Benjamin Blankertz · Giulio Jacucci
Luciano Gamberini · Anna Spagnolli
Jonathan Freeman (Eds.)

LNCS 9359

Symbiotic Interaction

4th International Workshop, Symbiotic 2015
Berlin, Germany, October 7–8, 2015
Proceedings

Springer

Lecture Notes in Computer Science 9359

More information about this series at http://www.springer.com/series/7409

Benjamin Blankertz · Giulio Jacucci
Luciano Gamberini · Anna Spagnolli
Jonathan Freeman (Eds.)

Symbiotic Interaction

4th International Workshop, Symbiotic 2015
Berlin, Germany, October 7–8, 2015
Proceedings

Springer

Editors

Benjamin Blankertz
TU Berlin, MAR 4-3
Berlin
Germany

Anna Spagnolli
Università di Padova
Padua
Italy

Giulio Jacucci
Department of Computer Science
University of Helsinki
Helsinki
Finland

Jonathan Freeman
Department of Psychology
Goldsmiths University
London
UK

Luciano Gamberini
Università di Padova
Padua
Italy

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-319-24916-2 ISBN 978-3-319-24917-9 (eBook)
DOI 10.1007/978-3-319-24917-9

Library of Congress Control Number: 2015949468

LNCS Sublibrary: SL3 – Information Systems and Applications, incl. Internet/Web, and HCI

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media
(www.springer.com)

Preface

This volume contains the papers presented at Symbiotic 2015: the International Workshop on Symbiotic Interaction held during October 7–8, 2015 at the TU Berlin.

Symbiotic 2015 was the fourth edition after the first held at the University of Padova during December 3–4, 2012, the second held on December 12, 2013, at Goldsmiths, University of London, and the third held at the University of Helsinki during October 30–31, 2014. For Symbiotic 2015, we solicited 23 high-quality submissions in three categories: papers, posters, and demos. The workshop gathered a long list of important scholars in many disciplines (see Program Committee list), and each anonymous paper was reviewed by three members. We accepted 11 full papers and eight short papers.

We believe that Symbiotic will continue to grow and attract more interest from disparate fields with the aim of investigating future relationships between computers and humans. Symbiotic 2015 was sponsored by the MindSee Project (http://mindsee.eu/) and was partially funded by the European Community (FP7 – ICT; Grant Agreement 611570).

August 2015

Benjamin Blankertz
Giulio Jacucci
Anna Spagnolli
Luciano Gamberini
Jonathan Freeman

Organization

Program Committee

Laura Acqualagna	Technische Universität Berlin, Germany
Oswald Barral	University of Helsinki, Finland
Benjamin Blankertz	Technische Universität Berlin, Germany
Marc Cavazza	Teesside University, Middlesbrough, UK
Ricardo Chavarriaga	École polytechnique fédérale de Lausanne, Switzerland
Cristina Conati	University of British Columbia,Vancouver, Canada
Mauro Conti	University of Padua, Italy
Peter Desain	Radboud University Nijmegen, The Netherlands
Manuel Eugster	Aalto University, Denmark
Stephen Fairclough	Liverpool John Moores University, Liverpool, UK
Siamac Fazli	Cognitive Neurotechnology Lab, Korea University
Jonathan Freeman	Goldsmiths University, London, UK
Luciano Gamberini	HIT Research Centre, University of Padua, Italy
Hans Gellersen	Lancaster University, Lancaster, UK
Dorota Glowacka	University of Helsinki, Finland
Giulio Jacucci	University of Helsinki, Finland
David Kirsh	UC San Diego, La Jolla, CA, USA
Juan C. Moreno	Cajal Institute, Spanish National Research Council
Gernot R. Müller-Putz	Graz University of Technology, Austria
Anton Nijholt	University of Twente, The Netherlands
Andreas Riener	Johannes Kepler University Linz, Austria
Tuukka Ruotsalo	Aalto University Espoo, Finland
Paul Sajda	Columbia University, New York, USA
Erin Solovey	Drexel University, Philadelphia, USA
Michiel Sovijärvi-Spapé	Aalto University, Espoo, Finland
Anna Spagnolli	University of Padua, Italy
Paul Verschure	Universitat Pompeu Fabra, Barcelona, Spain
Markus Wenzel	Technische Universität Berlin, Germany
Thorsten Zander	Team PhyPA – TU Berlin, Germany

Contents

EEG Filtering Optimization for Code–Modulated Chromatic Visual Evoked Potential–Based Brain–Computer Interface

Daiki Aminaka, Shoji Makino, and Tomasz M. Rutkowski(✉)

Life Science Center of TARA at University of Tsukuba,
1-1-1 Tennodai, Tsukuba, Ibaraki, Japan
tomek@bci-lab.info
http://bci-lab.info/

Abstract. We present visual BCI classification accuracy improved results after application of high– and low–pass filters to an electroencephalogram (EEG) containing code–modulated visual evoked potentials (cVEPs). The cVEP responses are applied for the brain–computer interface (BCI) in four commands paradigm mode. The purpose of this project is to enhance BCI accuracy using only the single trial cVEP response. We also aim at identification of the most discriminable EEG bands suitable for the broadband visual stimuli. We report results from a pilot study optimizing the EEG filtering using infinite impulse response filters in application to feature extraction for a linear support vector machine (SVM) classification method. The goal of the presented study is to develop a faster and more reliable BCI to further enhance the symbiotic relationships between humans and computers.

Keywords: Brain–computer interface · ERP · EEG classification · cVEP

1 Introduction

A brain computer interface (BCI) is a symbiotic device which facilitates human–machine interaction without dependence on any muscle or peripheral nervous system actions [7]. BCI employs human neurophysiological signals for a straight brainwave–based communication of a human with an external environment. Particularly, in the case of patients suffering from locked–in–syndrome (LIS) [4], amyotrophic lateral sclerosis (ALS) or coma, BCI could help them to communicate or complete various daily tasks (type letters or control their environments using Internet of Things technologies, etc.). The BCI shall create a feasible option for such patients to communicate with their families, friends or caretakers by using their trained and properly classified brainwaves only [7].

A code modulated visual evoked potential (cVEP) is proposed in this paper as a brain–computer interface (BCI) paradigm. The cVEP is a natural response to a visual stimulus generated with specific code–modulated, and also enhanced

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 1–6, 2015.
DOI: 10.1007/978-3-319-24917-9_1

with color modulation, sequences [2,3] while the user gases at the light source. The cVEP–based BCI is a stimulus–driven paradigm which does not require a long training, as compared to the imagery–driven paradigm [7].

Usually, cVEP's advantage is in its faster classification time comparing to other types of visual–BCIs using steady state visual evoked potentials (SSVEPs) or P300 responses. Theoretically a single classification interval could take less than 387.5 ms in our experiments, but actually the cVEPs have to be averaged to remove EEG noise, which multiplies the above mentioned minimum period. Usually the averaging procedure can take longer time, for example 1.9375 seconds as in our previous study based on five cVEPs' averaging [1], which limits this paradigm's advantage. In this paper, we present results of classification improvement after application of high– and low—pass filtering of EEG to create the faster cVEP–based BCI. A linear support vector machine (SVM) classifier is applied in the presented cVEP–based BCI research project.

The cVEPs used in this project are induced by four RGB light–emitting diodes (LEDs). We also utilize the higher flashing carrier frequency of 40 Hz (which is amplitude modulated with the proposed $m-sequences$) comparing to the classical setting of 30 Hz (limited to compare results with classical computer displays usually with 60 Hz refreshing rate) [2]. There are maximum of five consecutive positive pulses (continues light) and minimum of one positive/negative pulse of the LEDs in this experiment settings. If cVEP's frequency features would be evoked similarly as in a case of SSVEP, the steady–state response suppose shall appear in EEG frequency bandwidths of $6 \sim 30$ Hz or $8 \sim 40$ Hz according to our hypothesis. In other words, low–pass filtering with a cutoff frequency of 30 Hz or 40 Hz shall do the best job to remove unnecessary higher frequencies from EEG. Moreover, we propose to use chromatic green–blue stimuli [6] as a further extension in our project. We also compare our results with the classical monochromatic (white–black) set–up.

From now on the paper is organized as follows. In the following section we describe materials and methods used in this study. Next, results and discussion are presented. Conclusions together with future research directions summarize the paper.

