optimum exists before the maximum amount of control (the game is too easy and/or nothing surprising happens). A regression analysis showed that more variance was explained using a third order polynomial (34,9%) than using a linear (29,1%) trend would. The third order polynomial showed an optimum before the 100% control mark which was

also shown by comparing the medians of the conditions with a high amount of control. This supports our hypothesis and is also in line with the idea that users playing a game want to have some kind of challenge in a game. In this way, the unreliable input from a BCI can be used as a challenge in a game [99]. The result is validated only for this specific game, other games might show a curve looking slightly different but with the same characteristics. Games that are in itself a big challenge for the user might require 100% control. An interesting conclusion from this is that one wants to use a BCI for control, the game difficulty needs to be adjusted to balance the user skills and game challenge for the optimal experience of flow [29].

2.6 DISCUSSION AND CONCLUSIONS

In this paper we stressed the need for UX evaluation of BCI applications. While some research has been done on this, it remains largely an uncultivated area of research. However, we can learn from methods developed in the field of human-computer interaction.

To evaluate a BCI system several methods are available: observational analysis can be used in settings where the interaction of the user(s) with the BCI system as a whole is important. For example in the case study of Mind the Sheep! we showed that observational analysis is a useful method when evaluating systems in a realistic setting especially when users can also interact with each other. When overt user response to the system is limited, *id est*, in case of a clinical experiment, when the user is disabled, or in non observable settings such as for example a web-based experiment, observational analysis is less useful.

Neurophysiological measurements are a quantitative-objective method to assess UX. However, these techniques are still topic of research and most are not very reliable at the present. If a reliable neurophysiological method is used however, this provides a worthwhile source of information as the signal is continuous of nature, as opposed to for example a questionnaire.

Interviews are especially useful in explorative studies. Asking open (non-leading) questions can lead to the reason *why* a user does or does not like a certain aspect of the system or why the users did what they did. This information is hard to capture through other methods, as it is quite detailed in nature.

Questionnaires are quantitative of nature and answers to the questionnaire can easily be quantified to prove effects over groups of participants. This makes it a frequently used method to evaluate systems. Standardized questionnaires exist on various as-

pects of UX. However, if one wants to evaluate a system on all these aspects the user has to fill in hundreds of questions, with the risk of ‘questionnaire fatigue’ (filling in the questionnaire at random, or the same answer for each item) and users choosing for the safe middle option, because at some point, all questions seem to be the same. In the case study Hamster Lab, we showed that when a high number of participants is involved or data is gathered over multiple sessions it is possible to limit the items in the questionnaire to exactly what is needed to answer your research question.

Part II

PART II STUDIES

The three chapters in this part describe the three studies we did to get an answer to our research question. Each chapter starts with a short introduction, related work, methods and a discussion of the results from the study.

HOW MUCH CONTROL IS ENOUGH?
INFLUENCE OF UNRELIABLE INPUT ON USER
EXPERIENCE

Bram van de Laar,
Danny Plass-Oude Bos,
Boris Reuderink,
Mannes Poel,
Anton Nijholt.

IEEE Trans. on Systems, Man, and Cybernetics Part B, 43 (6),
pp. 1584-1592, 2013, ISI Impact 3.236.

3.1 ABSTRACT

Brain-computer interfaces (BCI) provide a valuable new input modality within human-computer interaction systems. However, like other body-based inputs such as gesture or gaze based systems, the system recognition of input commands is still far from perfect. This raises important questions, such as: What level of control should such an interface be able to provide? What is the relationship between actual and perceived control? And in the case of applications for entertainment in which fun is an important part of user experience, should we even aim for the highest level of control, or is the optimum elsewhere? In this paper we evaluate if we can modulate the amount of control and if a game can be fun with less than perfect control. In the experiment users ($n=158$) played a simple game in which a hamster has to be guided to the exit of a maze. The amount of control the user has over the hamster is varied. The variation of control through confusion matrices makes it possible to simulate the experience of using a BCI, while using the traditional keyboard for input. After each session the user completed a short questionnaire on user experience and perceived control. Analysis of the data showed that the perceived control of the user could largely be explained by the amount of control in the respective session. As expected, user frustration decreases with increasing control. Moreover, the results indicate that the relation between fun and control is not linear. Although at lower levels of control fun does increase with improved control, the level of fun drops just before perfect control is reached (with an optimum around 96%). This poses new insights for developers of games who want to incorporate some form of BCI or other modality with unreliable input in their game: for creating a fun game, unreliable input can be used to create a challenge for the user.¹

3.2 INTRODUCTION

Recent developments in interfaces show that there is a need for less artificial means of control. The most prominent examples of the moment are the Nintendo Wii and the Microsoft Kinect, both gesture interfaces. But speech, eye gaze, and other physiological measures are also promising a more intuitive way of interaction. By allowing the user to apply knowledge from previous interactions, for example from interacting with the real world or from interacting with comparable systems, the interface is easy to learn, easy to remember, and easy to use, which

---

¹ For more information and additional discussion, please refer to Chapter 7

are key aspects for usable systems [89, 14, 93]. Brain activity as input modality also has potential in this area, as it can provide some insight in the intention of the user, without depending on external expression. Unfortunately, most of the current systems are still in the phase of proving that using brain activity for control is even possible, and are therefore not making full use of the intuitiveness this input could provide.

3.2.1 *Unreliable input*

One thing physiology-based inputs have in common is that the interpretation of the input is often problematic. This is mainly because of the noisiness and ambiguity of the input, but also because of the problem of intentionality, see [31, 114, 63] for example. This noise and ambiguity will make an unreliable input channel. We try to clarify this through three examples:

3.2.1.1 *Keyboard*

In case of input through a keyboard, the keys which are pressed by a user are always recognized as what was typed in. There is (in case of a wired keyboard) no measurable noise between the keyboard and the computer and no problematic interpretation of the keystrokes. The only possibility for an error to occur is by uncontrolled motor activity of the user. A recognition or classification rate is not applicable in this case.

3.2.1.2 *Vision-based*

In case of a gaze, gesture or other vision-based interface: Interfaces based on gaze, gestures et cetera depend on visually capturing (part of) the user. Either eyes or limbs point in a certain direction or make a certain movement which is classified by the computer and translated into an action within the system. Users can make errors by not accurately carrying out what they intended to do. As the system depends on the captured image, different kinds of artefacts and noise can hinder classification of the correct action. Pixel noise, different lighting, occlusion, movement of the person, movement of the camera, vibrations et cetera all may have a negative impact on the recognition rate. Input can change over time, for example as the user becomes more fatigued and is less expressive, or as the sun sets and the lighting changes. Besides, there can also be a large variability between users, such as between children and elderly, or men and women. As an example of ambiguity: when somebody waves

their hand, it could mean ‘good bye’, ‘hello’, or even ‘no’. A final problem is intentionality (also referred to as the Midas touch problem [64]). Not all actions will have a purposeful intention related to it. What if the user was not waving at the system, but waving to get rid of a mosquito passing by? Or in the case of an eye tracker, the user will already look at the system simply to take in information. In that case, not every eye gaze is meant as an input command. Especially when the system is always on, there will be times when the user is not purposefully interacting with the system.

3.2.1.3 *Brain*

Brain-Computer Interfaces (BCI) are dependent on the brain activity generated by the user performing a task. This can either be the user attending a visual, auditory or tactile stimulus, or by generating brain activity by actively performing a mental task (e.g. imaginary motor movement, performing mental calculus). Errors can occur when the user is not attending the stimulus or not (correctly) performing the mental task. But because of the highly varying nature of the resulting measured signal, the system will make errors in recognizing the users intention to attend/not attend or to perform the mental task or not. Recognition rates vary greatly between users and within users over time, largely independent of the type of BCI [139]. Related to the Midas touch problem, in BCIs it is difficult to detect whether the user is performing a task or not (idling) [152].

In the last two cases, there is an inherent noise (with different causes) on the input channel to the system. This noise makes recognizing the users intentions non-trivial. We therefore define unreliable input as the system not being able to reliably recognize the users intention, caused by inherent noise on the input channel.

3.2.2 *BCI based input*

In the case of brain-computer interfaces, each of these problems can be even more difficult to solve as there is no way to obtain a ground truth. Systems based on visual input can be evaluated using expert assessments on recordings of for example gaze and movements. First, BCIs require robust and noise-free brain activity recordings which is problematic at user’s homes. For the general population, undergoing surgery to get electrodes implanted is also not a viable option, as the surgery, but also long-term implantation of these electrodes, are still too risky.

Finally, there is the issue of response time. If a system is to be used for direct control, the response time should be minimal. This means that systems that depend on indirect measures such as increased blood flow in more active brain areas will be considered too slow for this purpose. What the general home-user then is left with is electroencephalography (EEG). Unfortunately, this measurement is highly sensitive to noise, both from the environment and from the user's body. It has a good temporal resolution, but because it uses electrodes on the outside of the head, it is difficult to only measure what we need to measure. Instead a smeared out, attenuated signal from numerous interfering sources is measured. Even if all this would have been perfect, the brain is a very complex system, and specific areas may activate for different reasons. As an example, certain areas of the brain are involved in making gestures. These areas may also activate, however, when the user is imagining to make that gesture, or when the user is looking at somebody else making that gesture [35], [85]. The problem of intentionality also still remains, as the user's brain may be responding to something that is not at all related to this particular part of the interaction with the system. As a result, the interpretation of such physiological input modalities may never be perfect, and at the moment, brain input seems to be the least perfect of all.

In most studies concerning user control, the input itself is considered to be near perfect, although mistakes may still be made because the user is distracted, unskilled, or is unsure about what to do. The general solutions that are provided to solve problems caused by imperfect control are therefore generally in the range of: make sure that the system is responsive (small delays for feedback or updated system status), that the feedback is easily understood, and include an undo button [93]. There are very little guidelines for interfaces where the control input itself may be a critical issue. How many mistakes can be made before a system becomes unusable or unacceptable? One could argue that it is not the actual control that matters, only the user's perception, but how do these two relate? And do we even need to aim for perfect control? Especially in the case of entertainment applications, some imperfections may add to the challenge of the task up until the point the user gets frustrated. Keeping in mind a not perfect control when designing a game can keep the user in a state of flow while they deal with the problem and learn how to cope with the imperfect control [99]. In this study we let users utilize the keyboard for interaction and simulate unreliable input by manipulating the input from the keyboard and mapping it, with different levels of accuracy, to actions in the game.

3.3 RELATED WORK

In the next section we will outline the related work relevant to this study. We will start with what *control* actually is and how to measure it. Control which is not perfect includes errors. How to deal with errors and how users perceive control concludes this section.

3.3.1 *How to measure control?*

There are many different measures of performance of the recognition by the system. Accuracy is the easiest one to understand, as it is simply the percentage of correct interpretations. The opposite measure is the error rate. Precision (fraction of retrieved instances) and recall (fraction of relevant instances) are the de facto measures in the information retrieval and visual action recognition domain. However, when dealing with an interactive system the interval at which actions can be performed and speed of processing become important as well. Also, as accuracy and error are dependent on the number of classes and the ratio between samples from the different classes, more complex measures should be taken into account.

For example in the case of the interpretation of moving either left or right, a random choice should yield a 50% accuracy on average, but in the case of left, right, forward, and back, the random classifier would only achieve 25%. This means that performance of 75% accuracy in a 2-class system is very different from the same accuracy in a 4-class system.

For an example of acceptable accuracies within the BCI domain for a four class problem: Ware et al. have evaluated the level of acceptable and desirable accuracy in a 4-class steady-state visual evoked potential (SSVEP) based brain-computer interface, with five participants, by incrementally decreasing the accuracy of the interface. They found that accuracy levels of at least 77% were accepted and desired [141]. These results are based on a very limited number of participants, but it may be an initial indication of the level of control a 2 class SSVEP BCI should aim towards.

However, in case of a two class problem, if the user moves to the left 70% of the time, and the system would be a simple classifier that always selects the class with the highest prior probability (the class which has been used most in the past), it could already achieve an accuracy of 70%, where if the classes would be equal, the result would have been 50%. To address this problem, various performance measures have been designed. One of

these is the ROC which displays the ratio between the true positive rate (the fraction of correctly detected target movements) and the false positive rate (detecting a target movement). The related AUC value of the ROC that gives the probability that a target event (e.g. movement to the left) has a higher confidence than a non-target event (all other events) [40]. While the ROC and AUC-ROC give a reliable measure of performance when the prior probabilities of the events are unequal, they are inherently binary, i.e. they measure the performance for discerning only two classes. A performance measure for multiple classes that has been gaining popularity in the BCI field is the ITR, which measures the amount of information expressed in bits that can be communicated through an unreliable channel per unit of time. In this case, the unreliable channel is the BCI, and the user is supposed to use an optimal encoding strategy for its message. Current, non-invasive BCIs have an ITR of up to 10–25 bits per minute [146]. The type of BCIs with the highest ITRs are so called P300 speller systems. These systems make use of the Event Related Potential (ERP) associated to attending a stimulus in the EEG of the user. This ERP starts to show up 300 ms after the (visual) stimulus has been given to the user. By visually highlighting all letters of the alphabet the system can distill what letter the user was attending to [38].

The advantage of measuring performance with the ITR is that, when calculated based on MI [27], it is insensitive to unequal prior event probabilities, and incorporates both the precision of and the time needed to detect an event: The more time you take for a selection, the more data the system can gather about the input you are trying to provide, the higher the resulting interpretation accuracy. But while a higher accuracy will increase the throughput, it will generally take more time which reduces the ITR in turn. Therefore this measure gives a good indication of the trade-off between time and accuracy.

Because the ITR incorporates both the speed of communication, and the amount of information a single event contains, it might measure a quality that is as much related to the ability of the user to express its intent as possible with an objective measure. Similarly, the difficulty index ID in Fitts' law is also a measure of information, measured in bits [41]. The assumption of using an optimal encoding strategy made in the ITR might be difficult to achieve in practise however. Most BCI based spelling applications do not use text prediction, and based on context the optimal predictions might change. But while an optimal encoding *in the decoder* might not always be feasible, some environments might be forgiving enough to let the user exploit alternative control strategies to optimize their information throughput. For

example, an unreliable command to turn left might be replaced with turning right for a longer period when time and space permit. Finally, Furdea et al. proposed the Written Symbol Rate (WSR) for BCI based speller systems. This approach only accounts for correctly written symbols excluding corrections [43].

3.3.2 *Dealing with Errors in BCIs*

In most cases, it is better for the system not to take action on an input than to take the wrong action. For example, in the case of a P300 speller, to correct an incorrect character, one selection is needed to delete it, and another selection to then select the correct one. Now there are dynamic P300 spellers that do not make a selection until the confidence level for a specific option is above a certain threshold [123].

Such systems not only act when the certainty is high enough, but also use repeated inputs to increase the certainty for a specific selection. This last feature is part of many potential-based BCIs, as these features are very sensitive to noise and difficult to detect based on one repeat (single-trial). Systems based on other types of brain activity features could also make use of these same principles, by looking at the confidence levels of the classifier, or by combining successive classifications until a certain threshold is reached.

Quek et al. [115] developed a simulation tool for control in applications using a BCI but without the need of an actual participant and an EEG cap connected to the system. Various factors, such as error rate, non stationarities due to a changing state of mind, noise and delay are incorporated in their models. However useful this study is, for our study this approach was unusable for a couple of reasons. The study by Quek et al. only simulates two classes (i.e. left and right), needs EEG data from users to generate a model and would therefore be too time consuming and from a logistics point of view impossible for our intended experiment.

3.3.3 *Perception of BCI Control*

Various studies have been carried out not only focussing on accuracy but also taking into account how users perceived exerting control through a BCI. In a comparison of user experience between actual and imaginary movement control of a BCI game, users had better control with actual movement, which also resulted in higher alertness. Imaginary movement was perceived

as more challenging [134]. A more longitudinal study was done in which 14 participants played a BCI game repeatedly over a period of five weeks. They used three different mental task pairs during each session. The results indicate that the user preference for certain mental tasks is primarily based on the correct recognition of the tasks by the system, and secondly on the ease of task execution [112]. To further improve possible recognition rates Zander and Jatzev developed and demonstrated an application detecting whether the errors occur on the user or system side. This could allow for new ways of dealing with errors, to improve the usability of any system [149]. Control and dealing with errors are important constructs in evaluating input methods that are unreliable. When evaluating games or virtual environments on user experience, which can be presence in virtual worlds [145] or user experience of BCIs [135] one has to not only look at the absolute performance, but also at the qualitative aspects of the input channel, such as the learning curve, obtrusiveness and intuitiveness. Combining improved recognition rates, user centered design, intelligently handling the inevitable errors and proper user experience evaluation will all help generating a better user experience for BCIs.

3.3.4 *Perception of Control*

From psychology, it is known that people tend to overestimate their ability to control events; this effect is called the illusion of control. In a laboratory study, participants had a varying control over a pair of lights. Even when their actual control was none at all, the participants indicated they had some level of control (the illusion of control) [5]. Langer demonstrated that this illusion of control is stronger when certain factors are present, such as competition, individual choice, familiarity with the action or elements part of the action, and level of involvement [74]. Thompson et al. propose that this perception is already created simply from the intention to create a particular outcome (turn on the left light), and a possible connection between the action executed (press a button) and the outcome (left or right light on). If the expected outcome is positive, people tend to overestimate their level of control, whereas if the outcome is negative, people tend to downplay their amount of control [130]. Nijboer et al. showed that mastery confidence was related to BCI performance. However, this was shown only in 2 out of 4 participants [94].

3.4 METHODS

The goal of this study is to investigate whether we can modulate control in a game and what relation exists between control and fun.

Hypothesis 1: If we can construct a method that resemble the various accuracy levels in the same way one experiences when controlling a BCI and this is validated by the perceived control of participants, then we have a proper way of modulating control.

Hypothesis 2: If the modulation of control is related to the reported values for fun, then we can further investigate what the relation is between these two and whether an optimum or point of no improvement exists.

We need to look at various conditions with different amounts of accuracy to evaluate the experience of control. This requires a lot of data points, not attainable by doing experiments in a lab. We therefore set up an experiment which can be done by using a web browser. The experiment consists of a Flash® based game developed by a professional game studio. We varied the amount of control users had in the game and evaluated their user experience by administering a questionnaire after finishing the game. In the following sections we will elaborate on the design of the game, how we varied the amount of control and what data was used for analysis.

3.4.1 Experimental Design

The game consists of four levels, increasing in difficulty. Four levels in consecutive order make up one round, one or more rounds (usually five) make up one run. See Figure 4 for an overview. Participants started every round in a randomly assigned condition. In total there were 15 different conditions, each with a different amount of control (see Section 3.4.3). Before starting each round users were given instructions on the control of the game, what they were supposed to do to finish the game as well as how to complete the questionnaire. After finishing a round, the in-game questionnaire was administered and a new condition randomly chosen for the next round. Participants in this experiment finished at least one round, and there was no maximum number of rounds. This means participants could opt to quit the experiment after every round. Participants were however encouraged to play for five runs. Five runs would amount to an average experiment duration of 25 minutes including breaks.

Figure 4: Runs, rounds and levels.

Figure 5: A screenshot of the game that was used for the experiment.

3.4.2 Game

In the game used for the experiment, participants controlled a hamster by pressing the four arrow keys on the keyboard. The game setting is an evil laboratory where experiments on hamsters are carried out on computer-brain interfacing with the goal to control hamsters. Users can take control over one hamster to lead it to freedom. A screenshot of the game can be seen in Figure 5.

The amount of control of the user over the hamster varies randomly across rounds (see section 3.4.3). One round in the game consists of four sequential levels, shaped in the context of the evil laboratory. Respectively a cage, a labyrinth, an office room and the block where the laboratory is situated have to be escaped from. Each level is a maze with dead ends and some occasional obstacles. Touching the obstacles causes the player to ‘die’ and to be transported back to where they started in that level. When the user finishes a level, immediately the next level is presented. After the last level a questionnaire was presented (see section 3.4.5). After completing the questionnaire, the users could play another round, most probably with a different level of control. After five rounds the total amount of time they took to rescue five hamsters was recorded and compared to the times

that were already in the database to provide the user with a rank. This was supposed to motivate the participants to play again. If a certain round was too hard, the user had the opportunity to press a button 'skip round' to skip the current round and go directly to the questionnaire. After the start of a level, it took one minute for this button to become active.

3.4.3 Levels of control

The level of control in the game was manipulated by using specific control schemes for each of the conditions consisting of different levels of control.

The events that are used to control the game consist of directives to move in one of four directions, or not to move at all, resulting in *five* possible events at each evaluation of the game loop. When the user has perfect control — which is usually the case with button-based input — each directive of the user is directly translated in the corresponding action. With unreliable controllers, such as a BCI, it is possible that a different, unintended action is performed. For each directive, action pair, the probability of making this mistake can be denoted, which results in a so-called confusion matrix. The behaviour of an imperfect BCI with discrete output can be fully described by this confusion matrix.

The 15 conditions in the game are specified with a confusion matrix. We have chosen to start with the simple assumption that each class (four directives for movement, or no action) is detected correctly with the same probability (accuracy), and that all mistakes are equally likely:

$$C_a = \begin{bmatrix} a & e & e & e & e \\ e & a & e & e & e \\ e & e & a & e & e \\ e & e & e & a & e \\ e & e & e & e & a \end{bmatrix}, \quad (1)$$

where a is the percentage of correctly detected events, and $e = \frac{1-a}{5-1}$ is the rate for a specific confusion. Rows of C_a correspond to a specific directive (the ground truth) and sum to one, the columns correspond with a specific detection.

The specific accuracies are chosen such that the MI of the confusion matrices C^a is distributed evenly with 15 points over the whole possible range. Given the rate of events, the ITR can be

Figure 6: The relation between the MI and the accuracy for a five-class confusion matrix with the same probability for all correctly detected events, and the same probability of each pair of mistaking one event for another.

calculated from the MI. The accuracies of the different throughputs with regular increases is displayed in Figure 6. The logarithmic relation between the accuracy and MI results in relatively few conditions with low accuracies. Given an information (I) of 0.0, the resulting accuracy will be 20% (see Figure 6) which equals a random distribution over five classes.

The MI in bits is calculated as follows, with a discrete variable Y for the different directives, and discrete variable X for the different detections:

$$I(X; Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log_2 \frac{p(x, y)}{p_1(x) p_2(y)}, \quad (2)$$

where $p(x, y) = C_{x,y}^a$ is the joint probability distribution function of X and Y, and $p_1(x)$ and $p_2(y)$ are the marginal probability distribution functions of X and Y respectively. Note that the marginal probability distributions are assumed to be uniform. This is the same assumption used in [146] for ITR calculation.

In summary, we constructed 15 levels of control based on the equally spaced MI of control schemes with equal probabilities of correct detection, and equal probabilities for each confusion.

3.4.4 *Participants*

Participants to the experiments were invited to play the game through various means (e-mail, social media and mouth-to-mouth). Based on IP addresses 200 unique participants started a round. 351 rounds in total were started. 212 (60.4%) of these runs were continued through the four levels and properly filled in the questionnaire. The average age of all participants was 24.99 years ($SD = 7.76$). 8 (3.8%) participants refused or gave a bogus answer on the age question. 52.5% of the participants indicated male and 47.5% of the participants indicated female on the gender question. Only 45.8% of all participants replied to this question. 12 participants played the game for the maximum of 5 rounds and got a high score ranking in the game. The fastest to finish did this in a time of 21 minutes and 42 seconds, the slowest in 35 minutes and 3 seconds.