2 Materials and Methods

The experiments reported in this paper were performed in the Life Science Center of TARA, University of Tsukuba, Japan, and they were approved by the ethical committee of the Graduate School of Systems and Information Engineering at University of Tsukuba, Tsukuba, Japan (experimental permission no. 2013R7). The subjects agreed voluntarily to participate in the study. The visual stimulus generating LEDs were driven by square waves delivered from *ARDUINO UNO* micro–controller board. We used $m-sequence$ encoded flashing patterns [3] to create four commands of the cVEP–based BCI. The binary pseudorandom string $m-sequence$ with a length of 31 bits was used as follows [0100100001010111011000111110011]. The special feature of the $m-sequence$,

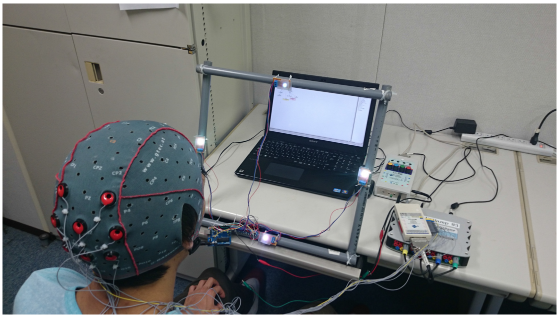


Fig. 1. The user seating in front of a frame with four visual stimulation chromatic LEDs used in this study. The picture was included with a permission of the photographed user.

which has been useful for the cVEP–based BCI paradigm design, was an unique autocorrelation function. The autocorrelation function had only a single peak at the $m - sequence's$ period. It was thus possible to introduce a circular shift of the $m - sequence$ denoted by τ, to create a set of another sequences with shifted autocorrelation functions, respectively. In this study, the shifted time length has been defined as $\tau = 7$ bits. Three additional sequences have been generated using shifts of τ, $2 \cdot \tau$ and $3 \cdot \tau$, respectively. During the online cVEP–based BCI experiments the four LEDs continued to flash simultaneously using the time–shifted $m - sequences$ as explained above. Two $m - sequence$ period lengths have been tested to investigate whether they would affect the cVEP response discriminability. The conventional full $m - sequence$ period of $T = 516.7$ ms, as in case of a conventional computer display with a refresh rate of 60 Hz (referred here as "a low flashing frequency") and the proposed $T = 387.5$ ms (referred as "a high flashing frequency") have been tested. The LED–based visual stimulus generator is presented in Figure 1. During the cVEP–based BCI EEG experiments the users were seated on a comfortable chair in front of the LEDs (see Figure 1). The distance between user's eyes and LEDs was about $30 \sim 50$ cm

Table 1. EEG signals recording conditions

Number of users	9 (8 males and 1 female)
Average age of users	26.4 years old (standard deviation of 7.0 years)
Single session length	8 and 11 s
$m - sequence$ length T	516.7 and 387.5 ms
$m - sequence$ shifts τ	116.7 and 87.5 ms
EEG amplifier	g.USBamp by g.tec with wet active g.LADYbird electrodes
Electrode locations	$O1, O2, Po3, Po4, P1, P2, Oz$ and Poz
Reference and ground	Left earlobe and FPz
Sampling frequency	512 Hz
Notch filter	Butterworth 4^{th} order stopping $48 \sim 52$ Hz
Band–pass filter	Butterworth 8^{th} order with a passband of $5 \sim 100$ Hz

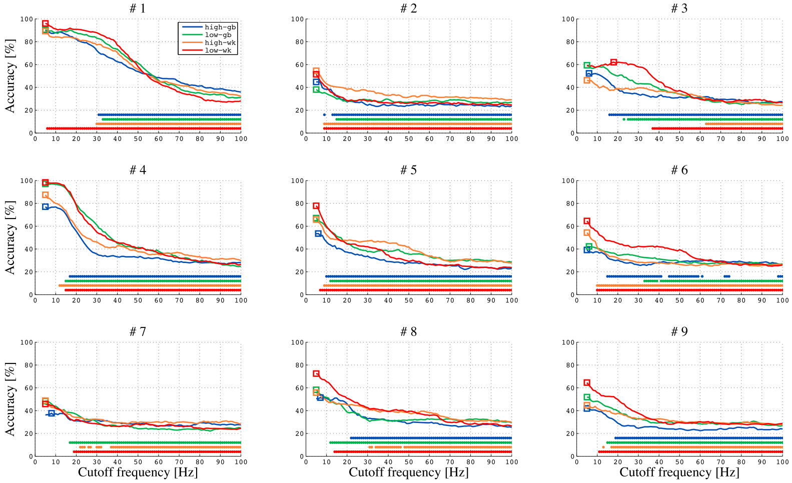


Fig. 2. The mean accuracy results of SVM–based classification after high–pass filtering. There are four results depicted for each user, namely from green–blue high carrier frequency (blue lines); low carrier frequency (green lines); white–black high carrier frequency (orange lines); low carrier frequency (red lines), respectively. Square markers show the maximum accuracies. Four horizontal lines, or dots, at the bottom of each panel depict the significant differences of classification accuracies between the non–filtered (raw EEG signals, of which accuracies are not shown here) and the filtered cVEPs ($p < 0.05$ of Wilcoxon–test). The theoretical chance level of the experiments was of 25%.

(chosen by the users for a comfortable view of the all LEDs). A notch filter was applied to remove power line interference of 50 Hz from EEG together with a band–pass filter to remove eye blinks and muscle–originating noise. Details of the EEG experimental set up are summarized in Table 1. To avoid user's eye blinks, each trial to gaze at a single LED was separated with pauses. The 60 cVEPs were collected for each of four LED flashing targets. An OpenViBE [5] bio–signal data acquisition and processing environment, together with in–house programmed in Python extensions, were applied to realize the online cVEP–based BCI paradigm. In the data acquisition phase, user gazed at four LEDs as instructed. The cVEPs to top LED were firstly collected for the classifier training and other were used for testing. The triggers indicating the onsets of the $m-sequences$ were sent to the amplifier directly from the *ARDUINO UNO* micro–controller to mark the beginning of each cVEP response. A linear SVM classifier was used in this study to identify which of the flickering patterns the user was gazing at. The cVEP response processing and classification steps were as follows: (i) for training purpose, the EEG cVEP responses to the top flashing LED (m–sequence with $\tau = 0$) were defined as $Y(t)$ and another three cVEPs (responses to bottom, right and left LEDs as shown in Figure 1) were created by circular shifting of the original $Y(t)$ by τ, $2 \cdot \tau$ and $3 \cdot \tau$ respectively; (ii) high–pass Butterworth IIR filters were applied to EEG with cutoff frequencies

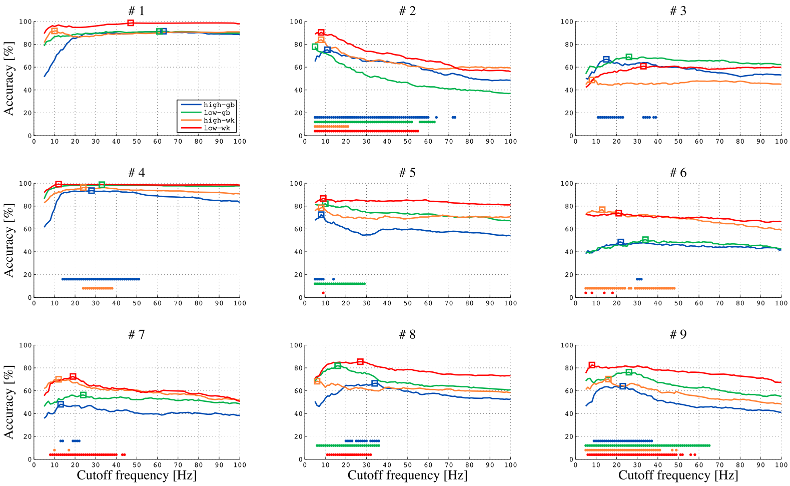


Fig. 3. The mean accuracy results of SVM–based classification after low–pass filtering. There are four results depicted for each user, namely from green–blue high carrier frequency (blue lines); low carrier frequency (green lines); white–black high carrier frequency (orange lines); low carrier frequency (red lines), respectively. Square markers show the maximum accuracies. Four horizontal lines, or dots, at the bottom of each panel depict the significant differences of classification accuracies between the non–filtered (raw EEG signals, of which accuracies are not shown here) and the filtered cVEPs ($p < 0.05$ of Wilcoxon–test). The theoretical chance level of the experiments was of 25%.

of a and b Hz, where $a \in \{6, 7, \ldots, 100\}$ Hz; (iii) four–class linear SVM classifier was trained using 60 filtered cVEPs for each flashing target, respectively; (iv) high–pass filters were applied similarly as in (ii) to EEG for testing dataset with 60 filtered cVEPs to four target $m - sequences$ linear SVM evaluations; (v) the above steps (ii)–(iv) were applied for the frequencies $a = 5, 6, \ldots, 100$ Hz. The above procedure steps (i)–(v) were also repeated by switching testing and training cVEPs to the top LED. Finally, four experiment types were conducted for each user by employing: the conventional low frequency; the proposed high frequency; and in each of the above setting in the two color modes with white–black and green–blue flashing LEDs.

3 Results

Results of the conducted cVEP–based BCI paradigm experiments are summarized in Figures 2 and 3. The accuracies were calculated for cVEPs induced by four types of stimulations as mentioned in previous section. The theoretical chance level of all experiments was of 25%. In the case of Figure 2, the mean high–pass filter cutoff frequency of four maximum classification accuracies for each user was of 5.58 Hz (standard deviation of 2.22 Hz). The significant differences of the above accuracies, as tested with pairwise Wilcoxon–test, between non–filtered and filtered cVEPs ($p < 0.05$) were observed as shown in form of horizontal lines in at the bottom of each panel in Figure 2. We next applied low–pass filtering to EEG for identifying the higher frequency features, which resulted with BCI classification accuracies as shown in Figure 3. Except for subject #1, the results have shown that low–pass cut–off frequencies within a range of $10 \sim 30$ Hz scored the best for all the stimulation types. The mean cutoff frequency of four maximum classification accuracies for each user was of 20.58 Hz (standard deviation of 14.32 Hz). There were significant differences among non–filtered and filtered result ($p < 0.05$), as evaluated with Wilcoxon–test, yet the frequencies values we user–dependent as shown in Figure 3.

4 Conclusions

The proposed LED flashing and cVEP response–based BCI paradigm with the chromatic (green–blue) stimulus has been discussed in this paper. We tested and optimized high– and low–pass filters for cVEP–based BCI accuracy improvement using linear SVM classifier. The conducted experiments verified the optimal filter bandwidth for the proposed cVEP feature extraction within the mean range of $5.58 \sim 20.58$ Hz (which shall round up to $6 \sim 21$ Hz taking into account the exact frequency steps used in the study). We originally hypothesized that the low–pass filleting at 30 or 40 Hz cutoff frequencies shall do the good job for cVEP unrelated noise removal, but the results of the presented experiments have shown that much lower cutoff frequencies of about 21 Hz are also feasible. For the future research, we plan to investigate further details of frequency features of cVEP, which is a broadband signal due to it's square wave pseudo–random components.

References

1. Aminaka, D., Makino, S., Rutkowski, T.M.: Chromatic and high-frequency cVEP-based BCI paradigm. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE Engineering in Medicine and Biology Society (2015). (accepted, in press)
2. Bin, G., Gao, X., Wang, Y., Hong, B., Gao, S.: VEP-based brain-computer interfaces: time, frequency, and code modulations [Research Frontier]. IEEE Computational Intelligence Magazine **4**(4), 22–26 (2009)
3. Bin, G., Gao, X., Wang, Y., Li, Y., Hong, B., Gao, S.: A high-speed BCI based on code modulation VEP. Journal of Neural Engineering **8**(2), 025015 (2011)
4. Plum, F., Posner, J.B.: The Diagnosis of Stupor and Coma. FA Davis, Philadelphia (1966)
5. Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., Lécuyer, A.: Openvibe: an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. Presence: Teleoperators and Virtual Environments **19**(1), 35–53 (2010)
6. Sakurada, T., Kawase, T., Komatsu, T., Kansaku, K.: Use of high-frequency visual stimuli above the critical flicker frequency in a SSVEP-based BMI. Clinical Neurophysiology (2014). (online first)
7. Wolpaw, J., Wolpaw, E.W. (eds.): Brain-Computer Interfaces: Principles and Practice. Oxford University Press (2012)

Using fNIRS for Prefrontal-Asymmetry Neurofeedback: Methods and Challenges

Gabor Aranyi, Marc Cavazza(✉), and Fred Charles

School of Computing, Teesside University, Borough Road, Middlesbrough TS1 3BA, UK
{g.aranyi,m.o.cavazza,f.charles}@tees.ac.uk

Abstract. Functional near-infrared spectroscopy (fNIRS) has become increasingly accessible in recent years, which allows this relatively low-cost and portable brain sensing modality for the application of brain-computer interfaces (BCI). Although there is a growing body of research on fNIRS-based BCI utilising users' covert psychophysiological activity, there is comparably less research on active BCI, where users engage in thinking strategies with the explicit intention of controlling the behaviour of an interactive system. We draw on four empirical studies, where participants received real-time neurofeedback (NF) of left-asymmetric increase in activation in their dorsolateral prefrontal cortex (DL-PFC), which has previously been identified as a correlate of approach-related motivational tendencies. We discuss methodological considerations and challenges, and provide recommendations about brain-signal selection and integration, NF protocol design, post-hoc and real-time applications of NF success criteria, continuous visual feedback, and individualised feedback based on the variations of the brain-signal in a reference condition.

Keywords: Prefrontal asymmetry · Functional near-infrared spectroscopy · (Affective) brain-computer interfaces · Neurofeedback

1 Introduction

The development of affective brain-computer interfaces (BCI) is a relatively recent trend [11], which has the potential to support new interactive systems, human-robot interaction [9], and cultural [14] or entertainment applications. One of the main challenges faced by affective BCI, especially if compared with BCI based on motor areas, is to identify a clear mapping between a target affective state and a BCI signal. This could be obtained from knowledge about the neural localisation of affective states: however, such knowledge is particularly elusive, as suggested by a recent review [15].

However, research in psychophysiology has identified a possible neural correlate for a specific affective dimension, known as approach/withdrawal [4], in the form of asymmetric activity in the prefrontal cortex (PFC). Approach/withdrawal behaves as a high-level affective dimension, and has been shown to play an important role in motivational processes, reward expectation, risk taking and

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 7–20, 2015.
DOI: 10.1007/978-3-319-24917-9_2

depression. It has originally been explored through EEG studies, which have defined asymmetry scores that can characterise this asymmetry. These behave as individual traits but are also subject to dynamic variations: furthermore, they can be controlled through neurofeedback (NF), as originally demonstrated by Rosenfeld and colleagues [19]. The potential use of prefrontal asymmetry to support affective BCI has been discussed in various reviews [17], although without reporting specific implementations.

Early work on using prefrontal asymmetry for BCI was based mostly on EEG signals. Wehbe et al. [27] reported passive measurement of EEG prefrontal asymmetry during computer gameplay; however, they claimed to be using it as a measure of arousal rather than approach. Karran et al. [14] explored the role of the PFC in subjects' aesthetic experiences. Our previous work explored EEG-based NF in the alpha band, with simultaneous fMRI analysis over NF epochs [12]. It confirmed that the affective strategies through which users controlled PFC alpha asymmetry corresponded to asymmetric activity in prefrontal regions (across areas BA9 and BA10), with no differences observed in pre-motor areas. These experiments are difficult to interpret any further, due to the small number of subjects, and the finding that the supine position is known to impair the ability of subjects to properly express approach [4]. In further experiments carried out in laboratory conditions, subjects achieved success rates of up to 73% with minimal training [7]. However, signal quality and stability during NF epochs remains an issue, and a limiting factor.

We posit that functional near-infrared spectroscopy (fNIRS) can provide an alternative, offering better signal quality and better resistance to motion artefacts, while also improving spatial resolution for the target brain areas. This is also supported by the finding that areas relevant to approach/withdrawal include the dorsolateral prefrontal cortex (DL-PFC) [25], whose localisation is accessible to fNIRS. Sitaram et al. [23] were amongst the first to suggest that signals based on metabolic activity could be equally suited to BCI than electrical signals. We were also inspired by recent experiments by Zotev et al. [28], which reported PFC-NF with both EEG and real-time fMRI. Finally, Naseer and Hong [18] have also reviewed recent uses of fNIRS for NF.

In this paper, we discuss methodological aspects of deploying fNIRS to implement BCI based on prefrontal asymmetry, under a NF paradigm. These are based on several experiments, one published [5] and the others accepted for publication or under review, during which we explored various settings for controlling approach, in affective contexts as diverse as empathy, anger or motivation. Rather than reproducing these studies here, we shall concentrate on specific elements of methodology, some common to all studies, such as optode selection and signal definition, and some more specific, such as the definition of baseline, control tasks compared to NF epochs, and calculations of statistical significance, both online and post hoc.

2 Brain-Signal Acquisition, Selection and Integration

We used fNIRS to operationalise BCI input based on asymmetric functional activation in the DL-PFC. Although the spatial resolution of fNIRS falls short that of fMRI and is limited to scanning the outer cortex, it has a number of advantages, such as lower susceptibility to motion artefacts and lower cost, that make it appropriate for application in BCI [3]. We followed the recommendation of Solovey et al. [24] for the use of fNIRS in HCI settings. We used an fNIR400 Optical Brain Imaging Station by Biopac Systems, with a 16-channel sensor with fixed 2.5cm source-detector separation (see [20] for channel locations). Data were collected with 2Hz sampling rate. This fNIRS device measures intensity changes in two wavelengths (730nm and 850nm) over time to calculate the change in oxygenated (HbO) and deoxygenated (HbR) haemoglobin concentration (in units of μMol/L) using the modified Beer-Lambert Law [1]. In order to provide real-time feedback based on brain activity measured by fNIRS, there is a need to select a single metric: HbO, HbR, or HbT (total haemoglobin; the sum of HbO and HbR). We conducted a pilot study to inform this decision.

The pilot study used a no-feedback paradigm including an approach task (watching pictures of delicious food under the instructions to imagine reaching out for the food and eating it) and a withdrawal task (watching pictures of spiders under the instructions to imagine escaping from the situation)[1]. We compared the metrics of HbO, HbR, and HbT to assess how well they are able to discriminate between the tasks. Based on literature co-authored by the developer of the fNIRS system used in our experiments (e.g., Ruocco et al. [20]), on literature applying HbO to affect-related manipulation in the DL-PFC [26], and to approach/withdrawal-related experimental manipulation [16], and based on our pilot study, we elected to use HbO for real-time application; we based post-hoc analyses on the same metric for consistency.

The haemodynamic response measured by fNIRS takes several seconds [3]. We took two approaches to accommodate for this approximately 7s delay: we either (a) simply removed the first 14 data points (corresponding to 7s sampled at 2Hz) of each epoch on each channel, or we (b) also included the 14 data points after the completion of the epoch (i.e., windowing; see Sarkheil et al. [22] for a similar approach).

The complexity of measured changes in blood oxygenation associated with the differential functional activation of the DL-PFC needs to be reduced to afford effective BCI input. This consists in deriving a single asymmetry metric from the continuous flow of oxygenation data from the input channels [8]. We averaged HbO values over the four leftmost and four rightmost channels (located over the left and right DL-PFC, respectively), then subtracted average Right from average Left. This metric reflects the inter-hemispheric difference in HbO change

[1] Pilot subjects (N = 4) confirmed positive attitude towards the food items and negative towards spiders. We additionally included an approach condition involving the same spider pictures with instructions to imagine approaching the spiders in protective clothing and swatting them.

in micromolar units (μMol/L). Note that this measure is relative to a baseline [2], and more importantly, it lacks an absolute zero point, as opposed to, for example, alpha-power asymmetry in EEG-based NF [6]. This has important practical consequences in defining and quantifying NF success. For example, as this operationalisation of asymmetry yields interval-level data, a ratio of task/no-task signals for defining and quantifying success (e.g., [22]) cannot be applied.