3.4.5 *Questionnaire*

The questionnaire was presented to the user within the flash game, before they could continue with another run in the game. The questionnaire included three open questions and six Visual Analogue Scale (VAS) items. Two open questions were included to gather basic demographics, namely age and gender. The third open answer box was for general remarks and additions.

The six VAS items measured the amount of fun, engagement and control the users experienced in the game. Table 1 shows the items for each construct. The way users answered the VAS was through a graphical slider. Of course, the experiment being carried out by means of a computer this VAS was digital but approaching the analogue domain with values ranging from 0 to 100. One click on the scale put the slider on the designated point, also indicated by the changing number right below the scale. Users could correct their answers until they clicked on the 'next' button, to go to the next page with three items.

3.5 RESULTS

In this section we will analyse the results from the questionnaire data. First we will report the data on the perceived control, to assert that our method of varying the amount of control is also experienced by the user in the correct way. In the section thereafter we will analyse the relationship between the amount of

Table 1: Constructs and items

Construct	Items
Control	1. I had the feeling the hamster did what I wanted it to do 2. I had the feeling the computer was following my commands
Frustration	3. I was frustrated while playing the game
Fun	4. Playing the game was fun
Empowerment	5. During the game: I felt empowered
Satisfaction	6. During the game: I felt satisfied

control induced by the confusion matrices on the amount of fun the user experiences.

3.5.1 *Perceived control*

First we cleaned up the data from obviously erroneous responses. If one of the following criteria were met the response would be removed from our dataset:

- A. Partly filled in questionnaires: if only the first or the first and second page with questions was filled in and not the subsequent page(s), the complete response was removed from the dataset.
- B. Meaningless questionnaires: if all questions were answered with the initialisation value (the VAS scale was initialised in the middle with a pointer), the complete response was removed from the dataset.

In section III.B we mentioned that participants had the ability to skip a round after one minute. We included these records as it still provides us with useful data. After the cleanup process we got 158 data points to base our analysis on. Looking at the cleaned questionnaire data we first describe the distribution of data points over the conditions. We expected an even distribution of data points over the conditions as these were randomly initialized. The minimum amount of data points per condition was 7 for the lower half of the accuracy scale and 10 for the higher part of the scale. A complete overview of the number of data points per condition can be seen in Table 2.

Table 2: Number of samples per condition.

Samples per condition.			
MUI	#	MUI	#
0	9	1.35	13
0.17	7	1.5	14
0.34	13	1.66	10
0.53	8	1.84	14
0.71	9	2.05	11
0.86	11	2.17	10
1.03	9	2.32	10
1.22	10		

We then constructed the scale ‘perceived control’. This scale was made up out of two items: “I had the feeling the hamster did what I wanted it to do.” and “I had the feeling the computer was following my commands.”. A reliability analysis of the proposed scale resulted in a Cronbach’s Alpha of 0.885 which made this scale a proper measurement of how much control the participants experienced in the game. As the range for each VAS is 0–100 and the scale is the addition of the separate item scores the range for the scale will be 0–200, the medians of these values are plotted in Figure 7. To validate if we indeed varied the amount of control as we intended a linear regression was carried out. This revealed a significant linear trend explaining 50% of the variance in the data ($R^2 = 0.499, p < 0.001$) as can be seen in Figure 7.

We analyse another item, frustration, in the same manner and this further validates our method of influencing control. Linear regression showed a significant negative trend and explained 13.3% of the variance ($R^2 = 0.133, p < 0.001$), shown in Figure 8.

3.5.2 Fun - Control

To validate our most important hypothesis, control positively influences fun, up to a certain point where fun is decreasing because the challenge is becoming less, we analyse the questionnaire data for fun and compare it to the amount of control. To assess whether a linear model or another model would explain our data better, we first performed a linear regression over all data points. This showed a significant trend ($R^2 = 0.291, p <$

Figure 7: The median values of Perceived control are plotted for all conditions (varying amounts of Control(MUI)) in the experiment. A linear regression through all data is plotted as well ($R^2 = 0.499$).

Figure 8: The median values of Frustration are plotted for all conditions (varying amounts of Control(MUI)). A linear regression through all data is plotted as well ($R^2 = 0.133$).

Models for curve fitting			
Model	R^2	AIC	delta(AIC)
1th order (linear)	0.291	977.683	n/a
2nd order (quadratic)	0.336	969.434	-8.249
3rd order (cubic)	0.349	968.305	-1.129
4th order	0.356	968.170	-0.135

Table 3: 1st through 4th order models with respective R^2 , AIC and the change in AIC from the lower order polynomial.

Figure 9: Medians of Fun vs. Control and the third order curvefit through all data points.

0.001) through the data. Second, third and higher order polynomials were also tested. The third order polynomial, which can be seen alongside the first order (linear) polynomial in Figure 9, proved to provide the best fit with an explained variance in the data of 34.9% ($R^2 = 0.349$, $p < 0.001$) taking into account that higher order polynomials yield a slightly higher R^2 but also require sacrificing another degree of freedom for every incrementation of the order. Another measure for this which takes the added complexity of higher order polynomials into account is the Akaike's Information Criterion (AIC) [4]. Less information loss is resembled in a lower AIC. While the difference in the AIC for the second and third order polynomial is substantial, the difference between the third and fourth order is only marginal compared to the former.

The third order polynomial as well as the medians show a clear downwards tendency from the linear trend after the $\text{mui}=2.05/\text{acc}=96.73\%$ point. This supports our hypothesis that control is needed for fun, upto a certain point, after which fun is not dependent on control anymore and other aspects of the UX start to play a role. Another interesting fact is that the condition in which movement is completely random (20% accuracy, $\text{mui}=0$) participants apparently found it relatively fun to play the game. When participants do have some amount of control e.g. the ($\text{mui}=0.17/41.22\%$ accuracy)-condition this effect is gone.

3.6 DISCUSSION

This study showed that it is possible to influence the amount of perceived control the user has in a game that has a 4 class, 2 dimensional control of navigation. This is supported by the fact that users reporting the amount of control they experienced showed a strong relation to the amount of control they were given. We hypothesized an illusion of control; an overestimation of control on the near perfect side of control. Our data did however not support this hypothesis. We also did not find support for an underestimation of control at the low side. A possible explanation for this might be that in our study, users are *given* a certain amount of control defined by the confusion matrix in a session. This given amount of control is something they cannot alter by increasing their effort. Whereas with a BCI user may alter their amount of control, to a certain extent by the effort they put in. In this case the effort put in by the user is highly dependent on the motivation they have, which may be related to the performance of a P300-based BCI [68].

The second part of our analysis showed that the amount of control largely explains the amount of fun one experiences. Although fun is dependent upon the amount of control, at a certain point an optimum is reached. Our analysis showed that in our experiment after 96.73% of accuracy the fun decreases. This could be explained by the concept of flow [29], where challenge is related to skills in which an optimum exists where a state of flow is achieved. In this state of flow the skills of the user and the challenge asked from the user are both high and in balance. If the skill increases, the user would shift into boredom, if on the other hand the challenge increases, the user shifts into anxiety or frustration. This is related to the results we see in our data. Users have a certain skill of controlling the hamster and have a varying amount of control (the challenge), up to a certain point this challenge is more suited to their skill, until the optimum is reached. After this the fun decreases and general user expe-

rience shifts into boredom. This explanation is also supported by some user comments, for example: *"This is the third time I played this game, and the hamster listened quite well. Especially if your first hamster were never obeyed anything (sic), a well-listening hamster is almost boring."* This mechanism of shifting towards boredom could therefore be the reason the amount of fun participants experienced is decreasing after the optimum. At the other end of the scale, the concept still holds: users reported being frustrated through the respective item in the questionnaire, as well as through the open question while playing the game with a low amount of control, for example: *"I haven't quite forgiven hamsters after the last game even though this one was better."* and *"Frustration thy name is hamster. You know he was perfectly happy in his cage..."*. This is also what is often seen in BCIs, if the recognition accuracy is too low, the BCI is just a frustration to the user. It might give the user false hope which leads to frustration.

Like in almost all of HCI research, the results in this study, especially the point for the highest amount of fun, are based on just this game we used to experiment with. Using another game would probably yield another optimum at a different amount of control. Probably even using other level layouts would alter this result. However, we show that this effect is apparent in this particular setting. The game mechanics also included obstacle-like enemies in the harder levels, which upon collision would send you back to the beginning of the level (resetting one's progress). This means our simulation also includes the effects of making an error and users experiencing the cost of an error when using a real BCI. Without the possibility to make errors that have such a high cost, the challenge would probably be smaller as one will eventually get to the end of the level. Trying to anticipate on erroneous movements (and steering clear of the obstacles) is part of the challenge in the last two levels. A smaller challenge with the same amount of control could decrease frustration for the low accuracy conditions and increase boredom (and decrease fun) for the high accuracy conditions. We think that therefore the optimum could shift to the left in such a game.

Still, the game can be quite challenging, even with perfect control, when there is a motivation to finish as quickly as possible. In games that are simpler, for example, that only use 2 classes for navigation, or are less challenging in game play, the optimum probably shifts to the lower end of the control scale. Games with some kind of BCI control often only have one dimension of control, so the previous is especially applicable to these kind of games.

Our results show that in the lower levels of control, the accuracy has an important influence on the user experience. And while a higher accuracy leads to a better user experience in the lower parts, the better the accuracy gets, the less influence it is going to have on the user experience. In this particular game we estimated this critical point to be around 96.73% (although around 86% one can already see the effect in the data). In comparison to state of the art BCIs this is a very high accuracy. There are however some trade offs that have to be made in designing and implementing a BCI. The trade off between speed and accuracy is an important issue for this discussion. The longer the BCI gets to measure the EEG of the respective mental task, the higher the chance it will make the right decision. However, this sacrifices speed. Another common trade off is the number of classes used in a BCI. While we tried to simulate a 5 class (4 directions + 'No operation') often BCIs are still implemented with a 2 or 3 class operation which results in a less complex game with either a one-dimensional control or a timed sequential control in which a direction to walk towards has to be chosen based on a timer. In the first case the accuracy of classified results will be higher than in a 4/5 class interaction but the richness of the interaction will be lower. This could mean that the point where either an optimum is reached or plateauing sets in will be lower than the point we found in our study. Therefore it might not be worthwhile to strive for the highest possible accuracy at the cost of other factors which might have a bigger influence on the user experience. In the second case, the richness of interaction might be higher because of more possible actions, but the sequential selection of commands could increase the latency of the interaction to a point where it might be just a frustration for the user.

Overall, in our results the exact point at which added accuracy does not improve the user experience anymore is not that important, however, the finding that there is such a point at all, is. We think these findings could be an advocacy for not only optimizing for accuracy but taking other factors into account as well. Even outside the field of BCI this finding could be useful. Essentially in all modes of interaction where uncertainty is a factor, like physiological computing or gesture based interfaces a trade-off exists between system response time and accuracy. Although recognition accuracies might be higher or lower and interaction options might be more diverse than in a typical BCI, the effect we found might be present in these interaction methods as well.

Until now, no perfect BCI exists. However, according to our findings a perfect BCI controlled game might not be a system which is only optimized for accuracy. At some point accuracy does

not influence the fun anymore and other factors start to become more important. Therefore, turning the shortcomings of a BCI into a challenge for the user, a challenge outside of the game itself, might be a possible way to create a fun game.

EVALUATING USER EXPERIENCE OF ACTUAL
AND IMAGINED MOVEMENT IN BCI GAMING

Bram van de Laar,
Danny Plass-Oude Bos,
Boris Reuderink,
Dirk Heylen.

International Journal of Games and Computer Mediated
Simulations. 2 (4), pp. 33-47, 2010
Parts of this have been published in Proceedings of the
International Conference on Artificial Intelligence and
Simulation of Behaviour (AISB 2009).

4.1 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 paper describes research conducted to find out what the differences are between using actual and imagined movement as modalities in a BCI game. Results show that there are significant differences in user experience and that actual movement is a more robust way of communicating through a BCI.¹

4.2 INTRODUCTION

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 electroencephalogram (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 [97].

Games are usually the first applications to adopt new paradigms, driven by the gamers' continuing search for novelty and challenges [98]. 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 [46, 77]. 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 [103]. This was done by comparing keyboard control with BCI control for a simple game called BrainBasher, and evaluating the user experience with the Game Experience Questionnaire [61]. Both this game and this questionnaire were also used for the research described here, comparing actual and imagined movement.

---

¹ For more information and additional discussion, please refer to Chapter 7

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 the phenomenon that some people are unable to use a BCI at all [120]. 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 [75].

While most research into using movement for BCI has focused on imagined movement, some clinical research shows that actual movement in fact elicits a more pronounced and therefore better discernable signal in the motor cortex [86].

Actual movement can also be used with other interfaces than a BCI. Interfaces such as a motion tracking system, for example, which is probably more reliable at this moment. One big potential advantage of a BCI however is that the measured brain signals are always preceding actual muscle activity at the limbs, and can be measured before the muscles activate. This advantage is amplified by the onset of a potential in preparation of a movement, the so called *Bereitschaftspotential*, or *Readiness Potential* [71, 124]. A very useful aspect of the RP is the lateralized readiness potential (LRP), where the preparation of left-sided movement is reflected in a potential occurring at the right motor cortex, and vice-versa. Krauledat et al. show that this lateralized readiness potential can be used to classify actual movement even before the movement itself is carried out [72]. This could give a gamer an advantage over other interfaces especially in fast paced, highly reactive games.

But using BCI can also provide other benefits. While measuring brain activity for detection of movement, whether actual or imagined, other information can be derived from the brain as well, such as the user's mental and emotional state. This could be used to make smarter applications which are more aware of the user.

4.3 RELATED WORK

A few BCI games based on imagined or actual movement do already exist. [108] designed a first-person shooter game in which the user could move using the keyboard, and turn by imagined movement. Players learned to control the BCI by experimenting; no instructions were given beforehand. Other examples include

the virtual environments of [79], the board game of [67], and the game BrainBasher [103] that we used in this study.

Both actual and imagined movement can be used for BCIs. Obviously, actual movement is a more natural and intuitive way for users to communicate and neurological process known as Event-Related Desynchronization (ERD) express themselves. All these games involve movement tasks, and are based on a [106]. 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.

Actual movement is characterized by a more pronounced and reliable signal in the motor cortex [86]. This more pronounced signal is a very welcome advantage in the world of BCI where every extra percent of accuracy is appreciated.

When looking at the success of the Nintendo Wii, it becomes clear that actual movement is well enjoyed by gamers ². 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 [65]. 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.

4.4 METHODS

The main question in this study is whether there are differences between imagined and actual movement in a BCI gaming environment. Some of the differences that will be looked into are the gaming experience for the user and the detectability of the signal from the EEG. We also looked at the generalizability of these BCI modalities for different user groups based on demographical characteristics.

4.4.1 Experiment Setup

To answer these questions an experiment has been carried out in which users communicate with the BCI game BrainBasher

² "Nintendo wiining the console war", December 2008. <http://www.igizmo.co.uk/articles/news/744-gaming-nintendo-wiining-console-war>

[103] 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 experiment design see Figure 10.

For the system to be able to recognize the user's actions, a training session is required to create a user-specific classifier. 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 Game Experience Questionnaire (GEQ) developed at the Game Experience Lab in Eindhoven [61]. 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.

The experiment is set up as a randomized cross-over experiment to eliminate sequence and learning effects induced by the succession of both tasks [21]. The subjects are randomly assigned to either group A or group B, taking care that both groups in the end have an equal number of subjects. Both groups perform the same tasks ('receive the same treatment') but in a different order. In Figure 10 this crossover design can be clearly seen halfway the diagram. After all experiments are done we compare the results of the actual movement sessions versus the results of the imagined movement game sessions. The new questionnaire has also been evaluated so it can be used for assessments of other BCI games and other BCI paradigms.

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

4.4.2 *BrainBasher*

The BCI game used for this research is *BrainBasher* [103]. 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

Figure 10: Overview of the experiment setup

Figure 11: The symbols for left and right hand movement.

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 session in which stimuli (in the form of symbols denoting the user actions, see Figure 11) 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 artifacts 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 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 12)

As already explained, the system is trained on the ERD which is generated by planning, imagining and executing motor functions. The activity comes from the motor cortex in which the movement for every limb is prepared. The motor cortex is the most backward positioned part of the frontal lobe, stretching in a narrow band from the ears to the top of the scalp.

Figure 12: A game session.

Figure 13: BrainBasher System View

Figure 14: Photograph of a subject playing BrainBasher.

We used a 32 electrode setup to measure the ERDs. These electrodes are fitted into the small plastic holes of an EEG cap. The small holes are filled with electrolytic gel to provide a proper electrical (low inductance) connection to the scalp on which the potentials can be measured. The cap has to be set up by someone else than the user, because all the plastic holes have to be filled with gel by hand. Although the part of the brain the ERDs can be measured is quite small, we used the full set of 32 electrodes to use spatial filtering and to eliminate sources of noise outside the head. Setting up the system will take somewhere between 10 and 15 minutes. In the Figure 14 the system with the cap, the electrodes and the game BrainBasher can be seen in action.

4.4.3 The BrainBasher BCI

A schematic view of the total system is shown in Figure 13. 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, sampled at 256Hz. For training the system, examples of the ERD for both the left hand and right hand are used to train a classifier to be used during the online game session. The EEG data processing pipeline is shown in Figure 17. First the raw data is re-referenced to the common average reference (CAR) by subtracting the mean of all channels to remove far away sources of noise [81]. After re-referencing a bandpass-filter isolates the frequency range (8-30Hz) in which the ERD occurs. Then we train spatial filters (spatial filters operate in the space domain, as opposed to filtering in the frequency domain) to suppress irrelevant signals with the common spatial patterns (CSP) algorithm [70] to isolate activity on the motor cortices. Now that

we have filtered the signal both in the time domain and in the spatial domain, we calculate the energy of the signal by taking its variance. A Linear Discriminant Analysis (LDA) is applied to make a final prediction based on these band power features. After training the BCI generates a new prediction every quarter of a second, based on the real-time EEG data. These predictions are used to play the game.

4.4.4 BCI performance evaluation

After the training session, we evaluate the accuracy of the classifier by a cross-validation procedure. It is to be expected that a classifier has to have a certain minimum performance for pleasurable experience. Another frequently used performance measure for BCIs is the information transfer rates (ITR), representing communicated information measured in bits per minute. The ITR more directly expresses the utility of a BCI for a user that intends to use the BCI to communicate, and is preferable to other measures because it combines both the quality and the speed of the BCI into a single performance measure. Current BCI are reported to obtain ITRs of up to 10–25 bits per minute [146].

The ITR is estimated by calculating how much information the predictions of the BCI give on the intentions of the users (the ground truth), know as the mutual information I :

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log_2 \frac{p(x,y)}{p_1(x) p_2(y)} \quad (3)$$

where X and Y are the (discrete) predictions and the ground truth respectively, $p(x,y)$ is the joint probability distribution function of X and Y , and $p_1(x)$ and $p_2(y)$ are the marginal probability distribution functions of X and Y respectively. Mutual information measures the decrease in uncertainty in a signal Y (the desired action) given signal X (a derivation of the EEG).

The duration of a trial is needed to calculate the ITR in bits per minute from the mutual information as follows:

$$ITR = 60 \times \frac{I(X;Y)}{\Delta t} \quad (4)$$

where Δt is the time in seconds between one trial and the next, assuming their class labels are independent of each other.

Figure 15: BrainBasher BCI Pipeline

4.4.5 Questionnaire design

To evaluate the user experience a questionnaire based on the GEQ [61] is developed. 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. Questions that are not applicable, (e.g. the questions about the storyline, the complexity and the flow of the game) were left out. On the other hand we added questions, specifically on the amount of control the user experiences. The amount of control is a trivial aspect when using mouse and keyboard. These are reliable ways of communicating with the computer compared to a BCI system which is not so reliable. Also items about user concentration and alertness were added. This is an important aspect because users will have to concentrate to use a BCI game and possibly get tired more quickly than normal.

The questionnaire consists of statements to which users can respond on a 5 point Likert-scale ranging from 'completely disagree' to 'completely agree'. Some examples of items in the questionnaire might be: "I liked playing the game.", "I felt the computer recognized my actions." or "I'm exhausted."

To analyse the results of the questionnaire, we will use Cronbach's Alpha [28]. 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 the 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. The goal for every scale is a value for Alpha larger than the commonly accepted threshold of 0.7 [26].

4.5 RESULTS

First we describe the demographics of the test subject pool, then we analyse 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 10 (50%) were male and 10 (50%) were female. 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. 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 were dominantly left-handed 15 (75%) were right-handed. 14 (70%) received an education higher than average. Computer usage and game experience varied a lot among participants: 8 (40%) participants used a PC for more than six hours a day, 5 (25%) used a computer on a less than daily basis. The same variance goes for game experience: 2 (10%) played games two hours a day, 8 (40%) on a weekly basis, 6 (30%) on a monthly basis and 4 (20%) never played a video game.

QUESTIONNAIRE CONSTRUCTION All participants filled in the questionnaire after both tasks without missing any questions. The responses on the same items for both movement tasks were stored in the same respective variables for scale analyses and in separate variables to analyse the differences in user experience between both tasks. Scale reliability 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. Correcting the scales for items that did not contribute to the scales consistency, e.g. deleting items with a low or negative Inter-Item Correlation, Standardised Alpha's ranged from 0.620 to 0.865 and all scales consisted of at least three items.

To evaluate the usefulness and dimensionality of the resulting scales, a factor analysis was done on all scales separately. Factor analysis is a multivariate analysis technique which can be used to analyse data with "large sets of correlated variables [..]" and as such can be usefull "as a means of examining and describing the internal structure of the covariance and correlation matrices involved". The method used in the factor analysis called Principal Components Analysis (PCA) yields one or more linear combinations of variables in which the maximum explained vari-

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 4: Constructed Scales including alpha and variance explained by 1st principal component.

ance in the data is attained. If one is looking for unidimensionality, then the first extracted component, id est, the one which explains the most variance, should explain by far the most variance in the data.

Looking at our data the first principal components extracted by the factor analyses of every scale explained more than 56% of the variance in the data, except for the Negative Experiences scale. Scree plots [24] also indicated strong unidimensionality across all scales except for the Negative Experiences scale, which turned out to be a two dimensional scale. 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 the cutoff for the number of dimensions. This gave a confirmation for the earlier findings. An obvious knee shape could be seen in all plots except for the Negative Experiences scale.

The corrected questionnaire consisted of 32 items divided over 8 scales. An overview of all corrected scales with their respective Alpha's and variance explained by the first dimension in factor analysis can be found in Table 1. The variance explained by the first factor measures to what extent a scale is measuring only one underlying attribute or construct.

DIFFERENCES IN USER EXPERIENCE The final corrected scales were used to compare the user experience for users performing

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 5: Paired *t*-Tests Scales, comparing imagined and actual movement

both kinds of movements. A direct comparison by means of paired *t*-tests was done. The results of these test can be seen in Table 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. 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 imagined movement. The same trend is also shown in the Fatigue scale, while there is no 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.