3 Protocol Design Considerations

There is a growing body of research on fNIRS-based NF [18]; however, fMRI-based NF research can also effectively inform fNIRS study design due to the comparability of the haemodynamic signal measured by the two neuroimaging modalities (see [3] for a comparison). We sought inspiration from fMRI-based studies [22,28] to inform study design, because of their relevance to the target mental activity (affective regulation), experimental task (up-regulation of activity in a target area using thinking strategies), and feedback operationalisation.

Protocols for experiments and interaction design for active BCI need to be tailored for supporting a feedback strategy and applying a success criterion, depending on the tasks the participants are required to carry out for interacting with the system. The length of individual epochs (short time periods with a specific task) and blocks (a sequence of epochs), and that of the entire protocol (number of blocks), needs to be manageable for participants, while it also needs to provide a sufficient quantity of data for the purpose of research and application. These considerations place constraints on how much data is to be collected, and how data collection can be structured in a way that it is most manageable for participants.

With regards to the length of each block and the overall number of blocks, the two main considerations are (a) the participants' ability to maintain focus on the mental activity, and (b) the amount of data necessary to support the use of the success criterion. To address (a), we asked participants to provide subjective difficulty ratings of each task involved in the experiments, and we also conducted post-use interviews to gather qualitative data about their interaction with the system. Regarding (b), when using a statistical success criterion, discussed in detail in the following section, we advise conducting a power analysis to determine the number of observations required within each epoch to detect an effect with a given magnitude (some data may need to be discarded in the filtering process).

These considerations are inherently related to how feedback is provided and NF success is defined. In the following sections, we illustrate the approach we took to addressing these challenges through two sets of experiments.

4 Criterion for Neurofeedback Success

We applied a statistical criterion to determine NF success. Specifically, we characterised NF success as a statistically significant increase in left-asymmetry during a NF epoch, compared to either (a) a baseline or (b) a reference epoch.

Since statistical significance depends on the sample size, or in the present case, the number of observations in a time-series we refer to as an epoch, we also calculated different effect-size measures (r and Cohen's d) to characterise the magnitude of NF success. We also implemented real-time, automated determination of NF success as well as post-hoc testing. Additionally, we explored if applying a non-parametric approach (bootstrapping) delivers practical benefits. We discuss the advantages and disadvantages of these approaches in the following sections through two sets of experiments. We describe the experimental protocols designed to support the application of the statistical success criterion, with variations along the following properties:

- Success evaluated against zero asymmetry or asymmetry during a reference epoch;
- Success evaluated real-time or post-hoc;
- Treatment of delay (trimming or windowing);
- Threshold characterisation and feedback mapping (fixed or personalised).

4.1 Evaluating NF Success Against Baseline

With the fNIRS system we applied, baseline is measured over 10s, against which the asymmetry scores collected during a NF epoch can be compared, without the need for a reference epoch. Since this baseline measurement consists in collecting light-intensity data on each channel as a reference for calculating oxygenation changes [1], the asymmetry metric we derive from the channels is zero for the baseline. Therefore, we determined NF success using the baseline criterion by performing a one-sample t-test on the asymmetry scores collected during the NF epoch against the test value zero. Performing this one-sample t-test [10] is computationally simple and can be implemented real-time by calculating the t value upon the completion of the NF epoch by dividing the mean of observed asymmetry values by the estimate of the standard error (Fig. 1a), which is then compared to a critical value to determine NF success.

We conducted two experiments using this success criterion (Fig. 2). In Experiment 1a, success was determined real-time, based on unfiltered data, using a parametric criterion, delay was treated by trimming, while threshold and mapping for the feedback channel were fixed (i.e. the same for each subject and each experimental trial). By comparison, in Experiment 1b, success was determined post-hoc, based on filtered data, using a distribution-free criterion, delay was treated by windowing, while threshold and mapping were also fixed.

Experiment 1a used 33s long NF epochs that contained 66 observations (2Hz sampling frequency), therefore we applied the t critical value for $p = .05$ (two-tailed) with 65 degrees of freedom (df) for each block: $t_{crit}(65) = 2.00$. The experimental software logged asymmetry values during the NF epoch, calculated the t value, and if it was larger than 2, the block was deemed successful. The experimental software did not test the parametric assumption of normality; however, post-hoc analyses using bootstrapping resampling method resulted in accepting the same epochs as successful.

(a) $$t = \frac{\bar{x} - \mu_0}{SD/\sqrt{n}}$$ (b) $$r = \sqrt{\frac{t^2}{t^2 + df}}$$ (c) $$d = \frac{\overline{NF} - \overline{Ref}}{\sqrt{\frac{(n_{NF} - 1)SD_{NF}^2 + (n_{Ref} - 1)SD_{Ref}^2}{n_{NF} + n_{Ref} - 2}}}$$

Fig. 1. Equations used in real-time implementation of NF success, where $\overline{x}$ is the mean of observed values, the test value μ_0 is zero, SD is the standard deviation, n is the number of observed values, and df is degrees of freedom. $\overline{NF}$ and $\overline{Ref}$ are the mean of asymmetry values during NF and the reference epoch, respectively.

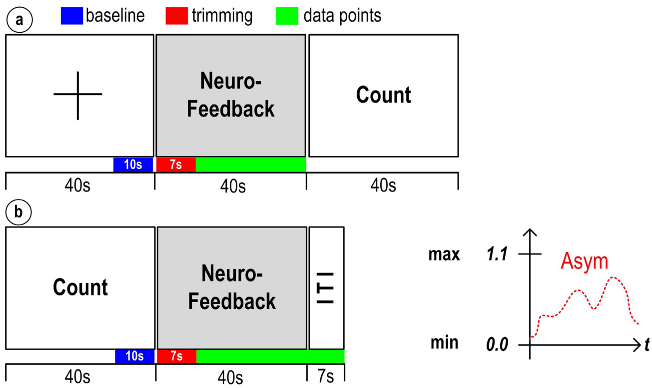


Fig. 2. Block design for Experiments 1a and 1b, where NF success is evaluated against baseline. Note that in 1b, an inter-trial interval (ITI) is added to the end to allow for windowing; baseline in 1b is measured during the reference task (see text), decreasing block length.

We argue that the computationally more demanding bootstrapping method should be favoured for post-hoc analysis, but in cases where real-time determination of NF success is important, the simple parametric criterion is sufficient. Experiment 1a also used the magnitude of the NF signal in a successful epoch as graded input to a computer system (mapped to the differential weighting of a search algorithm [5]). This was achieved by characterising the magnitude of NF success by calculating the effect-size measure r (Fig. 1b). Note that this calculation is not computationally demanding; therefore, it can be applied in real-time too. The effect-size measure r is interpreted the same as the correlation coefficient, it mitigates the difficulty of comparing fNIRS signals across subjects and blocks [21], and since its value is constrained between 0 and 1, it is convenient for mapping to graded input.

Another advantage of calculating an effect-size measure when a statistical criterion is applied to determine NF success is that it allows for evaluating the sensitivity of the set-up to detect changes in the asymmetry signal by quantifying the magnitude of increase in asymmetry the applied statistical criterion can detect. For example, in Experiment 1a, the smallest effect-size associated with a

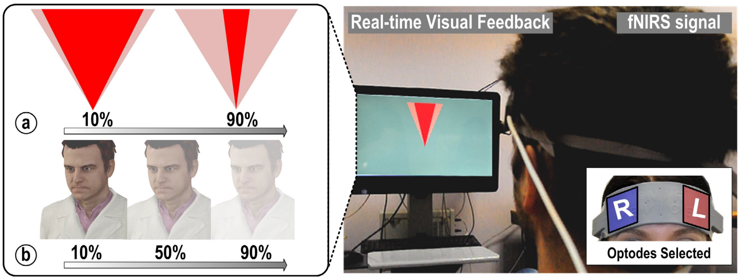


Fig. 3. Examples of continuous visual feedback in (a) Experiment 1a and (b) Experiment 2a.

successful block was r = .28, which demonstrates that we could reliably detect medium effect-sizes with 33s long NF epochs. Issues related to low power (i.e. asymmetry increases but it is not detected) can be mitigated by increasing the number of observations in an NF epoch, either by increasing its length or the sampling frequency.

Threshold for providing feedback during the NF epoch (i.e., the minimum reinforced signal magnitude) and mapping the magnitude of the left-asymmetry signal to feedback was fixed in both experiments, that is, each subject in each trial received feedback using a set of pre-defined parameters. The visual feedback channel in both experiments was a downward-pointing red triangle symbolising a light beam, which could be narrowed by up-regulating left-asymmetry (Fig. 3a). This feedback was conceptually related to the experimental context, which involved speeding up an algorithmic search process using a BCI [5][2]. As discussed above, the threshold was zero (i.e., no increase from baseline). We defined the maximum value for feedback empirically in a pilot study (1.1μMol/L), using a similar design under a no-feedback paradigm. Asymmetry values between the threshold and maximum were mapped linearly to the width of the light beam (updated with the same 2Hz frequency of the signal acquisition), which allows for providing continuous feedback. We successfully applied a similar feedback mapping strategy before in EEG-based NF [7].