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 and one for imagined movement. 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 ($W_+(20) = 48, p = 0.0328$). Looking at performance across

ITR for all subjects					
Subject	AM	IM	Subject	AM	IM
A1	3.8	4.2	B1	2.0	4.1
A2	8.7	2.3	B2	2.6	2.1
A3	2.4	2.0	B3	3.8	2.1
A4	4.5	3.8	B4	4.8	10.5
A5	6.5	3.8	B5	6.0	2.3
A6	4.1	1.8	B6	2.6	4.2
A7	2.3	2.3	B7	2.7	5.3
A8	3.3	3.9	B8	2.3	2.3
A9	x	2.4	B9	5.1	1.7
A10	x	x	B10	5.7	2.6

Table 6: ITR for actual and imagined movement for all subjects (in bit-s/minute), missing data is marked with x.

different 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 didn't 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).

The Information Transfer Rate or ITR (see Section 4.4.4) is another performance measure we calculated for every subject. This makes our results more comparable with other BCI's as well it is more informative than purely an accuracy rate. Note that the ITRs are calculated just for the training data and that therefor the maximum achievable ITR is 15 bits per minute. In the actual game session the bitrate can be higher, for the best subjects probably over 30 bits per minute. It is however difficult to give accurate numbers for this prediction because our BCI is self paced without a definite time window.

The calculated ITRs can be seen in Table 6. (Due to an unfortunate loss of data we were not able to calculate the ITR for all subjects.)

These findings support the accuracy measure in that actual movement provides a more usable signal on average.

4.6 DISCUSSION

Results from this study 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, we developed a questionnaire for evaluating BCI games. 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 [98].

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 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. This problem of not being able to be *understood* by a BCI is referred to as BCI illiteracy [120].

One could argue that perhaps the feedback of the movement classification in the top bar in the game screen could be a distraction. This could affect the performance in a negative way. On the other hand, the feedback bar can actually improve performance, as the users can see how close they are to the wanted classification. This improves their awareness of how to best perform the mental task, and also functions are a motivator when they are close to the target.

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 more time to learn them. This is possibly a ground for older people not performing well at imagined movement in general. Although it was made clear to every subject what kind of movement they had to imagine and how they could do it, some people still reported that they struggled with the concept of imagining a movement.

Future work could include research into the different ways of imagining movement. As McFarland et al. [86] 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 are a valuable addition.

The developed questionnaire seems to be a instrument that can aid us in evaluating differences in user experience between different modalities for BCI, but 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.

Although the game works in an online manner and the classification algorithm is fast enough to be computed real-time, 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 delay of up to two seconds. This delay sometimes leads to confusion and a lower positive affect towards the sys-

tem. Due to differences in brain signals and quality of recordings, it is possible that some users were able to produce a recognizable brain action faster (with less delay) than other users regardless of the fixed two seconds window. For example, if the CSP filters are able to separate the relevant signals better for one subject than an other, less noise is present in the two seconds window, and the classifier can be relatively certain about the class of the performed brain action even if the signal is only present in the last fragment of the classification window. In the presence of more noise, a larger part of the classification window needs to contain the signal for a positive detection, resulting in higher response times. The current problem that response times can be higher than with ordinary kinds of interfaces also call for inovative game design. If the game is designed for usage with a BCI, the user should still be able to get immersed into the game, instead of being distracted by the slower response of the system. On the other hand, a BCI can provide new kinds of games, using the users mental state (e.g. relaxation level, emotions) as a part of the game itself and in such a way create an even more immersive world than is possible using a mere keyboard and mouse [129].

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. Other BCI for movement, such as the LRP could also be explored to reduce the response delay. Preliminary research [116] indicates that the LRP of actual movements is already recognizable before a button press can be detected. Furthermore, the LRP can still be recognized reliably at rates of two strokes per second. This last property would make BCIs usable in fast paced games, which is something that traditional BCIs do not allow for.

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 communications and at the same time trying to imagine the movement. With the acquired data from the actual movement, the imagined movement could be classified.

EXPERIENCING BCI CONTROL; TO BCI OR
NOT TO BCI?

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

IEEE Transactions on Computational Intelligence and AI in
Games, 5 (2), pp. 176-184, 2013, ISI Impact 1.69

5.1 ABSTRACT

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

5.2 INTRODUCTION

Brain-Computer Interfacing (BCI) systems were in the past and still are motivated by the wish of paralyzed or otherwise disabled people to utilize new means of communication or to extend their mobility. Electroencephalography (EEG) is most often used to capture the electrical brain activity. Examples of these applications are the P300 based spelling systems in the case of communication [38] and the use of imagined movement [76] or SCPs [13] to control, for example, a wheelchair. The first BCI game was developed by Vidal et al. [138]: a simple game in which the user had to navigate through a maze by looking at off-screen fixation points flashing periodically. However, since those years the available computational power has greatly increased, hardware costs have fallen and the number of peer reviewed research articles on BCIs has increased rapidly [147]. From around the year 2000, more and more research groups

developed BCI games. For an overview, see the survey in the article by Plass-Oude Bos et al. [110]. As EEG headsets are becoming cheaper [151] and BCI technology more available, applications of BCIs in entertainment is becoming more interesting [44].

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

5.3 RELATED WORK

In this section we will look at other BCI games based on the user modulating the bandpower of certain frequency bands in their brain activity. We will also look at relevant user experience questionnaires and how duration estimation can play a role in user experience evaluation.

5.3.1 BCI Games

BCI games have been developed by several research groups, often as proof-of-concepts or to evaluate the use of mental task in an online application. One classic example of a BCI game which also inspired us to some extent is Brainball, developed by Hjelm et al. [56, 55]. The concept of the game is to relax more than your competitor. Thus two players are competing to be more relaxed, which makes it a paradoxical and fun game. In this study the ratio between frontal alpha and beta waves is taken as a measure of relaxedness.

Pineda et al. [107] developed a BCI for a 3D shooter game. Forward and backward movements were controlled by the keyboard, turning left and right by alpha levels over the motor cortex. As this study focussed on the ability of participants to learn to control their alpha/mu levels over the course of several

weeks the study only included 4 participants. They found that control over mu activity was easily obtained and maintained.

Instead of focussing only on bit rates and accuracies Plass-Oude Bos et al. applied a user-centered approach and found that the ease of executing a certain mental task is an important factor as well [110]. Gürkök et al. found that certain mental tasks for BCI games while being appropriate might not be the best option as the major modality for interaction. [50].

5.3.2 *User experience*

According to Witmer et al. [145] both involvement and immersion are necessary for experiencing presence. Witmer et al. introduced a presence questionnaire addressing these factors for specific use in virtual environments. Nowadays, games can be at least as realistic as virtual environments. A succesful game drags the user along with its immersive graphics, compelling characters and creative narrative. Users are then able to feel themselves present in the environment or world that is created by the game.

The sense of presence can be interrupted by distracting factors such as audiovisual stimuli that are not congruent with the virtual world or interfaces that are unnatural or faulty. At the intersection of these games which aim at providing the ideal circumstances for a feeling of high presence, good user experience is key to make games with BCIs succesful. Van de Laar et al. [134] made an attempt to evaluate a BCI game with a questionnaire based on the GEQ by IJsselsteijn et al. [61] with specific items on the role of the BCI. Van de Laar et al. [135] gave an overview of the several methods available to researchers to evaluate BCI games and when to use them. Administering a questionnaire is the best method for quantifying results. To be able to answer the *why* question, interviewing would be a better method.

5.3.3 *Measuring Immersion*

When people play a game and are immersed in the experience, they tend to forget the time. This experienced, subjective duration is considered by Waterworth [144] to be influenced by the amount of activity in the working memory. Users who are deeply engaged in experimental tasks or activities such as playing computer games might experience a different duration from those who are not engaged to that extent. Previous studies have not reached a consensus on the direction or the nature of this

relation. On one hand Waterworth et al. [142] suggested that a high level of presence would yield an overestimation of duration. On the other hand the study by IJsselsteijn et al. [60] stated the opposite; in the case of a high level of presence the user would underestimate the duration, no significant data supported this claim however. Davies [30] suggested that a high level of presence may cause the user to either under or overestimate the duration. In an experiment by Waterworth et al. [143] designed to assess the validity of their model posed in [142], participants looked at audiovisual streams with varying content and durations. They found relatively weak relations for some streams between presence and estimated duration. However, the experiment was non-interactive and the longest duration between being interrupted from the experience was 104 seconds. The various studies on measuring immersion by estimation duration do not agree with each other. In this study we will do an analysis of duration estimation in two conditions that are interactive and longer than previous studies to assess whether duration estimation can be of value in UX evaluation.

5.4 METHODS

5.4.1 *World of Warcraft*

World of Warcraft (WoW) is a Massive Multiplayer Online Role Playing Game (MMORPG), meaning that thousands of users log on to servers with their virtual characters to battle monsters or each other and evolve their character through an interactive narrative set in a fictional medieval virtual world. The game is immensely popular with currently more than 10 million subscribers. The objective of the game is essentially to level up, get better abilities and better gear. This is done by achieving experience points, gained by completing quests, slaying enemies and exploring the world.

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

5.4.1.1 *Bear shape*

The bear shape simulates the armor and health similar to a warrior, therefore providing a so-called Tank ability. In this shape the player is able to soak up damage, especially useful when

fighting more than one enemy. At the same time the bear has slow speed but high impact attacks at close range (melee).

5.4.1.2 *Druid shape*

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

5.4.2 α WoW

In our adaptation of WoW, called α WoW, we make use of the power in the alpha band over parietal regions. According to Cantero et al. [23] high alpha levels in the parietal lobe indicate a state of relaxed alertness. Also Barry et al. “[..]confirm the arousal link between alpha and electrodermal activity” [8]. We mapped this state of relaxed alertness to the shape of the Night Elf in druid shape who is strongly dependent on intelligence and mental concentration. The opposite state of relaxed alertness (the decrease in alpha band activity) would be a state of stress or agitation providing a natural relation to the bear which is eager to fight.

The relation of alpha levels to inhibitory processes and passive or idling states of certain parts of the brain is often studied but also often misunderstood as suggested by Basar et al. [10] and [11]. According to Basar et al. [9] and Pineda et al. [107] users can train their alpha levels. This is also what we observed during the many demos we gave of α WoW at international (e.g. AISB2009 [1]) and national conferences and popular-scientific events (e.g. TEDxAmsterdam 2009 [2]). One person often served as demo subject and learned to control alpha levels to such an extent that intentionally transforming every 5 seconds was possible.

5.4.3 *Experimental setup*

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

To acquire the EEG data we used Emotiv's EPOC. The Emotiv EPOC is an inexpensive wireless EEG device. Although inexpensive it provides more sensors than other inexpensive systems and high usability [127]. It has also been used in other BCI based games for example by Van Vliet et al. [137]. We fitted the set on the users head, tilted at an angle of about 25° (see Fig. 16) to be able to measure the parietal brain activity with the O1, O2, P7 and P8 channels. After tilting the headset these sensors are approximately located at the P1, P2, CP7 and CP8 positions respectively. Figure 18 shows randomly taken examples of EEG data segments indicative of a participant in a 'Stressed' state (high arousal), a 'Relaxed' state (low arousal) and a segment of data with eyes closed to compare.

A webcam with a built-in microphone was placed in the corners of both rooms to monitor the participant when the experimenter left the room and to record video and audio during the interview at the end of the experiment. An overview of the experimental setup (seen from the webcam) can be seen in Fig. 19.

We used 2 separate WoW accounts on 2 different servers but within the same game area. We made a default interface layout with offensive and defensive actions recommended for the respective character level. When leveling up a character, it was moved to an area that matched its competence and possible new actions were added to the action bar. Since the character's armor perishes over time, it was repaired after every session. Every participant played two sessions, one with and one without BCI control. See also section 5.4.4.1.

5.4.3.1 *Control signal*

The pipeline we used (see Figure 17) to assess the amount of alpha power in the parietal region starts with the acquisition of the EEG data from the Emotiv EPOC headset. The EEG data from channels O1, O2, P7 and P8 is selected and streamed to the analysis pipeline. The online data stream is cut into windows of two seconds with an inter window interval of 250 ms (4 Hz). For every window a Fast Fourier Transform (FFT) is computed for all channels and the band power from 8 – 13 Hz is averaged over all channels. An adaptive Z-score normalization with a time window of two minutes is used to convert the band power into a usable control signal and limit the influence of outliers. A short term smoothing including the previous two windows is applied with a 20%, 30% and 50% contribution respectively. Because it is nearly impossible to validate the control signal in the game setting, we also did a control experiment to validate it which we will explain in section 5.4.5.

Figure 16: Photograph of the Emotiv EPOC headset that was tilted forward about 25°

Figure 17: The pipeline including the signal processing steps as implemented in Alpha WoW. For a detailed description see section 5.4.3.1.

Figure 18: Examples of EEG data during Eyes closed, 'Relaxation' and 'Stressed' states.

Figure 19: Overview of the experimental setup

5.4.3.2 *Mapping the control signal to game actions*

The control signal is a 4 Hz signal in the $[0, 1]$ range. To map this signal to actions within WoW we interpret and convert the control signals into virtual key strokes. Every 250 ms the value of the control signal is rounded to the nearest first decimal and sent through a keystroke to the game to supply the user with feedback through means of an animated orange bar, known as the ‘stress bar’ (see Fig. 20). The longer the bar, the lower the value of the control signal and the lower the amount of alpha activity. As to control of the actual shape shifting we applied hysteresis, id est, direction specific thresholds of 0.3 and 0.7 to trigger the transformation. As the control signal is normalized the 0.3 and 0.7 thresholds roughly compare to either one standard deviation below and one standard deviation above the average of the signal. To transform to bear form a control signal of 0.3 or lower is needed, to transform from bear to druid a control signal of 0.7 or higher is needed.

To make sure the user really intended to transform after 1 second (4 windows) of dwell time an indication of the upcoming transformation is shown. In case of transforming from bear to druid the edges of the screen flashed a translucent blue for 1 second, in case of transforming from druid to bear the edges flashed a translucent red for 1 second. If the control signal still exceeds the threshold 1 second after the onset of the flash (8 windows in total), the actual transformation was triggered.

Figure 20: Screenshot of WoW with the add-on enabled. The orange bar (top left) provides feedback on the user's current alpha level. Bigger means less alpha.

5.4.4 Participants and instructions

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

5.4.4.1 Progression of the experiment

The participant was welcomed and took a seat at the laptop. To make the time assessment as unbiased as possible, participants were asked to turn off mobile phones and to hand them in together with their wrist watches for the duration of the experiment. WoW by default has a small on screen clock which we removed through a script at the start of every experiment. Because the desktop of the Windows environment was visible before launching WoW we hid the taskbar. The Emotiv EPOC headset

was installed on the participant's head and the experimenter made sure all sensors were operating at green signal quality levels, which indicates proper impedance levels according to the manufacturer of the headset. The participant received general instructions on playing WoW and played a tutorial session of 5 minutes to get acquainted with the controls and interface. The experimenter answered any questions posed by the participant. After this, either a session with or without BCI control started with the according instructions. After the BCI pipeline was started the normalized control signal was given the time to stabilize for 1 minute before the actual session started. If instructions were clear, the experimenter literally told the user to play the game until they did not want to play anymore. In the case that they wanted to stop the current session, the participant was instructed to ring a bell that was on the table he was sitting at. The experimenter left the room and started the stopwatch. If the user did not ring the bell within 30 minutes, the experimenter interrupted the participant playing the game, if the user did ring the bell, the experimenter immediately noted the time that had passed and went back into the experiment room. In either case, the experimenter immediately administered the questionnaire because the participants had to make an estimation of the duration of the session. After the questionnaire was finished, the experimenter prepared for the second session and gave the participant the according instructions. During this time, absolutely no mention of time passed was given, even when asked for by the participant. The second session followed the same course of events as the first. After the second session, the Emotiv EPOC was removed from the participants head and recording of video and audio started to capture the oral interview that followed.

5.4.4.2 Questionnaire

The questionnaire administered in this study was based on the presence questionnaire by Witmer et al. [145]. This questionnaire although published in 1998 is still the *de facto* questionnaire for measuring presence. According to Van Baren et al. [133] most other presence questionnaires are based on the one by Witmer et al. and are less well validated. We used the Involvement/Control scale and the Natural scale as the other sub scales in this questionnaire would not provide us with useful information. We also added two scales. The Transformation Control scale has specific items on the amount of control the user has over the transformation action. The items in this scale are based on [134], specified on BCI specific issues that would not be captured by the Involvement/Control Scale. The Fun/Achievement

scale was also added to provide us with information on the amount of fun, concentration and achievement the user experienced.

5.4.5 *Validation of the control signal*

To assess the performance users can achieve by means of the BCI pipeline used in the first experiment, we conducted a second experiment in a more controllable and clinical setting. The concept of the game is a ball that continuously falls down and has to be 'pushed' into the correct (green) basket, see figure 21. The bottom of the screen is divided into two equal parts (baskets) both randomly appearing either red or green during the fall of one ball. It takes 7 seconds for the ball to fall down. In this window of 7 seconds the participant has to push the ball to the left or to the right depending on which half of the screen is green. By generating high alpha levels the ball is pushed to the left, consequently, by generating low alpha levels the ball is pushed to the right. The same pipeline and control signal as defined in the previous section is used for this experiment, with the screen divided into ten equal parts, according to ten corresponding levels in the control signal. (0.1, 0.2, 0.3 etc.). Near real-time feedback is generated at a rate of 4 Hz as in the first experiment. Whenever the ball ends up in the green part the score is increased by one. Ten participants were recruited for this second experiment. Eight participants were male, two participants were female. Participants were not given a monetary reward. Participants trained for three sessions of 20 trials each. After the training 5 sessions were done resulting in 100 trials per person.

5.5 RESULTS

In this section we will first assess the internal consistency of the scales we used in our questionnaire. Next we will review our measure for duration estimation by the participants. Next we will do an analysis of variance on our data to assess whether users experienced the condition with BCI control differently (and if so in what way) than the condition without BCI control. Lastly we will review the results of the second experiment on how much control is attainable with the BCI pipeline that was used.

Figure 21: Screenshot of the ball/basket game.

5.5.1 *Item Analysis*

In our questionnaire we used the sub scales Involvement/Control and Natural from the Presence questionnaire by Witmer et al [145]. The reported Cronbach's Alpha for the whole questionnaire was 0.81. In our study we found the Cronbach's Alphas to be a bit lower, 0.733 for Involvement/Control and 0.655 for Natural. We constructed two other scales, one on Transformation Control, involving questions on the mechanism to shape-shift specifically and a scale we would call Fun/Achievement, which indicates whether the user had fun, liked to do things he/she could not do in real life, felt a sense of achievement and was concentrated on the game. Both showed good internal consistency, with a Cronbach's alpha on normalized items of 0.844 and 0.725 respectively.

5.5.2 *Duration Estimation*

As explained in Section 5.3.3 the relation between presence and duration estimation by the user is not that clear. Some studies point towards overestimation in the case of high presence, other studies state that underestimation is more prone to happen. For every participant an actual (objective) duration and an estimated (by the user) duration were recorded. Comparing the means over all users with a Wilcoxon signed ranks test yielded no significant differences for either the actual duration between conditions ($p = 0.456$) or estimated duration ($p = 0.227$) between conditions. To evaluate whether overestimation or under-

Table 7: Grand averages of participants. Durations per condition (min.) and number of over- and underestimations of duration.

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

estimation of duration was more prone to happen we counted the number of cases in which the difference in durations was bigger than zero and the cases in which the difference was smaller than zero. An overview can be seen in Table 7. This table also shows the average amount of over and underestimation.

To be able to see if there were differences between experienced and not experienced participants, we split up this analysis according to this group factor. An overview of these results can be seen in Table 8. This result hints at a slightly better notion of time for experienced participants. The average overestimation is lower and the percentage of participants who estimated the duration correctly is higher for this group.

Because there is no clear trend in these results we took the absolute value of the difference between the actual time and the estimated time according to the following equation:

$$\text{Duration\_diff} = |\text{Duration\_act.} - \text{Duration\_est.}| \quad (5)$$

Using a Pearson product-moment correlation to correlate this measure to the 4 scales in the questionnaire, compensating for multiple comparison with a Bonferroni correction yielded no significant results. A Wilcoxon test showed no significant difference between Duration_diff between the two conditions ($z = -1.144, p = 0.259$). However, looking at the Duration_diff measure within subjects over the two conditions showed a strong correlation ($\rho = 0.59, p < 0.001$).

Table 8: Averages of experienced and inexperienced participants. Durations per condition (min.), number of over and underestimations of duration and average amount of over and underestimations.

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

5.5.3 Analysis of Variance

A 2x2 one-way repeated measures MANOVA was conducted to compare scores on the four scales (Involvement, Naturalness, Transformation Control and Fun/Achievement) at Condition 1 (no BCI) and Condition 2 (BCI). Between subject factors are gender (female/male) and experience (yes/no). The overview of this analysis can be seen in Table 9 and Table 10. There was a significant effect for Condition (Wilk's Lambda = 0.426, $F(4,37) = 12.452$, $p < 0.001$, multivariate partial eta squared = 0.574). Further investigation of the univariate within-subjects tests showed that both the Involvement/Control scale ($F=20.369$, $p < 0.001$, partial eta squared = 0.337) and Transformation Control ($F=40.373$, $p < 0.001$, partial eta squared = 0.502) largely caused this effect. No significant interaction effect was found for Condition x Experience ($p = 0.421$), Condition x Gender ($p = 0.391$) nor Condition x Experience x Gender ($p = 0.784$). Between-subjects, the univariate test with source Experience showed a significant effect for the Natural scale ($F=5.655$, $p = 0.022$) with a positive direction, id est, Experienced players scored higher on Natural. All other test proved not significant.

5.5.4 Second experiment: accuracy

All 10 participants in the second experiment performed 100 trials. In table 11 we report the accuracy for every subject, the average accuracy (74.1%) the standard deviation and the corresponding Student's t confidence interval for a sample size of 10.

Table 9: Multivariate tests. Effect, Wilks' Lambda, F-measure, hypothesis df, Error df and p-value. 'Experience' is between-subjects, 'Condition' is within-subjects

Multivariate tests					
Effect	Λ	F	df ₁	df ₂	p
Experience	0.724	3.328	4	35	0.021
Condition	0.548	7.222	4	35	0.001

Table 10: Within-subjects and between-subjects univariate tests. Source of the effect, Effect, Mean Squares, F-measure, hypothesis df, Error df and p-value

Univariate tests						
Source	Effect	MS	F	df ₁	df ₂	p
Condition	Involvement	1.699	11.179	1	38	0.002
	Transf. Control	37.561	21.708	1	38	0.001
Experience	Natural	10.001	5.728	1	38	0.022

74.1% $\pm$ 9.95% = 64.15% to 78.15%. Random chance level for a binary BCI would be 50%.

5.6 DISCUSSION

The reliability analysis of the scales showed good internal consistencies for all scales, although somewhat lower than reported by Witmer et al. [145]. The cause of this lower internal consistency may be caused by the fact that we used a modern game in our experiment instead of a by today's standards old fashioned virtual environment. Another difference is in the population of participants that is becoming more and more educated in operating computers, controlling input devices and navigating through 3D immersive worlds. We constructed the Transformation Control scale to measure the amount of control over the transformation action. This scale showed a high internal consistency of 0.84 with only 6 items. A high internal consistency makes sense considering the specific nature of the scale. It measures the amount of control over the transformation either with or without BCI. The Fun/Achievement scale consists of 4 items on positive effects: having fun, like to do things in the game,

Table 11: Subject accuracies, average, standard deviation and Confidence Interval.

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

being able to concentrate and having a sense of achievement. Concerning concepts broader apart, but still achieving a good internal consistency.

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

Further analysis of the actual and estimated durations showed no clear over or underestimation of time. This is not in line with what most previous studies reported. Looking at the absolute difference between actual and estimated duration we saw no correlation to any of the scales, supporting the findings by Waterworth et al. [143] that duration estimation is not a clear indication of presence. Furthermore we saw a very strong correlation within subjects. This suggests that duration estimation is largely a personal trait. Further analysis indicated that experienced participants were a little better at estimating the duration. This may well be due to the lower workload involved because of familiarity with the interface and the workings of the game.

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

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

1. Participants found the EEG headset and BCI a novel and fun gadget to play with, notwithstanding the fact that it might have provided them with less involvement in the game. This might be due to a novelty effect of using brain activity and EEG devices in general. Some participants also stated something along these lines during the interview. Either they wanted to see if they could get better at controlling the transformation process or they said they liked participating in the experiment but would not pay to play with a similar setup including the BCI at home. This possible novelty effect is something that is hard to prove or disprove. It can cloud user experience evaluation of BCI games in general. The only way to investigate whether, and if so to what extent, the novelty effect is influencing these results is to perform several follow-up experiments with the same participants.
2. The BCI provided the users for which it worked better with an added modality to aid them playing the game. Several experienced WoW players stated in their interview that they could see such a kind of BCI as a nice challenge when playing with a low-level character. They also felt that their amount of control would improve over time, providing them with a useful extra modality in a game which always has a shortage of buttons to express the plethora of functionalities in such a complex game.

3. For the participants who experienced no or little control, the BCI did not interrupt their experience to such an extent that it became a frustration. Some participants also reported this. "I just kept on playing and tried to quickly adapt. At a certain point I memorized all the shape specific functions and their respective buttons."

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

The second experiment in which we evaluated the performance of the BCI pipeline in a more clinical setting showed that on average participants achieved an accuracy of 74.1%. This is significantly higher than one would expect from a system operating on random choices. Looking into the per subject accuracies, there were two distinct groups. One group (5 participants) who got barely above chance level control (56% - 67%) and one group (5 participants) which achieved good to excellent control (82% - 97%). As the participants are not the same as the participants who participated in the first experiment we cannot analyse the effect of a low or high accuracy on user experience. However, this result shows that control with the current BCI pipeline is definitely possible but whether the user has good control is subject dependent.

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

BRAINBRUSH, A MULTIMODAL APPLICATION
FOR CREATIVE EXPRESSIVITY

Bram van de Laar,
Ivo Brugman,
Femke Nijboer,
Mannes Poel,
Anton Nijholt. Sixth International Conference on Advances in

Computer-Human Interactions (ACHI 2013), pp. 62-67, 2013

6.1 ABSTRACT

We combined the new developments of multimodal Brain-Computer Interfaces (BCI) and wireless EEG headsets with art by creating BrainBrush. Users can paint on a virtual canvas by moving their heads, blinking their eyes and performing selections using a P300 BCI. A qualitative evaluation (n=13) was done. A questionnaire was administered and structured interviews were conducted to assess the usability and user experience of the system. Most participants were able to achieve good control over the modalities and able to express themselves creatively. The user experience of the modalities varied. The use of head movement was considered most positive with the use of eye blinks coming in second. Users were less positive about the use of the BCI because of the low reliability and higher relative cost of an error. Even though the reliability of the BCI was low, the BCI was considered to have an added value: the use of BCI was considered to be fun and interesting.¹

6.2 INTRODUCTION

The act of creative expression is considered by many to be a purely human ability and skill [3]. Creative expression allows humans to express their identity and it can take multiple forms: for instance making music, dancing, painting, writing or acting. Apart from these more traditional forms of art, new art forms have emerged, such as computer art, motion graphics and the use of virtual reality environments for art.

In recent years, a neurorevolution has taken place in which neurotechnology is a hot topic. For instance, research has been done on improving task performance using neurotechnology. Some athletes now use neurofeedback training to enhance their performance by managing the stress of training and competition [33]. Furthermore, Ros et al. have shown that the microsurgical skills of ophthalmic microsurgeons can be improved significantly with neurofeedback training [118]. Neurotechnology has even been applied to the field of economics, creating the science of neuroeconomics: combining neuroscience, economics and psychology to explain the human decision making process [119]. In neuromarketing, neurotechnology is used to research why consumers make certain choices: for instance, why consumers prefer either Pepsi or Coca Cola [84]

Brain-Computer Interfaces (BCIs) are another example of research in the field of neuroscience are. A BCI provides a direct

¹ For more information and additional discussion, please refer to Chapter 7

interface between a human brain and a computer, without using peripheral nerves or muscles. For years BCI research has mainly focused on assistive technology for people with disabilities. For instance, patients with the motor neuron disease Amyotrophic Lateral Sclerosis (ALS) can now benefit from BCIs like the P300 speller [37], brain-controlled wheelchairs [45] or BCIs to control their environment [57].

In recent years, BCI research has also focused on applications for healthy people. One example is the ‘NeuroPhone’ system developed by Campbell et al. which allows neural signals to drive mobile phone applications on an iPhone using a wireless EEG headset, for instance to dial a phone number using the same principles as the P300 speller [22].

Furthermore, research is being done on incorporating BCI technology in games. Plass-Oude Bos et al. developed AlphaWoW, incorporating BCI in the game ‘World of Warcraft®’ [113]. They used alpha waves in the EEG for automatic adaptation of the avatar shape from bear to elf and vice versa. Gürkök et al. used SSVEP for sheep herding in the game ‘Mind the Sheep!’ [49].

BCIs can also provide a unique link between the source of creativity, the brain, and art. Christoph De Boeck created a responsive environment, Staalhemel (<http://www.staalhemel.com>), where 80 steel segments are suspended in a room, above the visitor’s head. The visitors wear a portable EEG headset and as they walk through the room, tiny hammers are activated by their brainwaves, tapping rhythmical patterns on the steel segments.

Other examples of the use of Brain-Computer Interfaces for art include the ‘Brain-Computer-Music Interface’ which enables a disabled person to create music [88], and the ‘Câmara Neuronal’, a performance where the brain signals of the performer are translated into audio and visual compositions (<http://projects.jmartinho.net/3486412/Camara-Neuronal-Video-Teaser>).

The German artist Adi Hösle, in cooperation with the Institute of Medical Psychology and Behavioural Neurobiology at the University of Tübingen, designed the application ‘Brain Painting’, a painting application which is controlled using a BCI and enables paralyzed patients to express themselves creatively. In ‘Brain Painting’, all actions are performed using the P300 paradigm. The system uses two screens for the painter: one screen displays the P300 matrix while another, larger, screen shows the painting canvas. The standard P300 speller matrix, containing characters and numbers, was adapted to contain symbols indicating different colors, objects, grid sizes, object sizes,

transparency, zoom and cursor movement. By repeatedly making selections using this P300 speller matrix, users can paint pictures on the virtual canvas. In the first evaluation of the 'Brain Painting' application, with 3 ALS-patients and 10 healthy subjects, both the ALS-patients and the healthy subjects were able to use the application with high accuracies: during a copy-painting task, the ALS patients achieved an average copy-painting accuracy of 70.18% while the healthy subjects scored an average accuracy of 80.53% [91]. One participant in the 'Brain Painting' study, who was severely disabled due to ALS, described her experience with the system: *"I am deeply moved to tears. I have not been able to paint for more than 5 years. Today I again had butterflies in my stomach, a feeling that I have missed for so much, so much [sic]. I was so sad, I was plagued by fears of loss, I was in shock because I could not paint. For me the picture I have created is so typical for me, no other paints in my style, and despite five years of absence, I am simply an artist again; I'm back to life!"*. Even though this feedback is very positive, the artistic freedom of the painter is limited due to the fact the cursor cannot be moved freely; the cursor can only be moved in a predetermined grid.

In continuation of the 'Brain Painting' research, Holz et al. developed the 'Brain Drawing' application to overcome the cursor movement limitation of the 'Brain Painting' system. In the 'Brain Drawing' application, imagined movement is used to control the cursor when drawing. During the first evaluation with 1 subject, the subject performed a Copy Drawing task in which he was instructed to draw a simple object (circle, ellipse or rectangle) on a virtual canvas. Holz et al. considered 4 out of 36 copied drawings to be successful by visual inspection and the subject found it very difficult to draw. The subject had to focus his attention for a long period because he continuously had to imagine movement, which resulted in high workload [58].

Todd et al. used two different BCIs in their research on how creativity can be supported and assessed using a BCI [131]. With their first BCI, users could only control a drawing cursor in horizontal and vertical directions by looking at one of four LEDs placed at the top, bottom, left and right of the screen. The cursor would move in the direction of the LED the user looked at, and continue drawing in that direction until the user looked at another LED. For the second application, the four LEDs were mapped onto four shapes (circle, star, square and line). After choosing a shape, the shape would be drawn on the canvas. Users did not have any control over the position, size or color of the shapes. Todd et al. concluded that relying completely on the efficiency of a BCI for image production is not practical as BCI technology is not yet mature enough for 100% reliability.

A possible solution they suggest is to create a hybrid, or multimodal, BCI by combining a BCI with other input modalities such as an eye-tracker.

These examples show that the focus in BCI research should shift from reliability to usability and user experience as is also reported by the FutureBNCI roadmap [3]. This shift in focus is necessary in order for BCIs to migrate out of the lab and into society. Healthy persons can choose from various alternative input modalities. For healthy persons to choose for BCI, the user experience and usability must be adequate. Most people have never used a BCI and the novelty of this new technology can be a reason for people to decide to use a BCI instead of alternative input modalities, even if a BCI is less reliable and slower. However, if the user experience and usability are not good, people are expected to choose a different input modality which provides a better user experience and usability. Due to the fact that the focus in BCI research has mainly been on the reliability, no standardized methods to assess the user experience for BCIs exist yet. Gürkök et al., Plass-Oude Bos et al. and Van de Laar et al. addressed the need for standardized methods to assess the user experience for BCIs [51, 111, 136]. Van de Laar et al. proposed a questionnaire consisting of a core containing general questions and modules for the different kinds of mental tasks and ways of interacting with the BCI [135].

In this study, we combined the new developments of multimodal BCIs and wireless EEG headsets with art by creating a multimodal interactive system (called BrainBrush) which allows healthy persons, but possibly also patients, to express themselves creatively. We evaluated this system to explore how the different modalities contribute to the user experience and whether BCI has an added value for this system.

6.3 METHODS

First we will describe how BrainBrush was developed and how the design choices were made. Second we will outline how we set up our experiment to evaluate the system.

6.3.1 *Development of BrainBrush*

BrainBrush (see figure 1) was developed in a two step iterative process. The outcome of the first evaluation and how this influenced our design choices is however beyond the scope of this

paper. For a more in-depth description see the work by Brugman [18].

We aimed to design the BrainBrush system in such a way that it would be appealing to healthy persons and would also be usable for patients who had lost control from their neck down (e.g. due to spinal cord injury). We expected healthy persons would find it appealing to be able to operate the BrainBrush system using only head movement, eyeblinks and brain activity because it is completely different to how healthy people normally use a computer. For patients who do not have control over their limbs, it is a necessity to be able to operate the system using nothing else than the described modalities.

Head movement, the P300 speller and eyeblinks were used as the input modalities for the BrainBrush system. Head movement has been used successfully in the past as an input modality, for instance to control a cursor [34]. Furthermore, the P300 speller has proven to be a robust BCI paradigm [48] and has been used in many BCI systems. The P300 signal is a positive deflection in the ongoing EEG signal observed roughly 300ms post stimulus over the centro-parietal area of the skull [37]. Finally, eyeblinks have been used for communication systems for ALS patients, such as the system for making selections on a computer screen proposed by Takeshita et al. [128]. However, the eyeblinks are usually detected using a camera and image processing techniques, instead of using an EEG headset. Chambayil et al. have shown promising results for their virtual keyboard BCI which uses eyeblinks to select characters [25]. All three modalities can be captured by one device: the Emotiv EPOC, using the EEG sensors and built-in gyro sensors.

The task of creating a brushstroke requires the system to provide users with a way to signal when they want the paintbrush to be put to the canvas and to signal once again when they want to take the paintbrush off the canvas, thereby ending the brushstroke. This on/off switch for the paintbrush has been implemented by detecting intentional eyeblinks in the EEG data based on an algorithm by Plass-Oude Bos et al. [109]. A template was constructed out of an EEG recording with 40 intentional eye blinks. A threshold is set for the Euclidean distance between the template and the live EEG data over channels AF3 and AF4. Both template and window have a length of 103 samples. When the Euclidean distance is below the threshold on either of the channels the on/off switch is triggered.

Head movement was implemented by using the gyro sensors of the Emotiv EPOC headset. An application called the *Mouse Emulator* is provided with the device and was used to convert

Figure 22: User interface in Painting mode with the 'New Painting' menu option at the top and the 'Undo' and 'Redo' menu options at the bottom

head movement into cursor movement in order to move the paintbrush across the virtual canvas.

In BrainBrush, we use P300 speller grids for the selection of brushes and colors because the P300 speller is suitable for making selections from a large set of options in a relatively short period of time [6], compared to other Brain-Computer Interfacing paradigms. In a study performed by Guger et al. with healthy participants, 88.9% of the participants were able to achieve at least 80% accuracy using the P300 speller paradigm [48], showing the robustness of the P300 signal. Although, we suspect a lower accuracy in our system because of the lower signal quality of the EPOC hardware. For BrainBrush, we use the original grid structure, but instead of grids with characters, grids with pictures are used where each picture depicts a certain brush or color (see Figures 2 and 3). The *P3Speller* module of the BCI2000 framework is used for the implementation of the P300 speller grids [121]. The stimulus duration is set to 100ms and the inter-stimulus interval to 175ms. For the selection of a brush or color, 15 sequences of flashes are shown, meaning each row and column flashes 15 times. Therefore, the target is flashed 30 times in total.

Finally, BrainBrush uses the open source MyPaint application to provide us with a virtual canvas (see figure 1). An overview of the complete system can be seen in figure 25.

Figure 23: User interface during brush selection showing the 11 available brushes and the eraser

Figure 24: User interface during color selection showing the 12 available colors

Figure 25: BrainBrush system overview

6.3.2 Experimental Design

The BrainBrush system was evaluated on user experience with a qualitative user study. We want to gain an understanding of underlying reasons and motivations for both the positive and negative opinions the participants formed about the user experience with the system and the various input modalities. Participants were asked to fill out an informed consent form and provide their demographic details. They were given instructions on how to use the P3 Speller. They could try the P300 speller in a practice session, after any questions the experimenter proceeded to record a training session for the P300 classifier [73]. Training data consisted of a 10 letter word ('BRAINPOWER') with 15 sequences utilizing all 14 channels of the Emotiv EPOC. After the training session an online session was carried out in which participants tried to spell the 5 letter word 'PAINT' using the *P3Speller* without feedback so not to bias their expectations. Next, participants were given instructions on using BrainBrush. After 5 minutes of familiarizing themselves the free painting sessions started and the experimenter left the room. Participants could take as long as they wanted, up to 30 minutes. If they wanted to stop painting at an earlier point in time, they would ring a bell to signal the experimenter. After the free painting session, the experimenter would administer the SUS questionnaire [17] followed by an in-depth structured interview.

6.4 RESULTS

The group of participants for the user study consisted of 8 males (61.5%) and 5 females (38.5%). All participants were aged 20 to 29, the average age was 24.8 (standard deviation: 2.9). 11 participants (84.6%) are right-handed, 2 left-handed (15.4%). All participants had the Dutch nationality.

None of the participants reported relevant medical conditions. Out of 13 participants, 2 participants (15.4%) indicated that they had previously participated in BCI research. Four participants (30.8%) indicated they regularly exercise some form of creative expression: participant 001 paints, draws and designs daily for study-related purposes; participant 012 paints once a week as a form of recreation; participant 008 regularly uses Photoshop and participant 009 sometimes plays the drawing-game 'Draw Something' on the iPad.

During the interviews after the free painting session, the participants were asked what accuracy they thought they had been able to achieve with the brush and color selection during the

Table 12: P300 classification accuracy during copy spelling and free painting

Participant	Copy spelling		BrainBrush
	Result	Accuracy	Accuracy
001	R6INW	40%	Regularly incorrect
002	XAINI	80%	70%
003	PACNZ	60%	80%
004	RHINT	60%	Regularly incorrect
005	R5IZT	40%	Regularly incorrect
006		-	< 50%
007	LAINI	80%	At the beginning: 50% At the end: 100%
008	RAIPT	60%	70%
009	RAINI	80%	21 out of 24 sessions correct (88%)
010	RAINI	80%	70%
011		-	Brush selection: 50-60% Color selection: 90%
012	L5BNN	20%	Choice decided beforehand: 100% Otherwise: lower
013	RGUNC	20%	25-33%
Average		56.4%	

free painting session. These results together with the results from the online session are shown in table 12.

During the experiments of participants 006 and 011, the classifier was accidentally not loaded during the online spelling session. Therefore, no copy spelling results are available for these participants. For the other 11 participants, the average accuracy during copy spelling was 56.4% (standard deviation: 23.4%, minimum 20%, maximum 80%).

The average SUS score for all participants was 66.2 (standard deviation: 14.2, minimum: 45, maximum: 87.5). A score of 68 would be average for all systems tested with SUS. During the interviews participants were asked to order the three modalities on pleasantness. An overview of the results can be seen in figure 26.

A summary of all analyzed topics covered in the questionnaire can be seen in table 13.

Figure 26: Ordering of pleasantness of modalities: Head movement was most often ranked most pleasant, eyeblinks most often second and the P3Speller (BCI) most often last.

Table 13: Summary of interview results

	+	+/-	-	Total
Expectations	6	7	0	13
Transfer creative ideas	4	3	6	13
Draw picture in mind	3	7	3	13
Time spent painting	4	2	7	13
Purchasing the system	0	4	9	13
Eyeblinks	7	2	4	13
Head movement	7	4	2	13
P3Speller	4	3	6	13
BCI for creative expression	9	2	2	13
Combination of modalities	9	1	3	13
Fun	12	1	0	13

6.5 DISCUSSION AND CONCLUSION

In this study, we set out to develop a multimodal interactive system which allows persons to express themselves creatively and included Brain-Computer Interfacing technology. Using this multimodal interactive system, we wanted to research how the different modalities contribute to the user experience and whether BCI has an added value.

We developed the BrainBrush system, which lets users paint on a virtual canvas using their head movement for brush control, eyeblinks to turn the brush on and off, and the P300 speller to select different brushes and colors. The presented system is the first of its kind utilizing these modalities.

A user study with thirteen participants showed that the BrainBrush system does not enable all users to express themselves creatively. The group of users who were able to achieve good control over all three input modalities were able to express themselves creatively and made nice paintings. However, for all users to be able to achieve this, the reliability of the input modalities must be improved.

The head movement modality was considered to be the most pleasant. However, misalignment of the cursor and a lack of smoothness negatively influenced the user experience. Replacing the Emotiv Mouse Emulator program with new software to translate head movement to cursor movement, is expected to improve the user experience.

After the head movement modality, the use of eyeblinks to turn the brush on or off was considered to be the most pleasant. In general, the user experience for this input modality was positive. However, the user experience can be further enhanced by performing more research into the topic of intentional eyeblink detection and improving the detection.

The P300 speller was considered to be the least pleasant input modality. In contrast to the other two input modalities, the user experience was not positive: using the P300 speller was considered to be mentally tiring and it caused physical discomfort and frustration. Participants suggested improvements to the P300 training setup and to the BrainBrush system.

The multimodal aspect of the system was good for the user experience. The combination of the three input modalities within the BrainBrush system was considered to be positive: it was a combination that made sense to most users and the input modalities were well balanced.

Finally, concerning the value of BCI for a multimodal interactive system for creative expression, we can conclude that BCI does have an added value: even though there were some issues due to the BCI modality, the use of BCI was considered to be fun, cool and interesting.

We feel the BrainBrush system in its current state offers a good basis, and with the suggested improvements, the user experience should improve further.

Part III

PART III CONCLUSION

This part of the thesis consists of two chapters: the first contains a discussion on the previous chapters. In the second chapter we will review the results of the various studies and study their implications. Furthermore we will discuss any limitations and future work regarding the research in this thesis.

In this chapter we will elaborate on the experimental studies we conducted and discuss some points which were less relevant in their respective articles but are worth looking into more extensively.

7.1 DISCUSSION OF CHAPTER 3

In Figure 8 in Chapter 3 we saw an interesting sawtooth shape. We will also discuss the sample size and the validity of fitting a curve to our data.

7.1.1 *Dataplots*

In Chapter 3 we analysed the relations between control and frustration and between control and fun. If we recall Figures 8 and 9 one can observe some sort of sawtooth characteristic in the plotted data. One might argue that something else is influencing the values for frustration and fun apart from the MUI. This however cannot be the case. There are four arguments why the apparent sawtooth characteristic is probably a coincidence:

- The data points were acquired from participants who were randomly assigned in multiple play sessions, at least one and at most five, to the 15 conditions. If there was to be a certain order in which participants were assigned to the conditions, or if there were any differences other than just the amount of control this could point to some underlying factor influencing the shape of the plot. This was not the case.
- If one compares the two figures more closely, one can see that the sawtooth characteristic is completely different between the two plots. While in Figure 8 it is more apparent in the middle of the plot, in Figure 9 it is only apparent on the right side of the plot.
- If we also recall Figure 7 we can see that the sawtooth characteristic is not at all apparent in this plot. Hence the effect is probably not caused by an underlying factor modulated by the MUI value.

Figure 27: 2nd to 5th order curve fits, relating to the curve fit analysis in Chapter 3.

The somewhat erratic shape of the plots can probably be explained by the low signal-to-noise ratio of the item measuring the concept. Because of the nature of this experiment we had to use extremely short and simple questionnaires to even attain the sample size we got and to be able to carry out the analysis and answer our research question. If we could have used a more sophisticated questionnaire and have gotten the same sample size the plots might be a lot smoother and clearer because of the higher signal-to-noise ratio.

7.1.2 Curve fit and sample size

Although in absolute terms we had a large participant pool for a BCI-based study, the sample size per condition was still on the low side for psychometric analysis of the data. As in any result there is a margin of error for the curve fitting. But as all of higher order curves show the same characteristic, id est, the deviation from the linear trend for higher values of the 'Control' dimension, one can only conclude that the data itself undeniably shows this characteristic. When we have a look at Figure 27 we can see the second, third, fourth and fifth order curve

Linear	
R^2	0,291
dof	155
p value	<0.001
Quadratic	
R^2	0,336
dof	154
p value	<0.001
p value (increase in r^2)	<0.001
Qubic	
R^2	0,349
dof	153
p value	<0.001
p value (increase in r^2)	0.0265
Quartic	
R^2	0,356
dof	154
p value	<0.001
p value (increase in r^2)	0.073

Table 14: F-test of curve fit.

fits. The third, fourth and fifth order plots all show the same characteristic for the higher values of control.

To further ground these findings we did additional F-tests to test for significance of the increase in R^2 value with the degrees of freedom taken into account. The results of these test are summarized in Table 14.

The F-tests were performed by calculating the r^2 value for the regressions and increase thereof between the n^{th} order and $n - 1^{\text{th}}$ order and the degrees of freedom for both regressions respectively as parameters.

The results show a significantly increased explained variability going from a first to second and second to third order curve fit. The fourth order fit still explains the variability better but the F-test shows this increase is not significant ($p > 0.05$). The results of this test are in agreement with the test based on Aikaike Information Criterion (AIC) as performed in chapter 3.