A disadvantage of comparing asymmetry scores to a simple baseline to determine NF success is related to the difficulty of interpreting the increase in asymmetry during the NF epoch. As mentioned above, we measured a 10s baseline before each NF epoch, which defined the asymmetry as zero for the start of NF. However, the appropriateness of this baseline is predicated on the assumption that the asymmetry signal is at neutral level when the baseline is taken. Should the baseline be measured when there is high left-asymmetry, the reference-point zero at the start of NF would represent a state of already high left-asymmetry, making it difficult to detect left-asymmetry increase during NF.

[2] In short, this experiment used PFC left-asymmetry as an indicator of approach-related motivational tendency, whose value was mapped to speeding up the behaviour of a search algorithm.

To overcome this, we designed data-collection blocks in Experiment 1a in the following way. Each block consisted of three epochs: NF, Count and Rest (Fig. 2a). Baseline for an NF epoch was measured in the last 10s of the preceding Rest epoch, where subjects were instructed to look at a grey screen and relax. An epoch with a mental counting task was included after each NF epoch to distract subjects' attention from the thinking strategy used during NF and to promote asymmetry converging to baseline before baseline for the next block would be taken. We elected to use a mental counting task (counting backwards from a given number by increments of a given integer), because it is theoretically unrelated to left-asymmetry and it is one of the most commonly used prefrontal activities for fNIRS-based BCI [18].

In Experiment 1b, we modified the block design by excluding the Rest epoch and measuring baseline for the next block in the last 10s of the Count epoch following NF (Fig. 2b). This reduced block length, allowing for including more blocks in the same protocol. Furthermore, including the counting task during baseline measurement is a more strict control of subjects' mental activity than the rest instructions. Additionally, Experiment 1b did not use the NF signal for graded input to modify the behaviour of a system; therefore we determined NF success post-hoc using bootstrapping (1000 samples, 95% confidence intervals) on filtered data: we applied sliding-window motion artefact detection (SMAR), raw data were low-pass filtered using a finite impulse response (FIR) filter with order 20 and 0.1Hz cut-off frequency [1].

In summary, determining NF success by comparing to baseline allows for a simple block structure, but the baseline may not reflect a meaningful reference point to determine NF success if thought processes are not controlled during baseline measurement. Including a reference task for baseline may alleviate this, but other potential issues remain, for example, the perceptual differences of the stimulus subjects receive during baseline and NF, and having to rely on the same (pre-defined) criteria across blocks and subjects to provide feedback. In the next section, we discuss how including a reference epoch may improve study design.

4.2 Evaluating NF Success Against a Reference Epoch

As discussed above, although comparing an NF epoch to baseline for defining success is simple and time efficient, it is still useful to include a reference task either to promote up-regulated left-asymmetry to converge to baseline between blocks (Experiment 1a), or to control thought processes during baseline (Experiment 1b). However, both leave the data collected for the reference task underanalysed. A step forward to better utilising the collected data is to include the reference task in a separate epoch that is directly compared to the NF epoch. We implemented this in two experiments.

In Experiment 2a (Fig. 4a), success was determined real-time, based on data filtered for extreme values, using a parametric criterion, delay was treated by trimming, while threshold and mapping for the feedback channel were individualised (for each subject and each experimental trial). By comparison, in Experiment 2b (Fig. 4b), success was determined post-hoc, on filtered data, using a

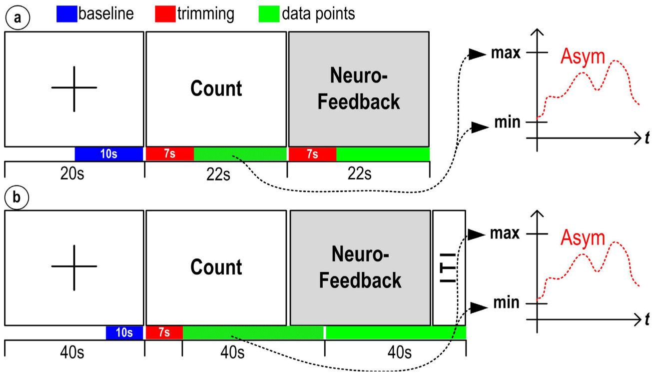


Fig. 4. Experiments 2a and 2b, where NF success is evaluated against a reference epoch. Note that in 2b, an inter-trial interval (ITI) is added to the end to allow for windowing. Feedback range in the NF epoch is determined using the distribution of asymmetry values during the reference epoch in each block to allow for individualised feedback.

distribution-free criterion; delay was treated by windowing, with individualised threshold and mapping. Both Experiments 2a and 2b applied the same visual stimulus across reference and NF epochs (but the visual stimulus was different across the experiments). The epochs were matched for length. Subjects rated the perceived difficulty of both the counting and NF tasks; statistical analysis revealed no significant difference in subjective difficulty, indicating that the two tasks were adequately matched. Baseline was measured at the start of each block under instructions to rest, but we defined NF success as a statistically significant increase in asymmetry from the reference epoch to the NF epoch.

A notable advantage of this approach to determining NF success is that it provides a control condition within the block; therefore, increase in asymmetry in this case can be readily attributed to change in mental activity. Furthermore, the asymmetry signal does not need to be at neutral level when the baseline is measured, because the success criterion only considers difference between the two epochs matched in length and stimulus.

Experiment 2a applied a real-time success criterion: the experimental software conducted an independent-samples t-test upon the completion of both epochs. Although the asymmetry values were collected from the same subject within the same block (with the same baseline), an independent-samples design is appropriate here, because the subject of analysis is the two population of asymmetry scores. The t value was calculated using unfiltered data; however, the experimental software removed outliers in each epoch (values outside three standard deviations from the mean) for the calculation, which can effectively remove noise resulting from movement artefacts [2]. Since effective epoch-length was 15s (trimmed), which contained at least 29 observations for each epoch sampled at 2Hz, the software used the t critical value of 2.05 with 28 degrees of freedom for p (two-tailed) = .05 as a threshold for success.

Conversely, Experiment 2b applied a post-hoc success criterion, where the t-test was calculated on filtered data (40s effective epoch length, windowed). In addition to the SMAR and FIR filters described in the previous section, we applied linear detrending on data from each channel [1]. Significance testing was conducted using a distribution-free approach.

In Experiments 1a and 1b, we used a 'one-size-fits-all' model for providing feedback, based on an empirically determined threshold for maximum feedback that was the same of each participant within each block. This approach promotes comparability of asymmetry values across blocks and subjects, but it does not consider individual differences, which can be quite substantial [21]. However, by analysing the distribution of asymmetry scores in a reference epoch, it is possible to devise personalised feedback within each block, taking into consideration the normal fluctuation of the asymmetry signal during a reference task. We recommend using a reference task (e.g., mental counting) for baseline if a simple set-up is preferred or there is an emphasis on collecting a large number of blocks from each participant. Otherwise, it is preferable to use a reference epoch with similar perceptual properties as the NF task, but including a different mental activity; this provides experimental control within a block, and allows for the application of individualised feedback to promote decreasing noise and increasing NF success. We implemented this in Experiments 2a and 2b in the following way.

We defined threshold for providing feedback during the NF epoch based on the asymmetry values collected during the reference epoch within the same block (Fig. 5). The threshold was defined as the mean of asymmetry values during the reference epoch plus 1.28 times their standard deviation[3]. Assuming normally distributed asymmetry values, this threshold would result in reinforcing only the top 10% of asymmetry values in the reference epoch. This approach to determine threshold is consistent with the original one of Rosenfeld et al. [19] for EEG-based frontal-asymmetry NF.

We provided continuous feedback, based on real-time changes in the magnitude of the asymmetry signal. For example, in Experiment 2a, the feedback channel was the image transparency of a virtual character (Fig. 3b), who was previously identified as mischievous, and the experimental subjects could make his image disappear from a virtual scene by expressing anger towards him, thereby up-regulating left-asymmetry [13]. Crossing the threshold during NF was mapped to 10% transparency of the virtual character, while reaching an empirically determined maximum asymmetry was mapped to 100% transparency, effectively removing the virtual character from the scene. The maximum asymmetry value for mapping was defined as the threshold plus the variation range of the asymmetry values during the reference epoch. Visual transparency was mapped linearly between the threshold and maximum value, updated with the same 2Hz frequency of the collection of asymmetry values.

[3] Outliers were removed by the experimental software to avoid extreme values, likely reflecting movement artefacts, exerting an unduly influence on the threshold.

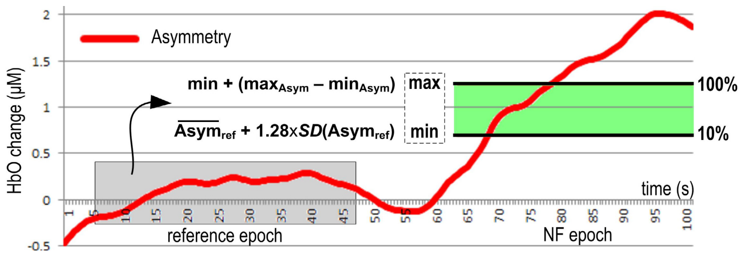


Fig. 5. Calculation of threshold (minimum) and maximum asymmetry values for linear mapping to continuous feedback. Note that the minimum feedback value is set to 10%.