When looking at the number of data points (Table 2) for the higher values of 'Control', one can also conclude that this effect is not based on outliers. However, the optimum in the curve is not at all an exact results and is probably heavily dependent on various factors such as sample size, design of the game levels and complexity of the game characteristics as also explained in Section 3.6 of Chapter 3. At this point our results are only to be considered as a step in the direction of forming a hypothesis. Before we can make more definitive conclusions, a bigger sample size and especially more variation in the design, the type and type of interaction of the game need to be considered.

7.2 DISCUSSION OF CHAPTER 4

In the experimental study we conducted for chapter 4 we looked at the differences between actual and imagined movement regarding accuracy of the two modalities but also in user experience. To evaluate the user experience in both modalities we modified an existing questionnaire and assessed to what extent it still was numerically consistent by means of the Cronbach's Alpha and a Principal Components Analysis (PCA) on the items in each scale. We designed the study to be of the randomized crossover type to exclude confounders such as learning effects. One might argue that a proper multivariate analysis such as a PCA requires a large sample size to get meaning out of the data. Often a number like $n=200$ or 20 times the number of items in the questionnaire is asked for. However, as the nature of this research is such, that questionnaires cannot be administered in parallel to large groups and every participant has to be physically be located inside the lab for the duration of the experiment (in the order of 2 to 3 hours) a sample size of 20 is already quite a hard task. Furthermore, a sample size of 200 or more is not always needed for the data to make sense in a PCA analysis. There are various other fields of research for which it is hard to get a large sample size while PCA analyses still have prove to make sense and are methodologically valid to use [19].

From a mathematical point of view a PCA mainly consists of estimating the covariance matrix. This covariance matrix has

$$(n * (n + 1))/2 \quad (6)$$

degrees of freedom. Therefore, with only 3 variables such as most scales in our questionnaire it is very feasible to make a

proper estimate of the covariance matrix with only 20 samples. Filling in 3 for Equation 6 gives us:

$$(3 * (3 + 1))/2 = 6 \quad (7)$$

In this case the number of samples (20) is more than three times bigger than our degrees of freedom. Therefore covariance estimation should not pose a problem for the cases where the number of variables = 3.

One might argue to what extent the paired observations of one participant in two different conditions are valid observations and can be used in the estimation of the same covariance matrix (and thus be used in the same PCA analysis). The paired observations are likely to be correlated as some aspects do not change between conditions. At the very least, they provide a more stable estimation of the covariation matrix because they are highly correlated. This is actually one of the strong points of the randomized crossover design of the study.

To further quantify to what extent the data is admissible to performing a PCA analysis on the data from this study, we performed more analyses on all 40 responses per item on all scales. More specifically, we performed Bartlett's test of Sphericity and Kaiser-Meyer-Olkin's (KMO) test which is a measure for sampling adequacy. In the case of a low number of samples Bartlett's test should result in significant values $p \leq 0.05$ and the KMO test should result in values ≥ 0.5 . See also [19], [20] and [32] for a more in-depth and mathematical discussion of the issue of using a low number of samples when performing a PCA.

The results of the post hoc tests are shown in Table 15. One can see that for all scales Bartlett's test is very significant ($p \leq 0.004$) and except for the Negative Experiences scale, all results for the KMO test are > 0.541 which complies to the constraints that were set for both tests. If we recall the discussion of the results in Chapter 4 we see that our initial doubtful findings on the Negative Experiences scale are further justified by this test that shows that this scale is not useful for this study and should be ignored. Perhaps when one repeats this study in a similar fashion but with a higher number of samples, we can find out whether we can confirm that this scale is constructed out of too many different aspects, thus resulting in a multi-dimensional scale not accurately measuring one concept, or is biased by certain factors skewing the scale in a certain direction. For the interested reader we present the scree plots showing the loading of the factors onto the principal components in Appendix ??.

Scale	Bartlett's test	KMO test
Alertness	$p < 0.001$	0.585
Goals	$p < 0.001$	0.541
Control	$p < 0.001$	0.665
Positive Experiences	$p < 0.001$	0.798
Challenge	$p < 0.001$	0.711
Fatigue	$p < 0.001$	0.666
Immersion	$p = 0.004$	0.620
Negative Experiences	$p < 0.001$	0.471

Table 15: Bartlett's test of sphericity and Kaiser-Meyer-Olkin test for the questionnaires scales used in Chapter 4

7.3 DISCUSSION OF CHAPTER 6

In Chapter 6 we did an experiment ($n=13$) with a qualitative nature. Part of the objective was to gather insights as to *why* users like or dislike certain aspects. In the experimental study we did for this chapter we also gathered some quantitative data. As we can see from the SUS questionnaire we can see a score of 66.2 which is only 1.8 points less than the average score for 500 studies. Normalizing the score 66.2 would put it at the 50% rank. For a system incorporating an EEG headset, which is not a mainstream technological device, this is quite promising. However, the background of the participants might have helped us raise the score a bit here. Most of the participants were students at a technologically oriented university. Usually, technology students have a higher acceptance of and more interest in new technologies. However, there also were some participants who drew, sketched or painted on a daily basis, for educational/professional or leisure purposes. It is interesting to see that a system incorporating a complex device such as an EEG device can still be so easily accepted by novel users. What we can learn from this chapter is that a new technology such as an EEG device or a BCI control mechanism can be accepted, as long as it is implemented in an intuitive manner.

CONCLUSIONS

In this section we will first have another look at the research question and at the studies we did to answer this question.

8.1 ANSWERING THE RESEARCH QUESTION

For this section it is important to recall our research question: what influence has the accuracy of a BCI on the user experience? We conducted several studies in which this theme was central.

In the first study we evaluated the relationship between accuracy, perceived control and whether the user had fun. In the second study we let the user do two different BCI paradigms in a game and evaluated both on accuracy and user experience. In the third and last study we evaluated a game with and without the use of a BCI. All studies were conducted with different (BCI-based) games. In the next sections we will discuss the findings from the studies.

8.1.1 *Control*

From the study in Chapter 3 we saw that users have a quite decent understanding of how much control they have over a character in a game. Although we hypothesized an underestimation in the lower part (it seems that a bit higher than random level seems still random) and an overestimation in the higher parts (almost all of the time seems all of the time) of the control scale, we could not validate this hypothesis.

When looking at the relationship between control and fun we see two interesting effects: the first is that low control is related to a low amount of fun. The linear relationship goes through the control scale until a point where it settles and at higher levels of control even decreases. Another way to formulate this is that at a certain point, the level of control is not anymore the deciding factor for the amount of fun experienced by the user. Other factors that influence UX come into play. Another reasonable explanation for this effect is that the added uncertainty of the imperfect control adds a certain challenge to the game. This chapter shows that BCI control can be an added challenge

for the user, a challenge which they can try to master just like mastering other abilities is needed for successfully completing a game.

8.1.2 *UX evaluation of a BCI game with direct control*

In Chapter 4 we saw that actual movement provides a better recognition accuracy than imagined movement in a BCI based game. We looked at UX more in-depth than in the first study by taking more extensive measures by means of a more detailed questionnaire. Through the more detailed and numerically consistent questionnaire we were able to better analyse the various aspects of UX in BCI games and how they influence each other. Participants indicated that they found more challenge in playing with the imagined movement modality. At the same time there was no significant difference between both modalities regarding positive and negative experiences. Combining these results with the results from the first study we can deduce that a lower level of control is not always a bad thing. The analysis of the data from this study indicates that participants were not experiencing significantly more negative aspects (e.g. frustration) although we saw that it was more mentally tiring to use imagined movement. Of course a novelty effect while using the imagined movement cannot be excluded, but remarks made by the participants generally pointed towards a preference for imagined movement. Although imagined movement gave a lower level of control and was more mentally tiring, participants really liked it, indicating that the level of control is not the critical factor in BCI games.

8.1.3 *UX evaluation of an additional BCI control in a popular game*

In Chapter 5 we set out to answer the question whether BCI control can be an added value to a modern video game. We recruited experienced WoW players and participants without any experience with WoW. We compared the actual duration of the two conditions (with and without BCI control) and our analysis showed they equally liked to play with and without BCI control. The questionnaires provided us with some more insights. Control and involvement was lower in the BCI condition, but the amount of fun was not significantly different in the two conditions. Several reasons might explain this: participants found the BCI a novel and interesting modality to play with, it provided them with a challenge and/or they could cope with or anticipate on the lack of control. Whether this novelty effect wears off

and user lose interest remains to be seen. Even without perfect control the addition of BCI control could make a game that gets users curious and interested in this new modality.

8.1.4 *BCI control in a multimodal application*

In Chapter 6 we implemented and evaluated an application that incorporated BCI control, eye blink detection and head movement. The evaluation was mainly done in a qualitative manner and in an iterative process so we could improve upon the application. There are no existing guidelines for this kind of applications yet and user centered, iterative design is a good way to design such an application. From the results in the final study we could see that most users were not cognitively overloaded with the combination of three modalities. Actually, switching between modalities in the way we implemented it in BrainBrush might have actually helped to lower the total workload.

8.2 OVERALL CONCLUSIONS

Overall the results from these chapters show that BCI accuracy is only important up to a certain point. (See chapters 3 & 4) If the BCI is even less accurate, one solution is to not base the interaction solely on the BCI control so that users can compensate for it by other means (Chapter 4). Another option is to give them the opportunity to learn to control the BCI, while using conventional control as an added challenge in the game (Chapter 5). Another possibility is to make the application more multimodal (as in Chapter 6). In this chapter we found out that BCI control can be used in combination with other modalities to make a hands-free painting application.

8.2.1 *Passive BCIs*

The results from chapter 4 showed that even when a BCI provides much lower control than a keyboard or mouse alternative, it is still possible to make an enjoyable game. For the short term acceptance of BCIs passive BCIs (as defined by Zander et al. [148]) are probably the most viable. One can make use of an added input channel, but does not have to rely on it for controlling the application. Especially in games, hybrid approaches in which the application monitors the brain activity, provides feedback but doesn't put the user in complete (lack of) control.

This way, the user can get used to the game and gradually gain control without getting frustrated.

8.2.2 *BCI for direct control*

Even the state-of-the-art studies do not provide us with a BCI with an accuracy high enough to directly replace a keyboard and/or mouse for healthy users. We should therefore be inventive in applying BCI in games and software in general to provide an advantage or challenge over just using a keyboard and/or mouse. Besides the noise in the signals coming from the EEG one possible reason might be that it can be hard to do something one never did before, for example imagining a movement. Various studies have shown that training has a positive impact on the accuracy one can achieve with a BCI. Often these studies require training in the order of a couple of hours per week for a month or two. It would be interesting to see what would happen if more longitudinal studies would occur. To compare, how long has one to train before one can properly put a keyboard to use at an acceptable speed? A big difference between a keyboard and a BCI is that a keyboard is much more intuitive. The key that is typed is what is shown on screen, if one typed the wrong letter, the screen immediately shows the mistake that was made. With a BCI things are, at least up until now, not so transparent. When e.g. an imagined movement is not detected or wrongly detected, does the error stem from the user? Or from the system? And in the case of the system, was it an artefact from the EEG? Or was the classifier not trained well enough? The challenge for future research on BCI for direct control is to improve the transparency of the system to the user and to determine whether the user made a mistake or the system. As HCI researchers we need to make an intuitive interactive system for users to get the proper feedback and motivate them to train for longer periods of time.

8.3 LIMITATIONS AND FUTURE WORK

It might be a cliché, but every study a researcher conducts generates a lot of data to validate his hypothesis, but there are always some findings that do not make sense, at least at first. Sometimes one can do further analyses on the data, discuss with colleagues and come up with a logical explanation. But it can also be the case that some findings just generate a lot of new questions for new studies. In this section we will elaborate on where the limitations in our study stop us from drawing further con-

clusions and our data feeds us with new hypothesis that warrant new studies to answer the questions that were raised.

8.3.1 *The games*

In every study and experiment choices have to be made depending on the number of measures to take, amount of time or effort needed. Also in the studies in this thesis a choice had to be made that every experiment was evaluated with one game. With only one game it is somewhat hard to generalize the results, but at least the results can be relevant or the amount of time needed to evaluate is feasible. This is one of the problems within HCI research in general. In prototype development certain design choices are made (type of game, interaction style, aesthetical choices, etc.). All these choices can have an influence on the outcome of the evaluation. As the number of choices and resulting number of combinations of choices is far bigger than the possibilities for evaluation of these choices it is very hard to be conclusive about any result when following this reasoning. However, one needs to be careful not to generalize too much and make a well argued reasoning why certain findings could be generalized and useful for other studies as well. We think that the effects we found cannot be directly generalized and that more studies have to be done to study the hypotheses raised by our data. At least we are convinced that similar findings will be generated by those studies.

8.3.2 *BCI hardware*

In the studies for the third and fourth chapter we used different EEG hardware to acquire the signal for our BCI. We used the BioSemi Active 2 in the study for the third chapter and the Emotiv EPOC for the study in the fourth chapter. Because the BioSemi system is a professional, wired, gelled EEG cap and the Emotiv a water-based, wireless system this might or might not have had an influence on a couple of aspects we couldn't control for:

- Participants might be weary of participating in an experiment in which they have to strap on a cap with connectors that will be electrically connected to their scalp by the use of gel.
- Participants might like the design of the Emotiv EPOC more, because it was smaller, shiny black and wireless. Also the previous issue of the gel of course comes into play.

- On the other hand, various anecdotal findings showed that users feel some pressure on their scalp from the EPOC that over time might be painful.
- Apart from possible usability issues outlined above, there is a difference in signal quality from both EEG hardware sets as found by Nijboer et al. [96].

Although we could not control these variables, they were probably not detrimental to our findings. In both studies we used the EEG sets, the participants were wearing the same EEG set in both conditions. Especially in chapter 4 this was an important choice as to not measure the effect of wearing or not wearing an EEG set while playing the game.

Regarding the difference in signal quality, we don't think the lower signal quality of the EPOC influenced the findings to a large extent. As shown in the validation experiment in chapter 4 we saw that some participants were able to achieve rather high accuracies using the Emotiv EPOC.

8.3.3 *Participants*

As in any experiment but especially in BCI experiments it is a lot of work to recruit participants. As BCI experiments often involve quite a long setup time with the participant required at the experiment location, the time invested by the participant is rather high, usually not less than an hour. This of course biases the 'kind' of participants that participate in these studies. Possible participants that are not open to such kind of technological hardware directly attached to their body or participants not willing to invest so much time in an experiment are the ones that are often missing from these experiments. What kind of influence this has on our results we cannot accurately assess as this is a very hard practical problem. The Emotiv EPOC helps solve this issue to a certain extent. The setup time is much lower than with professional gel-based systems. Also the burden on the participant is much less than with conventional systems. However, as mentioned before the performance is also lower. There will probably always be a trade-off between faster and easier to setup systems and more accuracy from more elaborate systems. The number of participants used in the studies vary, in the first study we had a lot of participants, but also a lot of 'conditions'. Although we had enough participants per condition and could paint a clear picture regarding the relation between fun and accuracy, with a data file with two to three times the number of participants/data points we would be able to make a much more convincing case. In the second study we

had 20 participants, evenly spread between males and females and with a spread in age as well. The spread over age was not very even though and a higher number of subjects might have revealed a relation between age and accuracy. Although we had decent number of subjects in the last study (42) and we had experienced as well as inexperienced WoW players participate we could not find many significant differences between these groups.

Part IV

APPENDICES

APPENDICES

A.1 QUESTIONNAIRE USED IN CHAPTER 3

Control:

1. I had the feeling the hamster did what I wanted it to do
2. I had the feeling the computer was following my commands

Frustration:

3. I was frustrated while playing the game

Fun:

4. Playing the game was fun

Empowerment:

5. During the game: I felt empowered

Satisfaction:

6. During the game: I felt satisfied

A.2 QUESTIONNAIRE USED IN CHAPTER 5

awow.pdf

Questionnaire

Subject ID: ____ Group: ____

1. Please give an estimation of the time you played in minutes and seconds: ____m ____s

The following questions use a 7 point scale. Please indicate your opinion by marking the corresponding answer. There are no wrong or right answers. It is about *your* experience.

2. I felt a sense of achievement.	not at all 0 0 0 0 0 0 0 very much
3. I did things I liked to do but could not do in real life.	not at all 0 0 0 0 0 0 0 very much
4. I socialized with other players.	not at all 0 0 0 0 0 0 0 very much
5. How much were you able to control events?	not at all 0 0 0 0 0 0 0 very much
6. How responsive was the environment to actions that you initiated (or performed)?	not at all 0 0 0 0 0 0 0 very much
7. How natural did your interactions with the environment seem?	not at all 0 0 0 0 0 0 0 very much
8. How completely were <i>all</i> of your senses engaged?	not at all 0 0 0 0 0 0 0 very much
9. How much did the visual aspects of the environment involve you?	not at all 0 0 0 0 0 0 0 very much
10. How natural was the mechanism which controlled movement through the environment?	not at all 0 0 0 0 0 0 0 very much
11. How aware were you of events occurring in the real world around you?	not at all 0 0 0 0 0 0 0 very much
12. How aware were you of your display and control devices?	not at all 0 0 0 0 0 0 0 very much
13. How compelling was your sense of objects moving through space?	not at all 0 0 0 0 0 0 0 very much
14. How inconsistent or disconnected was the information coming from your various senses?	not at all 0 0 0 0 0 0 0 very much
15. How much did your experiences in the virtual environment seem consistent with your real-world experiences?	not at all 0 0 0 0 0 0 0 very much
16. Were you able to anticipate what would happen next in response to the actions that you performed?	not at all 0 0 0 0 0 0 0 very much
17. How completely were you able to actively survey or search the environment using vision?	not at all 0 0 0 0 0 0 0 very much
18. How compelling was your sense of moving around inside the virtual environment?	not at all 0 0 0 0 0 0 0 very much
19. To what degree did you feel confused or disoriented at the end of the experimental session?	not at all 0 0 0 0 0 0 0 very much
20. How involved were you in the virtual environment experience?	not at all 0 0 0 0 0 0 0 very much
21. How distracting was the control mechanism?	not at all 0 0 0 0 0 0 0 very much
22. How much delay did you experience between your actions and expected outcomes?	not at all 0 0 0 0 0 0 0 very much
23. How quickly did you adjust to the virtual environment experience?	not at all 0 0 0 0 0 0 0 very much
24. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?	not at all 0 0 0 0 0 0 0 very much
25. How much did the control devices interfere with the performance of your activities?	not at all 0 0 0 0 0 0 0 very much
26. How well could you concentrate on these activities rather than on the mechanisms used to perform those activities?	not at all 0 0 0 0 0 0 0 very much
27. Did you learn new techniques that enabled you to improve your performance?	not at all 0 0 0 0 0 0 0 very much
28. Were you involved in the experimental task to the extent that you lost track of time?	not at all 0 0 0 0 0 0 0 very much
29. How challenging was the game for you?	not at all 0 0 0 0 0 0 0 very much
30. I was in control of shape shifting from bear to human.	not at all 0 0 0 0 0 0 0 very much
31. I was in control of shape shifting from human to bear.	not at all 0 0 0 0 0 0 0 very much
32. Over time I got better in shape shifting.	not at all 0 0 0 0 0 0 0 very much
33. Playing the game was fun.	not at all 0 0 0 0 0 0 0 very much
34. Shape shifting was intuitive.	not at all 0 0 0 0 0 0 0 very much
35. I had the feeling the system knew when I wanted to shift shape.	not at all 0 0 0 0 0 0 0 very much
36. I had the feeling the system had control over me when I wanted to shift shape.	not at all 0 0 0 0 0 0 0 very much
37. I could focus well on the shape shifting.	not at all 0 0 0 0 0 0 0 very much
38. It took a lot of effort to shift shapes.	not at all 0 0 0 0 0 0 0 very much
39. Over time shape shifting became easier.	not at all 0 0 0 0 0 0 0 very much
40. Playing the game was tiring.	not at all 0 0 0 0 0 0 0 very much
41. I could concentrate well on the game.	not at all 0 0 0 0 0 0 0 very much

A.3 QUESTIONNAIRE USED IN CHAPTER 6

SUS - System Usability Scale

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

BIBLIOGRAPHY

- [1] Demonstration of Alpha WoW at AISB 2009. Edinburgh, Scotland, 2009.
- [2] Demonstration of Alpha WoW at TEDxAmsterdam 2009, 2009. URL <http://www.tedxamsterdam.com/2009/alphaworldofwarcraft/>.
- [3] Future BNCI: A Roadmap for Future Directions in Brain / Neuronal Computer Interaction Research, 2012. URL http://future-bnci.org/images/stories/Future_BNCI_Roadmap.pdf.
- [4] H. Akaike. Information theory and an extension of the maximum likelihood principle. In *Second international symposium on information theory*, pages 267–281. Academiai Kiado, 1973.
- [5] L.G. Allan and H.M. Jenkins. The judgment of contingency and the nature of the response alternatives. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 34(1):1–11, 1980. doi: 10.1037/h0081013.
- [6] Fabio Aloise, Pietro Aricò, Francesca Schettini, Elena Lucano, Serenella Salinari, Fabio Babiloni, Donatella Mattia, and Febo Cincotti. Can the P300-based BCI training affect the ERPs? *International Journal of Bioelectromagnetism*, 13(3):148–149, 2011.
- [7] A. Barbero and M. Grosse-Wentrup. Biased feedback in brain-computer interfaces. *Journal of NeuroEngineering and Rehabilitation*, 7(1):34, 2010. ISSN 1743-0003.
- [8] Robert J. Barry, Adam R. Larke, Stuart J. Johnstone, Christopher A. Magee, and Jacqueline A. Rushby. Eeg differences between eyes-closed and eyes-open resting conditions. *Clinical neurophysiology*, 118(12):2765–2773, 2007.
- [9] E. Başar, C. Başar-Eroglu, R. Parnefjord, E. Rahn, and M. Schürmann. Evoked potentials: Ensembles of brain induced rhythmicities in the alpha, theta and gamma ranges. *Induced Rhythms in the Brain*, pages 155–181, 1992.
- [10] E. Başar, M. Schürmann, C. Başar-Eroglu, and S. Karakaş. Alpha oscillations in brain functioning: an integrative theory. *International Journal of Psychophysiology*, 26(1-3):5 –

- 29, 1997. ISSN 0167-8760. doi: 10.1016/S0167-8760(97)00753-8.
- [11] Erol Başar, Canan Başar-Eroglu, Sirel Karakaş, and Martin Schürmann. Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International Journal of Psychophysiology*, 39(2-3):241 – 248, 2001. ISSN 0167-8760. doi: 10.1016/S0167-8760(00)00145-8.
- [12] G. Bauer, F. Gerstenbrand, and E. Rimpl. Varieties of the locked-in syndrome. *Journal of Neurology*, 221(2):77–91, 1979.
- [13] N. Birbaumer, T. Elbert, AG Canavan, and B. Rockstroh. Slow potentials of the cerebral cortex and behavior. *Physiological Reviews*, 70(1):1–41, 1990.
- [14] Alethea L. Blackler, Vesna Popovic, and Douglas P. Mahar. Towards a design methodology for applying intuitive interaction. In Ken Friedman, Terence Love, Eduardo Côté-Real, and Chris Rust, editors, *WonderGround: 2006 Design Research Society International Conference*, Lisbon, 2006.
- [15] Margaret M. Bradley and Peter J. Lang. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, 1994.
- [16] Jeanne H. Brockmyer, Christine M. Fox, Kathleen A. Curtiss, Evan McBroom, Kimberly M. Burkhart, and Jacquelyn N. Pidruzny. The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, 45(4):624–634, 2009.
- [17] John Brooke. SUS - A quick and dirty usability scale. In Patrick W. Jordan, B. Thomas, Ian Lyall McClelland, and Bernard Weerdmeester, editors, *Usability Evaluation In Industry*, pages 189–194. CRC Press, 1996.
- [18] Ivo Brugman. A Multimodal Interactive System for Creative Expressions, August 2012. URL <http://essay.utwente.nl/61988/>.
- [19] Sergey V. Budaev. Using principal components and factor analysis in animal behaviour research: Caveats and guidelines. *Ethology*, 116(5):472–480, 2010. ISSN 1439-0310. doi: 10.1111/j.1439-0310.2010.01758.x. URL <http://dx.doi.org/10.1111/j.1439-0310.2010.01758.x>.