This approach of using a reference epoch to determine feedback range promotes NF success by tailoring the feedback mechanism to the individual subject by considering her own signal variation, with the additional benefit of determining the range of noise in the signal for which no feedback should be provided. Note that in this approach the reference epoch necessarily precedes the NF epoch.

Additionally, we calculated the Cohen's d effect-size measure to quantify the magnitude of NF success, which is characterised as the difference between mean asymmetry during the two epochs divided by the pooled standard deviation. This was also calculated real-time in Experiment 2a (Fig. 1c). The d value reflects the distance between the distribution of asymmetry values between the reference and NF epoch within the same block, which can be readily interpreted. For example, average d in successful blocks in Experiment 2a was 2.40, which corresponds to an average 23% overlap in asymmetry scores between NF and reference epochs, and there is a 96% chance that an asymmetry value picked randomly from the NF epoch will be larger than a randomly picked asymmetry value from the reference epoch. Calculating these measures can be useful for illustrating the magnitude of asymmetry up-regulation in NF. Although the magnitude of oxygenation changes can differ substantially across subjects and blocks, this approach relies on the distribution of observed asymmetry values within blocks, therefore data collected from different individuals at different times are comparable.

5 Summary of Recommendations

We presented two sets of experiments where we applied variations on protocol design to support fNIRS-based PFC-asymmetry NF for active BCI. In this section, we briefly summarise the key points and present recommendations for protocol design. Table 1 illustrates these points through a comparison of two experiments.

Protocol Length. NF demands focused attention and concentration (but it is also rewarding and interesting), which leads to fatigue over time: approximately 40s long epochs and 2min per block are manageable, provided there is enough rest between the blocks. A total protocol length of approximately 10min is comfortable, where fatigue towards the end that does not significantly impair the quality

Table 1. Summary of the key points illustrated through contrasting two experiments.

	Exp. 1a	Exp. 2b
Reference epoch	No	Yes (counting)
Threshold	0 asymmetry	Dynamic ($M + 1.28 \cdot SD$)
Maximum	Fixed (1.1)	Dynamic (min+range)
Statistical test	Parametric	Bootstrapping
Success test	Real-time	Post-hoc
Filtering	No	Yes (FIR, SMAR, detrending)
Delay treatment	Remove 7s	Windowing
Practice	3 blocks	1 block
Number of blocks	6	8
Number of subjects	11	10
Success rate[1]	73%	70%

[1] Indicates the percentage of subjects achieving NF success in at least half of the completed experimental blocks [5,7].

of data (analysis revealed that NF success was not significantly less likely towards the end of the protocol in the studies). A shorter block length is generally preferable, which allows for including more blocks in a protocol. Based on post-use interviews with 42 subjects who participated in our experiments, we advise to determine the length and number of blocks so that data collection fits within approximately 15–20min with instructions, practice, set-up and calibration.

Compensate for Delay in the Haemodynamic Response. This can be achieved by simply removing initial observations in an epoch or by windowing. Simple removing may be considered a safer solution when it is important to make sure that each data point was collected when feedback was present, but windowing helps to better utilise the collected data by boosting statistical power with increased sample size.

Using a Reference Task. We recommend using a reference task (e.g., mental counting) for baseline if a simple set-up is preferred or there is an emphasis on collecting a large number of blocks from each participant. Otherwise, it is preferable to use a reference epoch with similar perceptual properties as the NF task; this provides experimental control within a block, and allows for the application of individualised feedback.

Using a Statistical Success Criterion. Rather than relying solely on statistical significance, we advise to quantify NF success (e.g., by calculating effect-size measures). If the NF signal also serves as input and a real-time success criterion is required, a parametric approach is sufficiently robust, but simple data-screening should still be applied (e.g., removing outliers). Otherwise, a post-hoc success criterion should be preferred on filtered data and using a distribution-free method. Conduct a power analysis to inform the necessary length for the NF epoch.

Continuous, Real-Time Feedback. Feedback can be provided with the same frequency as the input signal is collected. With an empirically determined threshold and signal range, the feedback can reflect continuous variations of the input signal, as opposed to a limited set of categories. We recommend using a feedback channel that is conceptually or perceptually related to the NF task. Participants can accommodate delay in the feedback, but they need to be informed in advance to expect some delay. Additionally, we found it useful to instruct participants that some jitter may be also present in the feedback; however, this was not problematic enough to introduce smoothing to reduce fluctuations in the NF signal (e.g., moving average [6,28]).

References

1. Ayaz, H.: Functional near infrared spectroscopy based brain computer interface. PhD thesis, School of Biomedical Engineering Science & Health Systems, Drexel University (2010)
2. Ayaz, H., Izzetoglu, M., Shewokis, P.A., Onaral, B.: Sliding-window motion artifact rejection for functional near-infrared spectroscopy. In: Engineering in Medicine and Biology Society (EMBC), pp. 6567–6570. IEEE (2010)
3. Bunce, S.C., Izzetoglu, M., Izzetoglu, K., Onaral, B., Pourrezaei, K.: Functional near-infrared spectroscopy. IEEE Engineering in Medicine and Biology Magazine **25**(4), 54–62 (2006)
4. Carver, C.S., Harmon-Jones, E.: Anger and approach: Reply to watson (2009) and to tomarken and zald (2009). Psychological Bulletin **135**(2), 215–217 (2009)
5. Cavazza, M., Aranyi, G., Charles, F.: Brain-computer interfacing to heuristic search: first results. In: Vicente, J.M.F., Álvarez-Sánchez, J.R., de la Paz López, F., Toledo-Moreo, F.J., Adeli, H. (eds.) IWINAC 2015, Part I. LNCS, vol. 9107, pp. 312–321. Springer, Heidelberg (2015)
6. Cavazza, M., Aranyi, G., Charles, F., Porteous, J., Gilroy, S., Jackont, G., Klovatch, I., Raz, G., Keynan, N.J., Cohen, A., et al.: Frontal alpha asymmetry neurofeedback for brain-computer interfaces. In: Brain-Computer Interface Conference (2014)
7. Cavazza, M., Aranyi, G., Charles, F., Porteous, J., Gilroy, S., Klovatch, I., Jackont, G., Soreq, E., Keynan, N.J., Cohen, A., Raz, G., Hendler, T.: Towards empathic neurofeedback for interactive storytelling. In: Finlayson, M.A., Meister, J.C., Bruneau, E.G. (eds.) 2014 Workshop on Computational Models of Narrative of OpenAccess Series in Informatics (OASIcs), Dagstuhl, Germany, vol. 41, pp. 42–60. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik (2014)
8. Doi, H., Nishitani, S., Shinohara, K.: NIRS as a tool for assaying emotional function in the prefrontal cortex. Frontiers in Human Neuroscience **7**, Article 770 (2013)
9. Esfahani, E.T., Sundararajan, V.: Using brain-computer interfaces to detect human satisfaction in human-robot interaction. International Journal of Humanoid Robotics **8**(01), 87–101 (2011)
10. Field, A.: Discovering statistics using SPSS. Sage (2009)
11. Garcia-Molina, G., Tsoneva, T., Nijholt, A.: Emotional brain-computer interfaces. International Journal of Autonomous and Adaptive Communications Systems **6**(1), 9–25 (2013)
12. Gilroy, S.W., Porteous, J., Charles, F., Cavazza, M., Soreq, E., Raz, G., Ikar, L., Or-Borichov, A., Ben-Arie, U., Klovatch, I., et al.: A brain-computer interface to a plan-based narrative. In: International Joint Conference on Artificial Intelligence (IJCAI), pp. 1997–2005. AAAI Press (2013)

13. Harmon-Jones, E.: Clarifying the emotive functions of asymmetrical frontal cortical activity. Psychophysiology **40**(6), 838–848 (2003)
14. Karran, A.J., Fairclough, S.H., Gilleade, K.: A framework for psychophysiological classification within a cultural heritage context using interest. ACM Transactions on Computer-Human Interaction (TOCHI) **21**(6), Article 34 (2015)
15. Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., Barrett, L.F.: The brain basis of emotion: a meta-analytic review. Behavioral and Brain Sciences **35**(03), 121–143 (2012)
16. Morinaga, K., Akiyoshi, J., Matsushita, H., Ichioka, S., Tanaka, Y., Tsuru, J., Hanada, H.: Anticipatory anxiety-induced changes in human lateral prefrontal cortex activity. Biological Psychology **74**(1), 34–38 (2007)
17. Mühl, C., Heylen, D.: Cross-modal elicitation of affective experience. In: 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009, pp. 1–12. IEEE (2009)
18. Naseer, N., Hong, K.: fNIRS-based brain-computer interfaces: a review. Frontiers in Human Neuroscience **9**, Article 3 (2015)
19. Rosenfeld, J.P., Cha, G., Blair, T., Gotlib, I.H.: Operant (biofeedback) control of left-right frontal alpha power differences: Potential neurotherapy for affective disorders. Biofeedback and Self-Regulation **20**(3), 241–258 (1995)
20. Ruocco, A.C., Rodrigo, A.H., Lam, J., Di Domenico, S.I., Graves, B., Ayaz, H.: A problem-solving task specialized for functional neuroimaging: validation of the Scarborough adaptation of the Tower of London (S-TOL) using near-infrared spectroscopy. Frontiers in Human Neuroscience **8**, Article 185 (2014)
21. Sakatani, K., Takemoto, N., Tsujii, T., Yanagisawa, K., Tsunashima, H.: Nirs-based neurofeedback learning systems for controlling activity of the prefrontal cortex. In: Oxygen Transport to Tissue XXXV, pp. 449–454. Springer (2013)
22. Sarkheil, P., Zilverstand, A., Kilian-Hütten, N., Schneider, F., Goebel, R., Mathiak, K.: fMRI feedback enhances emotion regulation as evidenced by a reduced amygdala response. Behavioural Brain Research **281**, 326–332 (2015)
23. Sitaram, R., Caria, A., Birbaumer, N.: Hemodynamic brain-computer interfaces for communication and rehabilitation. Neural Networks **22**(9), 1320–1328 (2009)
24. Solovey, E.T., Girouard, A., Chauncey, K., Hirshfield, L.M., Sassaroli, A., Zheng, F., Fantini, S., Jacob, R.J.: Using fNIRS brain sensing in realistic HCI settings: experiments and guidelines. In: The 22nd Annual ACM Symposium on User Interface Software and Technology, pp. 157–166. ACM (2009)
25. Spielberg, J.M., Miller, G.A., Warren, S.L., Engels, A.S., Crocker, L.D., Banich, M.T., Sutton, B.P., Heller, W.: A brain network instantiating approach and avoidance motivation. Psychophysiology **49**(9), 1200–1214 (2012)
26. Tuscan, L.-A., Herbert, J.D., Forman, E.M., Juarascio, A.S., Izzetoglu, M., Schultheis, M.: Exploring frontal asymmetry using functional near-infrared spectroscopy: a preliminary study of the effects of social anxiety during interaction and performance tasks. Brain Imaging and Behavior **7**(2), 140–153 (2013)
27. Wehbe, R.R., Kappen, D.L., Rojas, D., Klauser, M., Kapralos, B., Nacke, L.E.: EEG-based assessment of video and in-game learning. In: CHI 2013 Extended Abstracts on Human Factors in Computing Systems, pp. 667–672. ACM (2013)
28. Zotev, V., Krueger, F., Phillips, R., Alvarez, R.P., Simmons, W.K., Bellgowan, P., Drevets, W.C., Bodurka, J.: Self-regulation of amygdala activation using real-time fmri neurofeedback. PloS One **6**(9), e24522 (2011)

A Neuroaesthetic Study of the Cerebral Perception and Appreciation of Paintings by Titian Using EEG and Eyetracker Measurements

Francesca Babiloni[1,2], Dario Rossi[1,2], Patrizia Cherubino[1,2], Arianna Trettel[2], Daniela Picconi[3], Anton Giulio Maglione[2], Giovanni Vecchiato[4], and Fabio Babiloni[1,2(✉)]

[1] Department of Molecular Medicine, University of Rome Sapienza, Rome, Italy
[2] BrainSigns Srl, Rome, Italy
fabio.babiloni@uniroma1.it
[3] Azienda Speciale Palaexpo, Rome, Italy
[4] Department of Physiology and Pharmacology, Sapienza University of Rome, Rome, Italy

Abstract. The neuroelectrical and the eye-movements activities were collected in a group of 27 healthy subjects during their visit of a fine arts gallery displaying twenty paintings of Titian. Evaluation of the appreciation of the paintings was performed by using the neuroelectrical approach-withdrawal index (AW). AW index was estimated for two groups of ten paintings of Titian: one group was related to religious matter while the second was related to portraits. In addition, it was compared the population AW indexes estimated in the first 10 seconds of the observation across all the selected paintings with the AW index estimated in the last 40 seconds of the observation. The number and the total duration (in seconds) spent on the eye fixations performed by the subjects during the observation of the paintings was also analyzed.

Results showed that the AW index was significantly higher during the observation of portraits than during the observation of the religious subjects ($p<0.007$). Interestingly, the grand average AW index estimated in the first 20 seconds of the observation of the paintings remains highly correlated with the AW index evaluated for the second part of the data (from 20 s to one minute) for all the 20 paintings examined ($r = 0.82$, $p<0.0001$). The number of eye fixations in the first 10 seconds of observation of the paintings that were most appreciated are significantly higher than the number of eye fixations for paintings that subjects did not like ($p<0.019$). Moreover, the total time spent on fixations for paintings that were liked by the subjects was significantly higher than the time spent on paintings that were not liked ($p<0.036$). Taken together, such results seem to suggest that the neuroelectrical correlates of the perception of “good” or “bad” paintings are generated in our brain within the first 10-20 seconds from the initial exposition of the subject to the painting.

Keywords: Neuroaestethic · EEG · Eyetracker · Titian

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 21–32, 2015.
DOI: 10.1007/978-3-319-24917-9_3

1 Introduction

The study of the cerebral perceptions related to fine art experiences has been started more than a decade ago by S. Zeki [1]. To now, the major part of the studies related to the neuroaesthetic discipline have been performed by using the hemodynamic correlates of brain activity through the functional Magnetic Resonance Imaging (fMRI), as reviewed in [2]. Only few studies have been performed by using the neuroelectrical correlates of brain activity during the fine arts perception [3–5]. Many of such studies put the subjects in front of a reproduction of the fine arts painting or sculpture on a computer screen, while only few are using real masterpieces as stimuli. Until few years ago, the possibility to collect cerebral activity from freely moving subjects was almost impossible from a technical point of view, due to the limitations on the acquisition hardware and the absence of efficient artifacts removing procedures. Nowadays, a body of methodologies able to collect and to analyze cerebral activity and eye-movements fixations during the execution of freely movements of subjects in open spaces has been developed [6–7]. Thus, it appears possible to gather cerebral activities even during the appreciation of real masterpieces in an art gallery.

In particular, we investigated the cerebral and eye movement activities correlated to the appreciation of paintings during a visit of a real art gallery in Rome. In fact, cerebral activity linked to the appreciation or rejection of the sensory inflow related to generic picture's observation has been index in literature by the unbalance of the EEG power spectra (EEG PSD) in the alpha band over the prefrontal areas [8–10]. In particular, it was demonstrated as a greater left prefrontal activity in the EEG PSD suggests a propensity to an engage with the sensory stimulus provided while a relatively greater right prefrontal activity suggests a modality of disengage from the stimulus proposed [8–9]. While the unbalance of the EEG PSD over the prefrontal cortices returns an index of the approach-withdrawal attitudes of the subject in front of the stimuli, the information about the modality of the painting exploration are not easily determined unless an eye-tracking device is used. By recording the eye-movements of the participants during the observation of the paintings, we were interested in the analysis of their scanning patterns during the observation of paintings they like when compared to the paintings they did not like.

2 Methodology

2.1 Experimental Design

The experiment has been performed at the "Scuderie del Quirinale", which is one of the major art galleries in Rome. The gallery hosted a collection of paintings from Titian (1488–1576). Twenty paintings were selected as stimuli for the subjects, ten related to portraits and ten related to religious matter. The twenty paintings selected for the analysis are presented in Fig. 1. During the experiment the gallery was closed to visitors. Ten of the selected paintings were related to a religious matter and ten were related to portraits of man or woman.

Twenty-seven healthy subjects (37.04±9 years, 14 males) were involved in the experiment. Informed consent was obtained from each subject after the explanation of the study, which was approved by the local institutional ethic committee. The experiment was conducted following the principles outlined in the Declaration of Helsinki of 1975, as revised in 2000. All the subjects underwent the same experimental procedure related to the acquisition of their EEG and eyetracking activities. In fact, before the visit each subject remained for one minute with the open eyes in a rest position.

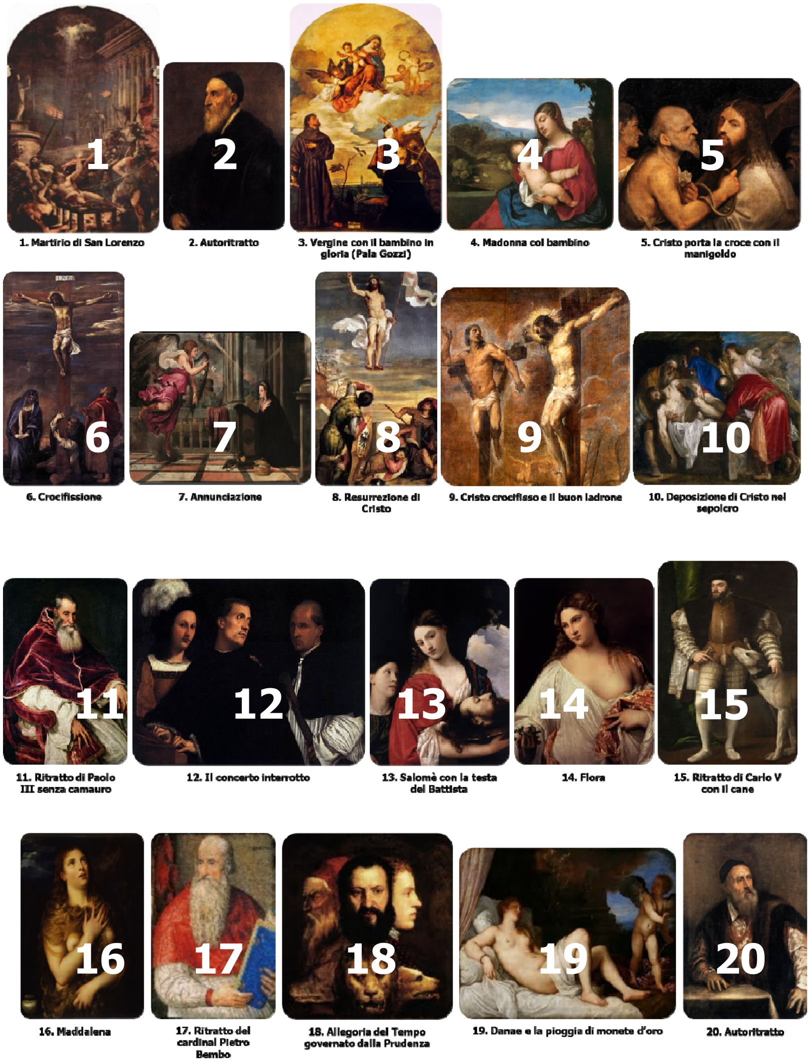


Fig. 1. The 20 masterpieces of Titian that have been presented to the subjects during the visit of the art gallery. The numbers on the paintings describe the visit sequence performed by all the investigated subjects.