- [20] Sergey V. Budaev. Multivariate methods and small sample size: combining with small effect size. *Manuscript*, 2011. URL <http://cogprints.org/8672/1/dochtermann-jenkins-response.pdf>.
- [21] Jr. Byron Wm. Brown. The crossover experiment for clinical trials. *Biometrics*, 36(1):69–79, 1980. ISSN 0006341X. URL <http://www.jstor.org/stable/2530496>.
- [22] Andrew Campbell, Tanzeem Choudhury, Shaohan Hu, Hong Lu, Matthew K. Mukerjee, Mashfiqui Rabbi, and Rajeev D.S. Raizada. NeuroPhone: brain-mobile phone interface using a wireless EEG headset. In *Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds*, MobiHeld '10, pages 3–8, New York, NY, USA, 2010. ACM. ISBN 978-1-4503-0197-8. doi: 10.1145/1851322.1851326.
- [23] J. Cantero, M. Atienza, C. Gómez, and R. Salas. Spectral structure and brain mapping of human alpha activities in different arousal states. *Neuropsychobiology*, 39(2):110–116, 1999. doi: 10.1159/000026569.
- [24] 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.
- [25] B. Chambayil, R. Singla, and R. Jha. Virtual keyboard BCI using Eye blinks in EEG. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2010 IEEE 6th International Conference on*, pages 466 –470, oct. 2010. doi: 10.1109/WIMOB.2010.5645025.
- [26] J.M. Cortina. What is coefficient Alpha? an examination of theory and applications. *Journal of Applied Psychology*, 78(1):98–104, 1993.
- [27] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, second edition, July 2006.
- [28] L.J. Cronbach. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3):297–334, 1951.
- [29] Mihaly Csikszentmihalyi. *Flow: The psychology of optimal experience*. New York: Harper & Row, 1990.
- [30] C. Davies. *Landscape, Earth, Body, Being, Space, and Time in the Immersive Virtual Environments Osmose and Ephémère*. The MIT Press, Cambridge, USA, 2003.

- [31] L. Deng and X. Huang. Challenges in adopting speech recognition. *Communications of the ACM*, 47(1):69–75, 2004.
- [32] Ned A. Dochtermann and Stephen H. Jenkins. Multivariate methods and small sample sizes. *Ethology*, 117(2):95–101, 2011. ISSN 1439-0310. doi: 10.1111/j.1439-0310.2010.01846.x. URL <http://dx.doi.org/10.1111/j.1439-0310.2010.01846.x>.
- [33] M Dupee and P Werthner. Managing the stress response: The use of biofeedback and neurofeedback with olympic athletes. *Biofeedback: Fall 2011*, 39(3):92–94, 2011.
- [34] G. Evans and P. Blenkhorn. A head operated joystick – experience with use. In *Proceedings of the CSUN conference on technology and persons with disabilities*, California State University, Northridge, 1999.
- [35] Luciano Fadiga, Laila Craighero, and Etienne Olivier. Human motor cortex excitability during the perception of others’ action. *Current Opinion in Neurobiology*, 15(2):213 – 218, 2005. ISSN 0959-4388. doi: 10.1016/j.conb.2005.03.013. URL <http://www.sciencedirect.com/science/article/pii/S0959438805000450>.
- [36] S.H. Fairclough. Fundamentals of physiological computing. *Interacting with computers*, 21(1-2):133–145, 2009. ISSN 0953-5438.
- [37] L.A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70(6):510 – 523, 1988. ISSN 0013-4694.
- [38] L.A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70(6):510 – 523, 1988. ISSN 0013-4694. doi: 10.1016/0013-4694(88)90149-6.
- [39] Mehrdad Fatourehchi, Ali Bashashati, Rabab K. Ward, and Gary E. Birch. EMG and EOG artifacts in brain computer interface systems: A survey. *Clinical Neurophysiology*, 118(3):480–494, 2007.
- [40] Tom Fawcett. An introduction to ROC analysis. 27(8): 861–874, 2005.
- [41] P. Fitts. The information capacity of the human motor system in controlling the amplitude of move-

- ment. *Journal of Experimental Psychology*, 47(6):381–391, 1954. URL <http://psycnet.apa.org/index.cfm?fa=fulltext.journal&jcode=xge&vol=121&issue=3&format=html&page=262&expand=1>.
- [42] D. Friedman, R. Leeb, C. Guger, A. Steed, G. Pfurtscheller, and M. Slater. Navigating virtual reality by thought: What is it like? *Presence: Teleoperators and Virtual Environments*, 16(1):100–110, 2007. ISSN 1054-7460.
- [43] A. Furdea, S. Halder, D.J. Krusienski, D. Bross, F. Nijboer, N. Birbaumer, and A. KÄEbler. An auditory oddball (p300) spelling system for brain-computer interfaces. *Psychophysiology*, 46(3):617–625, 2009. ISSN 1469-8986. doi: 10.1111/j.1469-8986.2008.00783.x. URL <http://dx.doi.org/10.1111/j.1469-8986.2008.00783.x>.
- [44] Future BNCI. a Roadmap for Future Directions in Brain / Neuronal Computer Interactions Research, 2012. URL http://future-bnci.org/images/stories/Future_BNCI_Roadmap.pdf.
- [45] F. Galán, M. Nuttin, E. Lew, P. W. Ferrez, G. Vanacker, J. Philips, and J. del R. Millán. A Brain-Actuated Wheelchair: Asynchronous and Non-Invasive Brain-Computer Interfaces for Continuous Control of Robots. *Clinical Neurophysiology*, 119(9):2159–2169, 2008. ISSN 1388-2457.
- [46] 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.
- [47] C. Groenegrass, C. Holzner, C. Guger, and M. Slater. Effects of P300-based BCI use on reported presence in a virtual environment. *Presence: Teleoperators and Virtual Environments*, 19(1):1–11, 2010. ISSN 1054-7460.
- [48] Christoph Guger, Shahab Daban, Eric Sellers, Clemens Holzner, Gunther Krausz, Roberta Carabalona, Furio Gramatica, and Guenter Edlinger. "how many people are able to control a p300-based brain-computer interface (bci)?" *Neuroscience Letters*, 462(1):94 – 98, 2009. ISSN 0304-3940.
- [49] H. Gürkök, G. Hakvoort, and M. Poel. Evaluating user experience with respect to user expectations in brain-computer interface games. In G. R. Müller-Putz, R. Scherer, M. Billinger, A. Krelinger, V. Kaiser, and C. Neuper, editors, *Proceedings of the 5th International*

- Brain-Computer Interface Conference, BCI 2011, Graz, Austria*, pages 348–351, Graz, Austria, September 2011. Verlag der Technischen Universität Graz.
- [50] H. Gürkök, G. Hakvoort, and M. Poel. Evaluating user experience in a selection based brain-computer interface game: A comparative study. In J. Coutinho Anacleto, S. Fels, N. Graham, B. Kapralos, M. Saif El-Nasr, and K. Stanley, editors, *Proceedings of the 10th International Conference on Entertainment Computing (ICEC 2011)*, Vancouver, BC, Canada, volume 6972 of *Lecture Notes in Computer Science*, pages 77–88, Berlin, October 2011. Springer Verlag.
  - [51] H. Gürkök, D. Plass-Oude Bos, B. L. A. van de Laar, F. Nijboer, and A. Nijholt. User Experience Evaluation in BCI: Filling the Gap. *International Journal of Bioelectromagnetism*, 13(1):54–55, July 2011.
  - [52] Hayrettin Gürkök, Gido Hakvoort, and Mannes Poel. Modality switching and performance in a thought and speech controlled computer game. In *Proceedings of ICMI 2011*. ACM, New York, NY, USA, 2011.
  - [53] G. Hakvoort, H. Gürkök, D. Plass-Oude Bos, M. Obbink, and M. Poel. Measuring immersion and affect in a brain-computer interface game. In *Human-Computer Interaction - INTERACT 2011, 13th IFIP TC 13 International Conference*. Springer, 2011. To appear.
  - [54] S.G. Hart and L.E. Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52:139–183, 1988.
  - [55] S.I. Hjelm. Research+ design: the making of brainball. *interactions*, 10(1):26–34, 2003.
  - [56] S.I. Hjelm and C. Browall. Brainball-using brain activity for cool competition. In *Proceedings of NordiCHI*, pages 177–188, 2000.
  - [57] Ulrich Hoffmann, Jean Marc Vesin, Touradj Ebrahimi, and Karin Diserens. An efficient P300-based brain-computer interface for disabled subjects. *Journal of Neuroscience Methods*, 167(1):115 – 125, 2008. ISSN 0165-0270.
  - [58] E. M. Holz, T. Kaufmann, D. Franz, A. Hösle, and A. Kübler. Brain Drawing: First Evaluation Results. In *Proceedings of the 3rd TOBI Workshop*, pages 109–110, Würzburg, Germany, March 2012.

- [59] J.E. Huggins, P.A. Wren, and K.L. Gruis. What would brain-computer interface users want? opinions and priorities of potential users with amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis*, (0):1–8, 2011.
- [60] W. IJsselsteijn, I. Bierhoff, and Y. Slangen-de Kort. Duration estimation and presence. In *Presence 2001 conference*, 2001. doi: 10.1.1.19.5785.
- [61] W. IJsselsteijn, Y. de Kort, K. Poels, A. Jurgelionis, and F. Bellotti. Characterising and Measuring User Experiences in Digital Games. *International Conference on Advances in Computer Entertainment Technology*, pages 1–4, 2007.
- [62] WA IJsselsteijn, YAW de Kort, and K. Poels. The Game Experience Questionnaire: Development of a self-report measure to assess the psychological impact of digital games. Manuscript submitted for publication, 2008.
- [63] R.J.K. Jacob and K.S. Karn. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind*, 2(3):573–605, 2003.
- [64] Robert J. K. Jacob. Eye movement-based human-computer interaction techniques: Toward non-command interfaces. *Advances in Human-Computer Interaction*, 1993.
- [65] M. Jeannerod. The representing brain neural correlates of motor intention and imagery. *Behavioural Brain Sciences*, 17:187–245, 1994.
- [66] Charlene Jennett, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66: 641–661, 2008.
- [67] Turan A. Kayagil, Ou Bai, Peter Lin, Steve Furlani, Sherry Vorbach, and Mark 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.
- [68] S. Kleih, F. Nijboer, S. Halder, and A. Kübler. Motivation modulates the p300 amplitude during brain-computer interface use. *Clinical Neurophysiology*, 121(7):1023–1031, 2010.
- [69] S. Kleih, F. Nijboer, S. Halder, and A. Kübler. Motivation modulates the p300 amplitude during brain-computer

- interface use. *Clinical Neurophysiology*, 121(7):1023–1031, 2010.
- [70] 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.
- [71] 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.
- [72] 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.
- [73] Dean J Krusienski, Eric W Sellers, François Cabestaing, Sabri Bayoudh, Dennis J McFarland, Theresa M Vaughan, and Jonathan R Wolpaw. A comparison of classification techniques for the P300 Speller. *Journal of Neural Engineering*, 3(4):299, 2006.
- [74] E.J. Langer. The illusion of control. *Journal of personality and social psychology*, 32(2):311–328, 1975.
- [75] Anatole Lécuyer, Fabien Lotte, Richard B. Reilly, Robert Leeb, Michitaka Hirose, and Mel Slater. Brain-computer interfaces, virtual reality, and videogames. *IEEE Computer*, 41(10):66–72, 2008.
- [76] R. Leeb, D. Friedman, G. Müller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller. Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic. *Intelligent Neuroscience*, 2007(2):1–12, 2007. ISSN 1687-5265. doi: <http://dx.doi.org/10.1155/2007/79642>.
- [77] 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. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 15(4):473–482, 2007.
- [78] 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.

- IEEE Trans. on Neural Systems and Rehab. Eng.*, 15(4):473–482, 2007.
- [79] 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.
  - [80] G. Lightbody, M. Ware, P. McCullagh, M.D. Mulvenna, E. Thomson, S. Martin, D. Todd, V.C. Medina, and S.C. Martinez. A user centred approach for developing brain-computer interfaces. In *2010 4th International Conference on Pervasive Computing Technologies for Healthcare*, pages 1–8, Piscataway, NJ, USA, 2010. IEEE.
  - [81] Kip Ludwig, Rachel Miriani, Nicholas Langhals, Michael Joseph, David Anderson, and Daryl Kipke. Using a Common Average Reference to Improve Cortical Neuron Recordings From Microelectrode Arrays. *J Neurophysiol*, 101(3):1679–1689, 2009. doi: 10.1152/jn.90989.2008. URL <http://jn.physiology.org/cgi/content/abstract/101/3/1679>.
  - [82] Regan L. Mandryk, M. Stella Atkins, and Kori M. Inkpen. A continuous and objective evaluation of emotional experience with interactive play environments. In *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1027–1036, New York, NY, USA, 2006. ACM.
  - [83] Regan L. Mandryk, Kori M. Inkpen, and Thomas W. Calvert. Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & Information Technology*, 25:141–158, 2006.
  - [84] Samuel M. McClure, Jian Li, Damon Tomlin, Kim S. Cypert, Latané M. Montague, and P. Read Montague. Neural correlates of behavioral preference for culturally familiar drinks. *Neuron*, 44(2):379 – 387, 2004. ISSN 0896-6273. doi: 10.1016/j.neuron.2004.09.019.
  - [85] 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.
  - [86] 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.
- [87] Niamh McNamara and Jurek Kirakowski. Functionality, usability, and user experience: three areas of concern. *interactions*, 13:26–28, November 2006. ISSN 1072-5520. doi: <http://doi.acm.org/10.1145/1167948.1167972>. URL <http://doi.acm.org/10.1145/1167948.1167972>.
- [88] E. R. Miranda, W. L. Magee, J. J. Wilson, J. Eaton, and R. Palaniappan. Brain-Computer Music Interfacing (BCMI) From Basic Research to the Real World of Special Needs. *Music and Medicine*, 3(3):134–140, 2011.
- [89] C. Mohs, J. Hurtienne, J.H. Israel, A. Naumann, M.C. Kindsmüller, H.A. Meyer, and A. Pohlmeyer. IUUI–Intuitive Use of User Interfaces. *Usability Professionals*, 6, 2006.
- [90] M. Moore Jackson and R. Mappus. Applications for Brain-Computer Interfaces. In Anton Nijholt and Desney Tan, editors, *Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction*, pages 89–103. Springer-Verlag London Ltd., London, UK, 2010.
- [91] J. Münßinger, S. Halder, S. Kleih, A. Furdea, V. Raco, A. Hösle, and A. Kübler. Brain Painting: First Evaluation of a New Brain-Computer Interface Application with ALS-Patients and Healthy Volunteers. *Front Neurosci*, 4:182. doi: 10.3389/fnins.2010.00182, 2010.
- [92] J.I. Münßinger, S. Halder, S.C. Kleih, A. Furdea, V. Raco, A. Hösle, and A. Kübler. Brain painting: First evaluation of a new brain-computer interface application with als-patients and healthy volunteers. *Frontiers in Neuroprosthetics*, 4, 2010.
- [93] J. Nielsen. *Usability engineering*. Morgan Kaufmann, 1993.
- [94] F. Nijboer, N. Birbaumer, and A. Kübler. The influence of psychological state and motivation on brain-computer interface performance in patients with amyotrophic lateral sclerosis - a longitudinal study. *Frontiers in neuroprosthetics*, 4:1–13, July 2010. ISSN 1662-453X.
- [95] F. Nijboer, N. Birbaumer, and A. Kübler. The influence of psychological state and motivation on brain-computer interface performance in patients with amyotrophic lateral sclerosis – a longitudinal study. *Frontiers in Neuroscience*, 4:55, 2010.

- [96] F. Nijboer, B. van de Laar, S. Gerritsen, A. Nijholt, and M. Poel. Usability of Three Electroencephalogram Headsets for Brain-Computer Interfaces: a Within Subject Comparison. *Interacting with Computers*, 2015. In press.
- [97] A. Nijholt, D. Tan, B. Allison, J. Milan, and B. Graimann. Brain-computer interfaces for HCI and games. In *CHI '08 extended abstracts on Human factors in computing systems*, pages 3925–3928, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-012-X. doi: <http://doi.acm.org/10.1145/1358628.1358958>.
- [98] A. Nijholt, D. Tan, G. Pfurtscheller, C. Brunner, J.R. Milán, B. Allison, B. Graimann, F. Popescu, B. Blankertz, and K. R. Müller. Brain-Computer Interfacing for Intelligent Systems. *IEEE Intelligent Systems*, 23(3):72–79, 2008.
- [99] A. Nijholt, D.P.O. Bos, and B. Reuderink. Turning shortcomings into challenges: Brain-computer interfaces for games. *Journal of Entertainment Computing*, 1(2):85–94, 2009.
- [100] D. Norman. Emotion & design: attractive things work better. *Interactions*, 9(4):36–42, 2002.
- [101] Michel Obbink, Hayrettin Gürkök, Danny Plass-Oude Bos, Gido Hakvoort, Mannes Poel, and Anton Nijholt. Social interaction in a cooperative brain-computer interface game. In Antonio Camurri and Cristina Costa, editors, *Intelligent Technologies for Interactive Entertainment*, volume 78 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 183–192. Springer Berlin Heidelberg, 2012. ISBN 978-3-642-30213-8. doi: 10.1007/978-3-642-30214-5_20. URL http://dx.doi.org/10.1007/978-3-642-30214-5_20.
- [102] H. Omar and A. Jaafar. Heuristics evaluation in computer games. In Zainab Abu Bakar, Tengku Mohd Tengku Sembok, Halimah Badioze Zaman, Peter Bruza, Fabio Crestani, Shalini R. Urs, and Zaiki Awang, editors, *2010 International Conference on Information Retrieval Knowledge Management*, pages 188–193, Piscataway, NJ, USA, 2010. IEEE.
- [103] 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.
- [104] D. Oude Bos and B. Reuderink. BrainBasher: a BCI Game. In *Extended Abstracts of the International Conference on Fun and Games*. Eindhoven TU/e, 2008.
  - [105] Emanuele Pasqualotto, Alessandro Simonetta, Veronica Gnisci, Stefano Federici, and Marta Olivetti Belardinelli. Toward a usability evaluation of bcis. *International Journal of Bioelectromagnetism*, 13(3):121–122, 2011.
  - [106] G. Pfurtscheller. Functional brain imaging based on ERD/ERS. *Vision Research*, 41(10-11):1257 – 1260, 2001. ISSN 0042-6989. doi: DOI:10.1016/S0042-6989(00)00235-2. URL <http://www.sciencedirect.com/science/article/B6T0W-436FSH7-3/2/62b0af7cf15d3b79c50baac3446f8e80>.
  - [107] J.A. Pineda, D.S. Silverman, A. Vankov, and J. Hestenes. Learning to control brain rhythms: making a brain-computer interface possible. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 11(2):181–184, 2003.
  - [108] 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.
  - [109] D. Plass-Oude Bos, M. Duvinage, O. Oktay, J. Delgado Saa, H. Guruler, A. Istanbulu, M. van Vliet, B. L. A. van de Laar, M. Poel, L. Roijendijk, L. Tonin, A. Bahramisharif, and B. Reuderink. Looking around with your brain in a virtual world. In *Proceedings of the 6th International Summer Workshop on Multimodal Interfaces, eNTERFACE'10*, pages 12–23. University of Amsterdam, 2010.
  - [110] D. Plass-Oude Bos, B. Reuderink, B. L. A. van de Laar, H. Gürkök, C. Mühl, M. Poel, A. Nijholt, and D. K. J. Heylen. Brain-computer interfacing and games. In D. Tan and A. Nijholt, editors, *Brain-Computer Interfaces. Applying our Minds to Human-Computer Interaction*, Human-Computer Interaction Series, pages 149–178. Springer Verlag, London, July 2010.
  - [111] D. Plass-Oude Bos, H. Gürkök, B. L. A. van de Laar, F. Nijboer, and A. Nijholt. User Experience Evaluation in BCI: Mind the Gap! *International Journal of Bioelectromagnetism*, 13(1):48–49, July 2011. ISSN 1456-7857.

- [112] D. Plass-Oude Bos, M. Poel, and A. Nijholt. A study in user-centered design and evaluation of mental tasks for BCI. *Advances in Multimedia Modeling*, pages 122–134, 2011.
- [113] Danny Plass-Oude Bos, Boris Reuderink, Bram L. A. van de Laar, Hayrettin Gürkök, Christian Mühl, Mannes Poel, Anton Nijholt, and Dirk K. J. Heylen. *Brain-Computer Interfacing and Games*, pages 149–178. Springer, London, June 2010.
- [114] R. Poppe. A survey on vision-based human action recognition. *Image and Vision Computing*, 28(6):976–990, 2010.
- [115] M. Quek, D. Boland, J. Williamson, R. Murray-Smith, M. Tavella, S. Perdakis, M. Schreuder, and M. Tangermann. Simulating the feel of brain-computer interfaces for design, development and social interaction. In *Proceedings of CHI 2011*. ACM, 2011.
- [116] Boris Reuderink, Anton Nijholt, and Mannes Poel. Affective Pacman: A frustrating game for brain-computer interface experiments. volume 9, pages 221–227, May 2009.
- [117] F. Rheinberg, R. Vollmeyer, and B.D. Burns. FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica*, 47(2):57–66, 2001. ISSN 0012-1924.
- [118] Tomas Ros, Merrick Moseley, Philip Bloom, Larry Benjamin, Lesley Parkinson, and John Gruzelier. Optimizing microsurgical skills with eeg neurofeedback. *BMC Neuroscience*, 10(1):87, 2009. ISSN 1471-2202. doi: 10.1186/1471-2202-10-87.
- [119] Aldo Rustichini. Neuroeconomics: what have we found, and what should we search for. *Current Opinion in Neurobiology*, 19(6):672 – 677, 2009. ISSN 0959-4388. doi: 10.1016/j.conb.2009.09.012.
- [120] C. Sannelli, M. Braun, M. Tangermann, and K.-R. Müller. 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.
- [121] G. Schalk, D.J. McFarland, T. Hinterberger, N. Birbaumer, and J.R. Wolpaw. BCI2000: a general-purpose brain-computer interface (BCI) system. *Biomedical Engineering, IEEE Transactions on*, 51(6):1034 –1043, june 2004. ISSN 0018-9294. doi: 10.1109/TBME.2004.827072.