Successively, all the subjects spent one minute observing a text explaining the content of the exposition on the wall of the gallery. This last sequence was taken as the EEG baseline for the successive analysis. For all the subjects the visit of the gallery consisted in the same sequence of painting's observation. Each one of the twenty paintings selected for the visit was observed for one minute by each subject.

This procedure was adopted in order to assure that all the subjects observed the selected paintings in the same temporal sequence. The naturalistic vision of the painting was performed in silence by the subjects, which were also asked to minimize their head and muscular movements in front to it. Fig. 2 shows the typical setup of the EEG and the eye-tracking recording device mounted on the subject.

Fig. 2. The typical setup for the recordings employing the international 10–20 system for the electrode montage and the ASL eyetracker device. The mobile EEG and eyetracker devices was worn by the subjects without any relevant physical efforts.

Figure 3 shows the free observation of the subject in front of a particular masterpiece involved in the experiment. The subject was free to adjust his/her optimal distance from the painting for the best possible perception.

After 1 minute of free vision, the experimenter asked the subject to rate the painting observed according to his/her perceived pleasantness (ranging from 1, ugly, to 10, beautiful) and then guided him/her to the next painting.

Fig. 3. The collection of the brain activity during the aesthetic observation of the painting "Annunciazione" by Titian. Note the EEG cap and the eye-tracking device mounted on the cap for the monitoring of the eye movements. EEG and eye tracking data were stored on the portable devices carried by the subjects in a little portable bag.

2.2 EEG Recordings and Signal Processing

The aim of this section was to describe the succession of signal acquisition and processing steps performed to estimate the approach/withdrawal (AW) index related to the appreciation or the rejection of each investigated painting of Titian in the analyzed population. In addition, it was also investigated the cerebral activity related to the appreciation or rejection of the paintings by each subject formed in the first 20 seconds of the naturalistic vision. To this purpose, the average value of the AW index estimated in the first 20 second of the vision of the painting was referred in the following as AW20. Such index was contrasted with the AW index averaged along the successive 40 seconds, called AW60 (from 21 seconds to the 60 seconds of the observation of the painting).

The following acquisition and analysis steps were then performed: 1) the collected EEG data were subjected to artifact removal by applying Independent Component Analysis (ICA) procedure supervised by an EEG expert. Usually up to two components related to the blink and eye movement artifacts were removed from the ICA space in each subject; 2) the gathered EEG data collected during the gallery visit were transformed in AW z-score values, by taking into account the cerebral activity collected during the baseline; 3) values of AW, AW20 and AW60 z-scores were computed for each subject and for each painting analyzed; 4) the grand average of the AW, AW20 and AW60 z-score values for each painting was then computed;

EEG was gathered by a portable 19-channel system (BEmicro, EBneuro, Italy). The international 10–20 system was used as guide for the electrode placement. The Fpz channel has been used as reference. Electrode impedances were below 5kΩ. Independent Component Analysis (ICA) was applied to the EEG to detect and remove components due to eye movements, blinks and muscular artefacts. The ICA procedure employed was supervised by an EEG expert and it was applied on every subject's data. Usually up to two components were removed from the ICA space since they are related to the blink and eye movement artifacts. The Individual Alpha Frequency (IAF) has been calculated for each subject in order to define the frequency bands of interest according to the method suggested in the scientific literature [11].

In particular, the Individual Alpha Frequency value was given by the frequency band (IAF-4, IAF+2) where the values are given in Hz. EEG traces were then segmented to extract and analyze the cerebral activity during the observation of the selected paintings. Each EEG trace has been band pass filtered in order to isolate the spectral components in the alpha band from the whole EEG spectrum. The filtered traces have been used to calculate the Global Field Power [GFP; 12]. We used the frontal electrodes to compute the GFP indexes in this study, by selecting the electrodes F7, F3, Fp1, Fz, Fp2, F4, F8 of the International 10–20 montage. Such selection was performed to evaluate the Approach/Withdrawal Index (AW index).

The formula that defines the AW index is the following:

$$AW = GFP\alpha\_right - GFP\alpha\_left \quad (1)$$

where the GFPα_right and GFPα_left stand for the GFP calculated among right (Fp2, F4, F8) and left (F7, F3, Fp1) electrodes, in the alpha band, respectively. The AW index was then normalized returning a z-score values across all the experiment for each subject. In fact, such index has been defined by taking into account the frontal EEG asymmetry's theory by Davidson and coworkers [9, 13]. After, the values of the z-score AW index across the time spent in the observation of each painting were averaged, returning a single value of AW for each subject and for each painting. These AW values were also grand-averaged across the subjects, to return an average AW value for each painting investigated.

The AW evaluation was also estimated by taking into account the EEG data from the first 20 seconds (AW20) as well as the data from the 20^{th} second to the end of observation for each paint (AW60).

2.3 Eye-Tracker Analysis

The eye-tracker device employed (ASL technologies, USA) returned information about the displacement of the eye gaze for each subjects during the observation of the twenty paintings analyzed. In particular, the number and the total duration in seconds of the eye fixations were taken as indexes of the scanning behavior of the subject during the one minute of free observation of the painting. A eye fixation was defined when the eye gaze of the subject remains stable in a square of 40x40 pixels of the target for a period of about 240–280 ms during a free scene viewing. Fixations were related to the focus of the attention during the scanning of a presented stimulus. Information on fixation pattern was also corroborated by the estimation of the total time in seconds spent on fixations by the subjects on the naturalist vision of the paintings. Fixations were collected for each subject and for all the paintings during the first 10 seconds and for the entire duration of the observation allowed (60 seconds).

2.4 Statistical Analysis

Behavioral. Each subject at the end of each observation of the painting returned a verbal score ranging from 1 to 10. An ANOVA with the main factors CONTENT (at two levels; Religious, Portrait) and PAINTING (at ten levels) was performed by using the explicit score as dependent variable.

EEG. The statistical analysis on EEG is aimed to understand if :

1) the AW index estimated in the acquired sample population differs between the paintings related to the portraits when compared to the those of religious content;
2) there is a difference in the appreciation generated by the subjects in the first 20 seconds of the observation of the painting when compared to the second part of such free observation. This comparison is addressed by the use of the AW20 and AW60 indexes and the correlation analysis across all the paintings considered.

An analysis of variance (ANOVA) has been performed on the AW values with the main factors CONTENT (with the two levels Portrait and Religious) and PAINTING (with ten levels) at the 5% significance level.

The grand averages of the AW20 and AW60 index for all the paintings evaluated were subjected to the correlation analysis (through the estimation of the Pearson coefficient). This was made to assess if the values of the appreciation or rejection of the paintings by the population that were formed in the first 20 seconds of the naturalistic vision were similar to those estimated in the last 40 seconds of the painting's view.

Eye-fixations. The statistical analysis performed on the eyetracker data was related to the hypothesis that the number of fixations and the time spent on such fixations could differ between paintings that the subjects like most from the paintings that the subjects did not like. To this aim, the number of eye fixations and the total time in seconds spent in fixations on each painting was evaluated for each subject. In each subject, the painting that received the best score and the painting that receive the lowest verbal score were selected. For these particular couple of paintings, the number of fixations along different time length of observation (10 and 60 seconds) and the total time spent on fixations were estimated. A paired Student's t-test was then performed separately on each one of these variables (e.g. the number of fixations at 10 and 60 seconds and the total time spent on fixations). Protection against the multiple execution of several univariate t-tests was adopted by using the False Discovery Rate procedure (FDR), at a 5% significance level.

3 Results

3.1 Behavioral Results

The ANOVA performed on the verbal scores for each painting observed returned a significant value for the factor PAINTING (F = 2.92; p<0.01) as well as for the interaction CONTENT x PAINTING (F = 9.42; p<0.00007). No significance was obtained for the main factor CONTENT (F = 0.66; p = 0.43).

3.2 Brain Activity Related to the Paintings Observation

The average AW index estimated in all the subjects for the 20 paintings selected is shown in Fig. 4. The AW index is expressed as a z-score value. When positive, it suggests an appreciation of the investigated group for the particular painting observed while vice-versa when negative. The average evaluation of the sample group analyzed differed greatly across the 20 paintings considered in the gallery. In fact, oscillations of the AW index between +2.5 and -1.5 were observed. The number on the abscissa of Fig. 4 represents the progressive number of the painting encountered by all the subjects during the fine art gallery visit, as described in Fig. 1.