- [122] M.J. Scherer. *Living in the State of Stuck : How Assistive Technology Impacts the Lives of People with Disabilities*. Brookline Books, third edition, 2000.
- [123] H. Serby, E. Yom-Tov, and G.F. Inbar. An improved p300-based brain-computer interface. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 13(1):89–98, 2005.
- [124] Hiroshi Shibasaki and Mark Hallett. What is the bereitschaftspotential? *Clinical Neurophysiology*, 117:2341–2356, 2006.
- [125] M. Slater and A. Steed. A virtual presence counter. *Presence: Teleoperators & Virtual Environments*, 9(5):413–434, 2000. ISSN 1054-7460.
- [126] Erin Treacy Solovey, Audrey Girouard, Krysta Chauncey, Leanne M. Hirshfield, Angelo Sassaroli, Feng Zheng, Sergio Fantini, and Robert J.K. Jacob. Using fNIRS brain sensing in realistic HCI settings: Experiments and guidelines. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology*, pages 157–166, New York, NY, USA, 2009. ACM.
- [127] Kenyon Stamps and Yskandar Hamam. Towards inexpensive bci control for wheelchair navigation in the enabled environment - a hardware survey. In Yiyu Yao, Ron Sun, Tomaso Poggio, Jiming Liu, Ning Zhong, and Jimmy Huang, editors, *Brain Informatics*, volume 6334 of *Lecture Notes in Computer Science*, pages 336–345. Springer Berlin Heidelberg, 2010. ISBN 978-3-642-15313-6. doi: 10.1007/978-3-642-15314-3_32. URL http://dx.doi.org/10.1007/978-3-642-15314-3_32.
- [128] Kazuhiro Takeshita, Akihiko Uchibori, Yoshiki Mizukami, Taiji Satoh, Kanya Tanaka, and Shigeru Uchikado. A communication system for ALS patients using eye blink. *International Journal of Applied Electromagnetics and Mechanics*, 18(1-3):3–10, 2003.
- [129] D. Tan and N. Nijholt, editors. *Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction*. Springer Verlag, 2010.
- [130] S.C. Thompson, W. Armstrong, and C. Thomas. Illusions of control, underestimations, and accuracy: A control heuristic explanation. *Psychological Bulletin*, 123(2): 143, 1998.

- [131] D. A. Todd, P. J. McCullagh, M. D. Mulvenna, and G. Lightbody. Investigating the use of brain-computer interaction to facilitate creativity. In *Proceedings of the 3rd Augmented Human International Conference, AH '12*, pages 19:1–19:8, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1077-2.
- [132] N. Tractinsky, A.S. Katz, and D. Ikar. What is beautiful is usable. *Interacting with Computers*, 13(2):127–145, 2000.
- [133] Joy van Baren and Wijnand IJsselsteijn. Measuring Presence: A Guide to Current Measurement Approaches. In *OmniPres Project, Deliverable 5*, 2004. URL <http://www8.informatik.umu.se/~jworth/PresenceMeasurement.pdf>.
- [134] B. L. A. van de Laar, B. Reuderink, D. Plass-Oude Bos, and D. K. J. Heylen. Evaluating user experience of actual and imagined movement in bci gaming. *International Journal of Gaming and Computer Mediated Simulations*, 2(4): 33–47, October 2010. ISSN 1942-3888.
- [135] B. L. A. van de Laar, H. Gürkök, D. Plass-Oude Bos, F. Nijboer, and A. Nijholt. Perspectives on user experience evaluation of brain-computer interfaces. In C. Stephanidis, editor, *Universal Access in Human-Computer Interaction. Users Diversity. 6th International Conference, UAHCI 2011, Held as Part of HCI International 2011, Orlando*, volume 6766 of *Lecture Notes in Computer Science*, pages 600–609, Berlin, June 2011. Springer Verlag.
- [136] Bram van de Laar, Femke Nijboer, Hayrettin Gürkök, Danny Plass-Oude Bos, and Anton Nijholt. User Experience Evaluation in BCI: Bridge the Gap. *International Journal of Bioelectromagnetism*, 13(3):157–158, 2011.
- [137] M. van Vliet, A. Robben, N. Chumerin, N.V. Manyakov, A. Combaz, and M.M. Van Hulle. Designing a brain-computer interface controlled video-game using consumer grade eeg hardware. In *Biosignals and Biorobotics Conference (BRC), 2012 ISSNIP*, pages 1–6. IEEE, 2012.
- [138] J.J. Vidal. Real-time detection of brain events in EEG. *Proceedings of the IEEE*, 65(5):633–641, 1977.
- [139] Carmen Vidaurre and Benjamin Blankertz. Towards a cure for bci illiteracy. *Brain Topography*, 23:194–198, 2010. ISSN 0896-0267. URL <http://dx.doi.org/10.1007/s10548-009-0121-6>. 10.1007/s10548-009-0121-6.

- [140] R. Vilimek and T. Zander. Bc (eye): Combining eye-gaze input with brain-computer interaction. *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, pages 593–602, 2009.
- [141] M.P. Ware, P.J. McCullagh, A. McRoberts, G. Lightbody, C. Nugent, G. McAllister, M.D. Mulvenna, E. Thomson, and S. Martin. Contrasting levels of accuracy in command interaction sequences for a domestic brain-computer interface using SSVEP. In *Biomedical Engineering Conference (CIBEC), 2010 5th Cairo International*, pages 150–153. IEEE, 2010.
- [142] E.L. Waterworth and J.A. Waterworth. Focus, locus, and sensus: The three dimensions of virtual experience. *CyberPsychology & Behavior*, 4(2):203–213, 2001.
- [143] J. A. Waterworth and E. L. Waterworth. Being and time: judged presence and duration as a function of media form. *Presence: Teleoper. Virtual Environ.*, 12(5): 495–511, October 2003. ISSN 1054-7460. doi: 10.1162/105474603322761298.
- [144] J.A. Waterworth. Memory mechanisms and the psychophysical scaling of duration. *Perception*, 1985.
- [145] Bob G. Witmer and Michael J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoper. Virtual Environ.*, 7(3):225–240, June 1998. ISSN 1054-7460. doi: 10.1162/105474698565686.
- [146] Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, and Theresa M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113:767–791, 2002.
- [147] J.R. Wolpaw and E.W. Wolpaw. Brain-computer interfaces: Something new under the sun. *Brain-Computer Interfaces: Principles and Practice*, page 3, 2012.
- [148] T. O. Zander, C. Kothe, S. Welke, and M. Roetting. Utilizing secondary input from passive brain-computer interfaces for enhancing human-machine interaction. In *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, volume 5638 of *Lecture Notes in Computer Science*, pages 759–771, Germany, 2009. Springer-Verlag Berlin Heidelberg.
- [149] T.O. Zander and S. Jatzev. Detecting affective covert user states with passive brain-computer interfaces. In *Affective*

- Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–9. IEEE, 2009.
- [150] T.O. Zander, C. Kothe, S. Jatzev, and M. Gaertner. Enhancing human-computer interaction with input from active and passive brain-computer interfaces. *Brain-Computer Interfaces*, pages 181–199, 2010.
  - [151] B. Zhang, J. Wang, and T. Fuhlbrigge. A review of the commercial brain-computer interface technology from perspective of industrial robotics. In *Automation and Logistics (ICAL), 2010 IEEE International Conference on*, pages 379–384. IEEE, 2010.
  - [152] Dan Zhang, Yijun Wang, Xiaorong Gao, Bo Hong, and Shangkai Gao. An algorithm for idle-state detection in motor-imagery-based brain-computer interface. *Intell. Neuroscience*, 2007:5–5, January 2007. ISSN 1687-5265. doi: 10.1155/2007/39714. URL <http://dx.doi.org/10.1155/2007/39714>.
  - [153] C. Zickler, V. Di Donna, V. Kaiser, A. Al-Khodairy, S. Kleih, A. Küebler, M. Malavasi, D. Mattia, S. Mongardi, C. Neuper, M. Rohm, and R. Rupp. BCI applications for people with disabilities: Defining user needs and user requirements. In *Assistive Technology from Adapted Equipment to Inclusive Environments: AAATE2009*, pages 185–189, Amsterdam, the Netherlands, 2009. IOS Press.

SIKS Dissertatiereeks

====
1998
====

- 1998-1 Johan van den Akker (CWI)
DEGAS - An Active, Temporal Database of Autonomous Objects
- 1998-2 Floris Wiesman (UM)
Information Retrieval by Graphically Browsing Meta-Information
- 1998-3 Ans Steuten (TUD)
A Contribution to the Linguistic Analysis of Business Conversations
within the Language/Action Perspective
- 1998-4 Dennis Breuker (UM)
Memory versus Search in Games
- 1998-5 E.W.Oskamp (RUL)
Computerondersteuning bij Straftoemeting

====
1999
====

- 1999-1 Mark Sloof (VU)
Physiology of Quality Change Modelling;
Automated modelling of Quality Change of Agricultural Products
- 1999-2 Rob Potharst (EUR)
Classification using decision trees and neural nets
- 1999-3 Don Beal (UM)
The Nature of Minimax Search
- 1999-4 Jacques Penders (UM)
The practical Art of Moving Physical Objects
- 1999-5 Aldo de Moor (KUB)
Empowering Communities: A Method for the Legitimate User-Driven
Specification of Network Information Systems
- 1999-6 Niek J.E. Wijngaards (VU)
Re-design of compositional systems
- 1999-7 David Spelt (UT)
Verification support for object database design
- 1999-8 Jacques H.J. Lenting (UM)
Informed Gambling: Conception and Analysis of a Multi-Agent
Mechanism for Discrete Reallocation.

====
2000
====

- 2000-1 Frank Niessink (VU)
Perspectives on Improving Software Maintenance
- 2000-2 Koen Holtman (TUE)

Prototyping of CMS Storage Management

- 2000-3 Carolien M.T. Metselaar (UVA)
Sociaal-organisatorische gevolgen van kennistechnologie;
een procesbenadering en actorperspectief.
- 2000-4 Geert de Haan (VU)
ETAG, A Formal Model of Competence Knowledge for User Interface Design
- 2000-5 Ruud van der Pol (UM)
Knowledge-based Query Formulation in Information Retrieval.
- 2000-6 Rogier van Eijk (UU)
Programming Languages for Agent Communication
- 2000-7 Niels Peek (UU)
Decision-theoretic Planning of Clinical Patient Management
- 2000-8 Veerle Coup (EUR)
Sensitivity Analysis of Decision-Theoretic Networks
- 2000-9 Florian Waas (CWI)
Principles of Probabilistic Query Optimization
- 2000-10 Niels Nes (CWI)
Image Database Management System Design Considerations,
Algorithms and Architecture
- 2000-11 Jonas Karlsson (CWI)
Scalable Distributed Data Structures for Database Management

=====
2001
=====

- 2001-1 Silja Renooij (UU)
Qualitative Approaches to Quantifying Probabilistic Networks
- 2001-2 Koen Hindriks (UU)
Agent Programming Languages: Programming with Mental Models
- 2001-3 Maarten van Someren (UvA)
Learning as problem solving
- 2001-4 Evgueni Smirnov (UM)
Conjunctive and Disjunctive Version Spaces with
Instance-Based Boundary Sets
- 2001-5 Jacco van Ossenbruggen (VU)
Processing Structured Hypermedia: A Matter of Style
- 2001-6 Martijn van Welie (VU)
Task-based User Interface Design
- 2001-7 Bastiaan Schonhage (VU)
Diva: Architectural Perspectives on Information Visualization
- 2001-8 Pascal van Eck (VU)
A Compositional Semantic Structure for Multi-Agent Systems Dynamics.
- 2001-9 Pieter Jan 't Hoen (RUL)
Towards Distributed Development of Large Object-Oriented Models,
Views of Packages as Classes

2001-10 Maarten Sierhuis (UvA)
Modeling and Simulating Work Practice
BRAHMS: a multiagent modeling and simulation language
for work practice analysis and design

2001-11 Tom M. van Engers (VUA)
Knowledge Management:
The Role of Mental Models in Business Systems Design

====
2002
====

2002-01 Nico Lassing (VU)
Architecture-Level Modifiability Analysis

2002-02 Roelof van Zwol (UT)
Modelling and searching web-based document collections

2002-03 Henk Ernst Blok (UT)
Database Optimization Aspects for Information Retrieval

2002-04 Juan Roberto Castelo Valdueza (UU)
The Discrete Acyclic Digraph Markov Model in Data Mining

2002-05 Radu Serban (VU)
The Private Cyberspace Modeling Electronic Environments
inhabited by Privacy-concerned Agents

2002-06 Laurens Mommers (UL)
Applied legal epistemology;
Building a knowledge-based ontology of the legal domain

2002-07 Peter Boncz (CWI)
Monet: A Next-Generation DBMS Kernel For Query-Intensive Applications

2002-08 Jaap Gordijn (VU)
Value Based Requirements Engineering: Exploring Innovative
E-Commerce Ideas

2002-09 Willem-Jan van den Heuvel(KUB)
Integrating Modern Business Applications with Objectified Legacy Systems

2002-10 Brian Sheppard (UM)
Towards Perfect Play of Scrabble

2002-11 Wouter C.A. Wijngaards (VU)
Agent Based Modelling of Dynamics: Biological and Organisational Applications

2002-12 Albrecht Schmidt (Uva)
Processing XML in Database Systems

2002-13 Hongjing Wu (TUE)
A Reference Architecture for Adaptive Hypermedia Applications

2002-14 Wieke de Vries (UU)
Agent Interaction: Abstract Approaches to Modelling, Programming and
Verifying Multi-Agent Systems

2002-15 Rik Eshuis (UT)
Semantics and Verification of UML Activity Diagrams for Workflow Modelling

2002-16 Pieter van Langen (VU)
The Anatomy of Design: Foundations, Models and Applications

2002-17 Stefan Manegold (UVA)
Understanding, Modeling, and Improving Main-Memory Database Performance

====

2003

====

2003-01 Heiner Stuckenschmidt (VU)
Ontology-Based Information Sharing in Weakly Structured Environments

2003-02 Jan Broersen (VU)
Modal Action Logics for Reasoning About Reactive Systems

2003-03 Martijn Schuemie (TUD)
Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy

2003-04 Milan Petkovic (UT)
Content-Based Video Retrieval Supported by Database Technology

2003-05 Jos Lehmann (UVA)
Causation in Artificial Intelligence and Law - A modelling approach

2003-06 Boris van Schooten (UT)
Development and specification of virtual environments

2003-07 Machiel Jansen (UvA)
Formal Explorations of Knowledge Intensive Tasks

2003-08 Yongping Ran (UM)
Repair Based Scheduling

2003-09 Rens Kortmann (UM)
The resolution of visually guided behaviour

2003-10 Andreas Lincke (UvT)
Electronic Business Negotiation: Some experimental studies on the interaction
between medium, innovation context and culture

2003-11 Simon Keizer (UT)
Reasoning under Uncertainty in Natural Language Dialogue using Bayesian Networks

2003-12 Roeland Ordelman (UT)
Dutch speech recognition in multimedia information retrieval

2003-13 Jeroen Donkers (UM)
Nosce Hostem - Searching with Opponent Models

2003-14 Stijn Hoppenbrouwers (KUN)
Freezing Language: Conceptualisation Processes across ICT-Supported Organisations

2003-15 Mathijs de Weerd (TUD)
Plan Merging in Multi-Agent Systems

2003-16 Menzo Windhouwer (CWI)
Feature Grammar Systems - Incremental Maintenance of Indexes to
Digital Media Warehouses

2003-17 David Jansen (UT)
Extensions of Statecharts with Probability, Time, and Stochastic Timing

2003-18 Levente Kocsis (UM)
Learning Search Decisions

====
2004
====

- 2004-01 Virginia Dignum (UU)
A Model for Organizational Interaction: Based on Agents, Founded in Logic
- 2004-02 Lai Xu (UvT)
Monitoring Multi-party Contracts for E-business
- 2004-03 Perry Groot (VU)
A Theoretical and Empirical Analysis of Approximation in Symbolic Problem Solving
- 2004-04 Chris van Aart (UVA)
Organizational Principles for Multi-Agent Architectures
- 2004-05 Viara Popova (EUR)
Knowledge discovery and monotonicity
- 2004-06 Bart-Jan Hommes (TUD)
The Evaluation of Business Process Modeling Techniques
- 2004-07 Elise Boltjes (UM)
Voorbeeldig onderwijs; voorbeeldgestuurd onderwijs, een opstap naar abstract denken, vooral voor meisjes
- 2004-08 Joop Verbeek(UM)
Politie en de Nieuwe Internationale Informatiemarkt, Grensregionale politieke gegevensuitwisseling en digitale expertise
- 2004-09 Martin Caminada (VU)
For the Sake of the Argument; explorations into argument-based reasoning
- 2004-10 Suzanne Kabel (UVA)
Knowledge-rich indexing of learning-objects
- 2004-11 Michel Klein (VU)
Change Management for Distributed Ontologies
- 2004-12 The Duy Bui (UT)
Creating emotions and facial expressions for embodied agents
- 2004-13 Wojciech Jamroga (UT)
Using Multiple Models of Reality: On Agents who Know how to Play
- 2004-14 Paul Harrenstein (UU)
Logic in Conflict. Logical Explorations in Strategic Equilibrium
- 2004-15 Arno Knobbe (UU)
Multi-Relational Data Mining
- 2004-16 Federico Divina (VU)
Hybrid Genetic Relational Search for Inductive Learning
- 2004-17 Mark Winands (UM)
Informed Search in Complex Games
- 2004-18 Vania Bessa Machado (UvA)
Supporting the Construction of Qualitative Knowledge Models
- 2004-19 Thijs Westerveld (UT)
Using generative probabilistic models for multimedia retrieval
- 2004-20 Madelon Evers (Nyenrode)

Learning from Design: facilitating multidisciplinary design teams

=====

2005

=====

- 2005-01 Floor Verdenius (UVA)
Methodological Aspects of Designing Induction-Based Applications
- 2005-02 Erik van der Werf (UM))
AI techniques for the game of Go
- 2005-03 Franc Grootjen (RUN)
A Pragmatic Approach to the Conceptualisation of Language
- 2005-04 Nirvana Meratnia (UT)
Towards Database Support for Moving Object data
- 2005-05 Gabriel Infante-Lopez (UVA)
Two-Level Probabilistic Grammars for Natural Language Parsing
- 2005-06 Pieter Spronck (UM)
Adaptive Game AI
- 2005-07 Flavius Frasincar (TUE)
Hypermedia Presentation Generation for Semantic Web Information Systems
- 2005-08 Richard Vdovjak (TUE)
A Model-driven Approach for Building Distributed Ontology-based Web Applications
- 2005-09 Jeen Broekstra (VU)
Storage, Querying and Inferencing for Semantic Web Languages
- 2005-10 Anders Bouwer (UVA)
Explaining Behaviour: Using Qualitative Simulation in Interactive Learning Environments
- 2005-11 Elth Ogston (VU)
Agent Based Matchmaking and Clustering - A Decentralized Approach to Search
- 2005-12 Csaba Boer (EUR)
Distributed Simulation in Industry
- 2005-13 Fred Hamburg (UL)
Een Computermodel voor het Ondersteunen van Euthanasiebeslissingen
- 2005-14 Borys Omelayenko (VU)
Web-Service configuration on the Semantic Web; Exploring how semantics meets pragmatics
- 2005-15 Tibor Bosse (VU)
Analysis of the Dynamics of Cognitive Processes
- 2005-16 Joris Graaumans (UU)
Usability of XML Query Languages
- 2005-17 Boris Shishkov (TUD)
Software Specification Based on Re-usable Business Components
- 2005-18 Danielle Sent (UU)
Test-selection strategies for probabilistic networks
- 2005-19 Michel van Dartel (UM)
Situated Representation

2005-20 Cristina Coteanu (UL)
Cyber Consumer Law, State of the Art and Perspectives

2005-21 Wijnand Derks (UT)
Improving Concurrency and Recovery in Database Systems by
Exploiting Application Semantics

====
2006
====

2006-01 Samuil Angelov (TUE)
Foundations of B2B Electronic Contracting

2006-02 Cristina Chisalita (VU)
Contextual issues in the design and use of information technology in organizations

2006-03 Noor Christoph (UVA)
The role of metacognitive skills in learning to solve problems

2006-04 Marta Sabou (VU)
Building Web Service Ontologies

2006-05 Cees Pierik (UU)
Validation Techniques for Object-Oriented Proof Outlines

2006-06 Ziv Baida (VU)
Software-aided Service Bundling - Intelligent Methods & Tools
for Graphical Service Modeling

2006-07 Marko Smiljanic (UT)
XML schema matching – balancing efficiency and effectiveness by means of clustering

2006-08 Eelco Herder (UT)
Forward, Back and Home Again - Analyzing User Behavior on the Web

2006-09 Mohamed Wahdan (UM)
Automatic Formulation of the Auditor's Opinion

2006-10 Ronny Siebes (VU)
Semantic Routing in Peer-to-Peer Systems

2006-11 Joeri van Ruth (UT)
Flattening Queries over Nested Data Types

2006-12 Bert Bongers (VU)
Interactivation - Towards an e-cology of people, our technological environment, and the arts

2006-13 Henk-Jan Lebbink (UU)
Dialogue and Decision Games for Information Exchanging Agents

2006-14 Johan Hoorn (VU)
Software Requirements: Update, Upgrade, Redesign - towards a Theory of Requirements Change

2006-15 Rainer Malik (UU)
CONAN: Text Mining in the Biomedical Domain

2006-16 Carsten Riggelsen (UU)
Approximation Methods for Efficient Learning of Bayesian Networks

2006-17 Stacey Nagata (UU)
User Assistance for Multitasking with Interruptions on a Mobile Device

2006-18 Valentin Zhizhkun (UVA)

Graph transformation for Natural Language Processing

- 2006-19 Birna van Riemsdijk (UU)
Cognitive Agent Programming: A Semantic Approach
- 2006-20 Marina Velikova (UvT)
Monotone models for prediction in data mining
- 2006-21 Bas van Gils (RUN)
Aptness on the Web
- 2006-22 Paul de Vrieze (RUN)
Fundaments of Adaptive Personalisation
- 2006-23 Ion Juvina (UU)
Development of Cognitive Model for Navigating on the Web
- 2006-24 Laura Hollink (VU)
Semantic Annotation for Retrieval of Visual Resources
- 2006-25 Madalina Drugan (UU)
Conditional log-likelihood MDL and Evolutionary MCMC
- 2006-26 Vojkan Mihajlović (UT)
Score Region Algebra: A Flexible Framework for Structured Information Retrieval
- 2006-27 Stefano Bocconi (CWI)
Vox Populi: generating video documentaries from semantically annotated media repositories
- 2006-28 Borkur Sigurbjornsson (UVA)
Focused Information Access using XML Element Retrieval
- ====
2007
====
- 2007-01 Kees Leune (UvT)
Access Control and Service-Oriented Architectures
- 2007-02 Wouter Teepe (RUG)
Reconciling Information Exchange and Confidentiality: A Formal Approach
- 2007-03 Peter Mika (VU)
Social Networks and the Semantic Web
- 2007-04 Jurriaan van Diggelen (UU)
Achieving Semantic Interoperability in Multi-agent Systems: a dialogue-based approach
- 2007-05 Bart Schermer (UL)
Software Agents, Surveillance, and the Right to Privacy: a Legislative Framework for Agent-enabled Surveillance
- 2007-06 Gilad Mishne (UVA)
Applied Text Analytics for Blogs
- 2007-07 Natasa Jovanovic' (UT)
To Whom It May Concern - Addressee Identification in Face-to-Face Meetings
- 2007-08 Mark Hoogendoorn (VU)
Modeling of Change in Multi-Agent Organizations
- 2007-09 David Mobach (VU)
Agent-Based Mediated Service Negotiation
- 2007-10 Huib Aldewereld (UU)

Autonomy vs. Conformity: an Institutional Perspective on Norms and Protocols

- 2007-11 Natalia Stash (TUE)
Incorporating Cognitive/Learning Styles in a General-Purpose Adaptive Hypermedia System
- 2007-12 Marcel van Gerven (RUN)
Bayesian Networks for Clinical Decision Support: A Rational Approach to Dynamic Decision-Making under Uncertainty
- 2007-13 Rutger Rienks (UT)
Meetings in Smart Environments; Implications of Progressing Technology
- 2007-14 Niek Bergboer (UM)
Context-Based Image Analysis
- 2007-15 Joyca Lacroix (UM)
NIM: a Situated Computational Memory Model
- 2007-16 Davide Grossi (UU)
Designing Invisible Handcuffs. Formal investigations in Institutions and Organizations for Multi-agent Systems
- 2007-17 Theodore Charitos (UU)
Reasoning with Dynamic Networks in Practice
- 2007-18 Bart Orriens (UvT)
On the development an management of adaptive business collaborations
- 2007-19 David Levy (UM)
Intimate relationships with artificial partners
- 2007-20 Slinger Jansen (UU)
Customer Configuration Updating in a Software Supply Network
- 2007-21 Karianne Vermaas (UU)
Fast diffusion and broadening use: A research on residential adoption and usage of broadband internet in the Netherlands between 2001 and 2005
- 2007-22 Zlatko Zlatev (UT)
Goal-oriented design of value and process models from patterns
- 2007-23 Peter Barna (TUE)
Specification of Application Logic in Web Information Systems
- 2007-24 Georgina Ramírez Camps (CWI)
Structural Features in XML Retrieval
- 2007-25 Joost Schalken (VU)
Empirical Investigations in Software Process Improvement

====
2008
====

- 2008-01 Katalin Boer-Sorbán (EUR)
Agent-Based Simulation of Financial Markets: A modular,continuous-time approach
- 2008-02 Alexei Sharpanskykh (VU)
On Computer-Aided Methods for Modeling and Analysis of Organizations
- 2008-03 Vera Hollink (UVA)
Optimizing hierarchical menus: a usage-based approach
- 2008-04 Ander de Keijzer (UT)
Management of Uncertain Data - towards unattended integration

- 2008-05 Bela Mutschler (UT)
Modeling and simulating causal dependencies on process-aware information systems from a cost perspective
- 2008-06 Arjen Hommersom (RUN)
On the Application of Formal Methods to Clinical Guidelines, an Artificial Intelligence Perspective
- 2008-07 Peter van Rosmalen (OU)
Supporting the tutor in the design and support of adaptive e-learning
- 2008-08 Janneke Bolt (UU)
Bayesian Networks: Aspects of Approximate Inference
- 2008-09 Christof van Nimwegen (UU)
The paradox of the guided user: assistance can be counter-effective
- 2008-10 Wauter Bosma (UT)
Discourse oriented summarization
- 2008-11 Vera Kartseva (VU)
Designing Controls for Network Organizations: A Value-Based Approach
- 2008-12 Jozsef Farkas (RUN)
A Semiotically Oriented Cognitive Model of Knowledge Representation
- 2008-13 Caterina Carraciolo (UVA)
Topic Driven Access to Scientific Handbooks
- 2008-14 Arthur van Bunningen (UT)
Context-Aware Querying; Better Answers with Less Effort
- 2008-15 Martijn van Otterlo (UT)
The Logic of Adaptive Behavior: Knowledge Representation and Algorithms for the Markov Decision Process Framework in First-Order Domains.
- 2008-16 Henriette van Vugt (VU)
Embodied agents from a user's perspective
- 2008-17 Martin Op 't Land (TUD)
Applying Architecture and Ontology to the Splitting and Allying of Enterprises
- 2008-18 Guido de Croon (UM)
Adaptive Active Vision
- 2008-19 Henning Rode (UT)
From Document to Entity Retrieval: Improving Precision and Performance of Focused Text Search
- 2008-20 Rex Arendsen (UVA)
Geen bericht, goed bericht. Een onderzoek naar de effecten van de introductie van elektronisch berichtenverkeer met de overheid op de administratieve lasten van bedrijven
- 2008-21 Krisztian Balog (UVA)
People Search in the Enterprise
- 2008-22 Henk Koning (UU)
Communication of IT-Architecture
- 2008-23 Stefan Visscher (UU)
Bayesian network models for the management of ventilator-associated pneumonia
- 2008-24 Zharko Aleksovski (VU)
Using background knowledge in ontology matching
- 2008-25 Geert Jonker (UU)
Efficient and Equitable Exchange in Air Traffic Management Plan Repair using Spender-signed Currency
- 2008-26 Marijn Huijbregts (UT)

Segmentation, Diarization and Speech Transcription: Surprise Data Unraveled

- 2008-27 Hubert Vogten (OU)
Design and Implementation Strategies for IMS Learning Design
- 2008-28 Ildiko Flesch (RUN)
On the Use of Independence Relations in Bayesian Networks
- 2008-29 Dennis Reidsma (UT)
Annotations and Subjective Machines - Of Annotators, Embodied Agents, Users, and Other Humans
- 2008-30 Wouter van Atteveldt (VU)
Semantic Network Analysis: Techniques for Extracting, Representing and Querying Media Content
- 2008-31 Loes Braun (UM)
Pro-Active Medical Information Retrieval
- 2008-32 Trung H. Bui (UT)
Toward Affective Dialogue Management using Partially Observable Markov Decision Processes
- 2008-33 Frank Terpstra (UVA)
Scientific Workflow Design; theoretical and practical issues
- 2008-34 Jeroen de Knijf (UU)
Studies in Frequent Tree Mining
- 2008-35 Ben Torben Nielsen (UvT)
Dendritic morphologies: function shapes structure

====
2009
====

- 2009-01 Rasa Jurgelenaite (RUN)
Symmetric Causal Independence Models
- 2009-02 Willem Robert van Hage (VU)
Evaluating Ontology-Alignment Techniques
- 2009-03 Hans Stol (UvT)
A Framework for Evidence-based Policy Making Using IT
- 2009-04 Josephine Nabukenya (RUN)
Improving the Quality of Organisational Policy Making using Collaboration Engineering
- 2009-05 Sietse Overbeek (RUN)
Bridging Supply and Demand for Knowledge Intensive Tasks - Based on Knowledge, Cognition, and Quality
- 2009-06 Muhammad Subianto (UU)
Understanding Classification
- 2009-07 Ronald Poppe (UT)
Discriminative Vision-Based Recovery and Recognition of Human Motion
- 2009-08 Volker Nannen (VU)
Evolutionary Agent-Based Policy Analysis in Dynamic Environments
- 2009-09 Benjamin Kanagwa (RUN)
Design, Discovery and Construction of Service-oriented Systems
- 2009-10 Jan Wielemaker (UVA)
Logic programming for knowledge-intensive interactive applications

- 2009-11 Alexander Boer (UVA)
Legal Theory, Sources of Law & the Semantic Web
- 2009-12 Peter Massuthe (TUE, Humboldt-Universitaet zu Berlin)
Operating Guidelines for Services
- 2009-13 Steven de Jong (UM)
Fairness in Multi-Agent Systems
- 2009-14 Maksym Korotkiy (VU)
From ontology-enabled services to service-enabled ontologies (making ontologies work in e-science with ONTO-SOA)
- 2009-15 Rinke Hoekstra (UVA)
Ontology Representation - Design Patterns and Ontologies that Make Sense
- 2009-16 Fritz Reul (UvT)
New Architectures in Computer Chess
- 2009-17 Laurens van der Maaten (UvT)
Feature Extraction from Visual Data
- 2009-18 Fabian Groffen (CWI)
Armada, An Evolving Database System
- 2009-19 Valentin Robu (CWI)
Modeling Preferences, Strategic Reasoning and Collaboration in Agent-Mediated Electronic Markets
- 2009-20 Bob van der Vecht (UU)
Adjustable Autonomy: Controlling Influences on Decision Making
- 2009-21 Stijn Vanderlooy (UM)
Ranking and Reliable Classification
- 2009-22 Pavel Serdyukov (UT)
Search For Expertise: Going beyond direct evidence
- 2009-23 Peter Hofgesang (VU)
Modelling Web Usage in a Changing Environment
- 2009-24 Annerieke Heuvelink (VUA)
Cognitive Models for Training Simulations
- 2009-25 Alex van Ballegooij (CWI)
"RAM: Array Database Management through Relational Mapping"
- 2009-26 Fernando Koch (UU)
An Agent-Based Model for the Development of Intelligent Mobile Services
- 2009-27 Christian Glahn (OU)
Contextual Support of social Engagement and Reflection on the Web
- 2009-28 Sander Evers (UT)
Sensor Data Management with Probabilistic Models
- 2009-29 Stanislav Pokraev (UT)
Model-Driven Semantic Integration of Service-Oriented Applications
- 2009-30 Marcin Zukowski (CWI)
Balancing vectorized query execution with bandwidth-optimized storage
- 2009-31 Sofiya Katrenko (UVA)
A Closer Look at Learning Relations from Text
- 2009-32 Rik Farenhorst (VU) and Remco de Boer (VU)
Architectural Knowledge Management: Supporting Architects and Auditors

- 2009-33 Khiet Truong (UT)
How Does Real Affect Affect Affect Recognition In Speech?
- 2009-34 Inge van de Weerd (UU)
Advancing in Software Product Management: An Incremental Method Engineering Approach
- 2009-35 Wouter Koelewijn (UL)
Privacy en Politiegegevens; Over geautomatiseerde normatieve informatie-uitwisseling
- 2009-36 Marco Kalz (OUN)
Placement Support for Learners in Learning Networks
- 2009-37 Hendrik Drachsler (OUN)
Navigation Support for Learners in Informal Learning Networks
- 2009-38 Riina Vuorikari (OU)
Tags and self-organisation: a metadata ecology for learning resources in a multilingual context
- 2009-39 Christian Stahl (TUE, Humboldt-Universitaet zu Berlin)
Service Substitution – A Behavioral Approach Based on Petri Nets
- 2009-40 Stephan Raaijmakers (UvT)
Multinomial Language Learning: Investigations into the Geometry of Language
- 2009-41 Igor Berezhnyy (UvT)
Digital Analysis of Paintings
- 2009-42 Toine Bogers
Recommender Systems for Social Bookmarking
- 2009-43 Virginia Nunes Leal Franqueira (UT)
Finding Multi-step Attacks in Computer Networks using Heuristic Search and Mobile Ambients
- 2009-44 Roberto Santana Tapia (UT)
Assessing Business-IT Alignment in Networked Organizations
- 2009-45 Jilles Vreeken (UU)
Making Pattern Mining Useful
- 2009-46 Loredana Afanasiev (UvA)
Querying XML: Benchmarks and Recursion
- ====
2010
====
- 2010-01 Matthijs van Leeuwen (UU)
Patterns that Matter
- 2010-02 Ingo Wassink (UT)
Work flows in Life Science
- 2010-03 Joost Geurts (CWI)
A Document Engineering Model and Processing Framework for Multimedia documents
- 2010-04 Olga Kulyk (UT)
Do You Know What I Know? Situational Awareness of Co-located Teams in Multidisplay Environments
- 2010-05 Claudia Hauff (UT)
Predicting the Effectiveness of Queries and Retrieval Systems
- 2010-06 Sander Bakkes (UvT)
Rapid Adaptation of Video Game AI

2010-07 Wim Fikkert (UT)
Gesture interaction at a Distance

2010-08 Krzysztof Siewicz (UL)
Towards an Improved Regulatory Framework of Free Software. Protecting user freedoms in a world of software communities and eGovernments

2010-09 Hugo Kielman (UL)
A Politiele gegevensverwerking en Privacy, Naar een effectieve waarborging

2010-10 Rebecca Ong (UL)
Mobile Communication and Protection of Children

2010-11 Adriaan Ter Mors (TUD)
The world according to MARP: Multi-Agent Route Planning

2010-12 Susan van den Braak (UU)
Sensemaking software for crime analysis

2010-13 Gianluigi Folino (RUN)
High Performance Data Mining using Bio-inspired techniques

2010-14 Sander van Splunter (VU)
Automated Web Service Reconfiguration

2010-15 Lianne Bodestaff (UT)
Managing Dependency Relations in Inter-Organizational Models

2010-16 Sicco Verwer (TUD)
Efficient Identification of Timed Automata, theory and practice

2010-17 Spyros Kotoulas (VU)
Scalable Discovery of Networked Resources: Algorithms, Infrastructure, Applications

2010-18 Charlotte Gerritsen (VU)
Caught in the Act: Investigating Crime by Agent-Based Simulation

2010-19 Henriette Cramer (UvA)
People's Responses to Autonomous and Adaptive Systems

2010-20 Ivo Swartjes (UT)
Whose Story Is It Anyway? How Improv Informs Agency and Authorship of Emergent Narrative

2010-21 Harold van Heerde (UT)
Privacy-aware data management by means of data degradation

2010-22 Michiel Hildebrand (CWI)
End-user Support for Access to
Heterogeneous Linked Data

2010-23 Bas Steunebrink (UU)
The Logical Structure of Emotions

2010-24 Dmytro Tykhonov
Designing Generic and Efficient Negotiation Strategies

2010-25 Zulfiqar Ali Memon (VU)
Modelling Human-Awareness for Ambient Agents: A Human Mindreading Perspective

2010-26 Ying Zhang (CWI)
XRPC: Efficient Distributed Query Processing on Heterogeneous XQuery Engines

2010-27 Marten Voulon (UL)
Automatisch contracteren

2010-28 Arne Koopman (UU)
Characteristic Relational Patterns

2010-29 Stratos Idreos (CWI)
Database Cracking: Towards Auto-tuning Database Kernels

- 2010-30 Marieke van Erp (UvT)
Accessing Natural History - Discoveries in data cleaning, structuring, and retrieval
- 2010-31 Victor de Boer (UVA)
Ontology Enrichment from Heterogeneous Sources on the Web
- 2010-32 Marcel Hiel (UvT)
An Adaptive Service Oriented Architecture: Automatically solving Interoperability Problems
- 2010-33 Robin Aly (UT)
Modeling Representation Uncertainty in Concept-Based Multimedia Retrieval
- 2010-34 Teduh Dirgahayu (UT)
Interaction Design in Service Compositions
- 2010-35 Dolf Trieschnigg (UT)
Proof of Concept: Concept-based Biomedical Information Retrieval
- 2010-36 Jose Janssen (OU)
Paving the Way for Lifelong Learning; Facilitating competence development through a learning path specification
- 2010-37 Niels Lohmann (TUE)
Correctness of services and their composition
- 2010-38 Dirk Fahland (TUE)
From Scenarios to components
- 2010-39 Ghazanfar Farooq Siddiqui (VU)
Integrative modeling of emotions in virtual agents
- 2010-40 Mark van Assem (VU)
Converting and Integrating Vocabularies for the Semantic Web
- 2010-41 Guillaume Chaslot (UM)
Monte-Carlo Tree Search
- 2010-42 Sybren de Kinderen (VU)
Needs-driven service bundling in a multi-supplier setting - the computational e3-service approach
- 2010-43 Peter van Kranenburg (UU)
A Computational Approach to Content-Based Retrieval of Folk Song Melodies
- 2010-44 Pieter Bellekens (TUE)
An Approach towards Context-sensitive and User-adapted Access to Heterogeneous Data Sources, Illustrated in the Television Domain
- 2010-45 Vasilios Andrikopoulos (UvT)
A theory and model for the evolution of software services
- 2010-46 Vincent Pijpers (VU)
e3alignment: Exploring Inter-Organizational Business-ICT Alignment
- 2010-47 Chen Li (UT)
Mining Process Model Variants: Challenges, Techniques, Examples
- 2010-48 Milan Lovric (EUR)
Behavioral Finance and Agent-Based Artificial Markets
- 2010-49 Jahn-Takeshi Saito (UM)
Solving difficult game positions
- 2010-50 Bouke Huurnink (UVA)
Search in Audiovisual Broadcast Archives

2010-51 Alia Khairia Amin (CWI)
Understanding and supporting information seeking tasks in multiple sources

2010-52 Peter-Paul van Maanen (VU)
Adaptive Support for Human-Computer Teams: Exploring the Use of Cognitive Models of Trust and Attention

2010-53 Edgar Meij (UVA)
Combining Concepts and Language Models for Information Access

====
2011
====

2011-01 Botond Cseke (RUN)
Variational Algorithms for Bayesian Inference in Latent Gaussian Models

2011-02 Nick Tinnemeier(UU)
Organizing Agent Organizations. Syntax and Operational Semantics of an Organization-Oriented Programming Language

2011-03
Jan Martijn van der Werf (TUE)
Compositional Design and Verification of Component-Based Information Systems

2011-04 Hado van Hasselt (UU)
Insights in Reinforcement Learning; Formal analysis and empirical evaluation of temporal-difference learning algorithms

2011-05 Base van der Raadt (VU)
Enterprise Architecture Coming of Age - Increasing the Performance of an Emerging Discipline.

2011-06 Yiwen Wang (TUE)
Semantically-Enhanced Recommendations in Cultural Heritage

2011-07 Yujia Cao (UT)
Multimodal Information Presentation for High Load Human Computer Interaction

2011-08 Nieske Vergunst (UU)
BDI-based Generation of Robust Task-Oriented Dialogues 2011-09 Tim de Jong (OU)
Contextualised Mobile Media for Learning

2011-10 Bart Bogaert (UvT)
Cloud Content Contention

2011-11 Dhaval Vyas (UT)
Designing for Awareness: An Experience-focused HCI Perspective

2011-12 Carmen Bratosin (TUE)
Grid Architecture for Distributed Process Mining

2011-13 Xiaoyu Mao (UvT)
Airport under Control. Multiagent Scheduling for Airport Ground Handling

2011-14 Milan Lovric (EUR)
Behavioral Finance and Agent-Based Artificial Markets

2011-15 Marijn Koolen (UvA)
The Meaning of Structure: the Value of Link Evidence for Information Retrieval

2011-16 Maarten Schadd (UM)
Selective Search in Games of Different Complexity

2011-17 Jiyin He (UVA)

Exploring Topic Structure: Coherence, Diversity and Relatedness

- 2011-18 Mark Ponsen (UM)
Strategic Decision-Making in complex games
- 2011-19 Ellen Rusman (OU)
The Mind 's Eye on Personal Profiles
- 2011-20 Qing Gu (VU)
Guiding service-oriented software engineering - A view-based approach
- 2011-21 Linda Terlouw (TUD)
Modularization and Specification of Service-Oriented Systems
- 2011-22 Junte Zhang (UVA)
System Evaluation of Archival Description and Access
- 2011-23 Wouter Weerkamp (UVA)
Finding People and their Utterances in Social Media
- 2011-24 Herwin van Welbergen (UT)
Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior
- 2011-25 Syed Waqar ul Qounain Jaffry (VU)
Analysis and Validation of Models for Trust Dynamics
- 2011-26 Matthijs Aart Pontier (VU)
Virtual Agents for Human Communication - Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots
- 2011-27 Aniel Bhulai (VU)
Dynamic website optimization through autonomous management of design patterns
- 2011-28 Rianne Kaptein(UVA)
Effective Focused Retrieval by Exploiting Query Context and Document Structure
- 2011-29 Faisal Kamiran (TUE)
Discrimination-aware Classification
- 2011-30 Egon van den Broek (UT)
Affective Signal Processing (ASP): Unraveling the mystery of emotions
- 2011-31 Ludo Waltman (EUR)
Computational and Game-Theoretic Approaches for Modeling Bounded Rationality
- 2011-32 Nees-Jan van Eck (EUR)
Methodological Advances in Bibliometric Mapping of Science
- 2011-33 Tom van der Weide (UU)
Arguing to Motivate Decisions
- 2011-34 Paolo Turrini (UU)
Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations
- 2011-35 Maaïke Harbers (UU)
Explaining Agent Behavior in Virtual Training
- 2011-36 Erik van der Spek (UU)
Experiments in serious game design: a cognitive approach
- 2011-37 Adriana Burlutiu (RUN)
Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference
- 2011-38 Nyree Lemmens (UM)
Bee-inspired Distributed Optimization

2011-39 Joost Westra (UU)
Organizing Adaptation using Agents in Serious Games

2011-40 Viktor Clerc (VU)
Architectural Knowledge Management in Global Software Development

2011-41 Luan Ibraimi (UT)
Cryptographically Enforced Distributed Data Access Control

2011-42 Michal Sindlar (UU)
Explaining Behavior through Mental State Attribution

2011-43 Henk van der Schuur (UU)
Process Improvement through Software Operation Knowledge

2011-44 Boris Reuderink (UT)
Robust Brain-Computer Interfaces

2011-45 Herman Stehouwer (UvT)
Statistical Language Models for Alternative Sequence Selection

2011-46 Beibei Hu (TUD)
Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work

2011-47 Azizi Bin Ab Aziz(VU)
Exploring Computational Models for Intelligent Support of Persons with Depression

2011-48 Mark Ter Maat (UT)
Response Selection and Turn-taking for a Sensitive Artificial Listening Agent

2011-49 Andreea Niculescu (UT)
Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality

====
2011
====

2012-01 Terry Kakeeto (UvT)
Relationship Marketing for SMEs in Uganda

2012-02 Muhammad Umair(VU)
Adaptivity, emotion, and Rationality in Human and Ambient Agent Models

2012-03 Adam Vanya (VU)
Supporting Architecture Evolution by Mining Software Repositories

2012-04 Jurriaan Souer (UU)
Development of Content Management System-based Web Applications

2012-05 Marijn Plomp (UU)
Maturing Interorganisational Information Systems

2012-06 Wolfgang Reinhardt (OU)
Awareness Support for Knowledge Workers in Research Networks

2012-07 Rianne van Lambalgen (VU)
When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions

2012-08 Gerben de Vries (UVA)
Kernel Methods for Vessel Trajectories

2012-09 Ricardo Nisse (UT)
Trust and Privacy Management Support for Context-Aware Service Platforms

- 2012-10 David Smits (TUE)
Towards a Generic Distributed Adaptive Hypermedia Environment
- 2012-11 J.C.B. Rantham Prabhakara (TUE) Process Mining in the Large: Preprocessing, Discovery, and Diagnostics
- 2012-12 Kees van der Sluijs (TUE)
Model Driven Design and Data Integration in Semantic Web Information Systems
- 2012-13 Suleman Shahid (UvT)
Fun and Face: Exploring non-verbal expressions of emotion during playful interactions
- 2012-14
Evgeny Knutov(TUE)
Generic Adaptation Framework for Unifying Adaptive Web-based Systems
- 2012-15 Natalie van der Wal (VU)
Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes
- 2012-16 Fiemke Both (VU)
Helping people by understanding them - Ambient Agents supporting task execution and depression treatment
- 2012-17 Amal Elgammal (UvT)
Towards a Comprehensive Framework for Business Process Compliance
- 2012-18 Eltjo Poort (VU)
Improving Solution Architecting Practices
- 2012-19 Helen Schonenberg (TUE)
What's Next? Operational Support for Business Process Execution
- 2012-20 Ali Bahramisharif (RUN)
Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing
- 2012-21 Roberto Cornacchia (TUD)
Querying Sparse Matrices for Information Retrieval
- 2012-22 Thijs Vis (UvT)
Intelligence, politie en veiligheidsdienst: verenigbare grootheden?
- 2012-23 Christian Muehl (UT)
Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction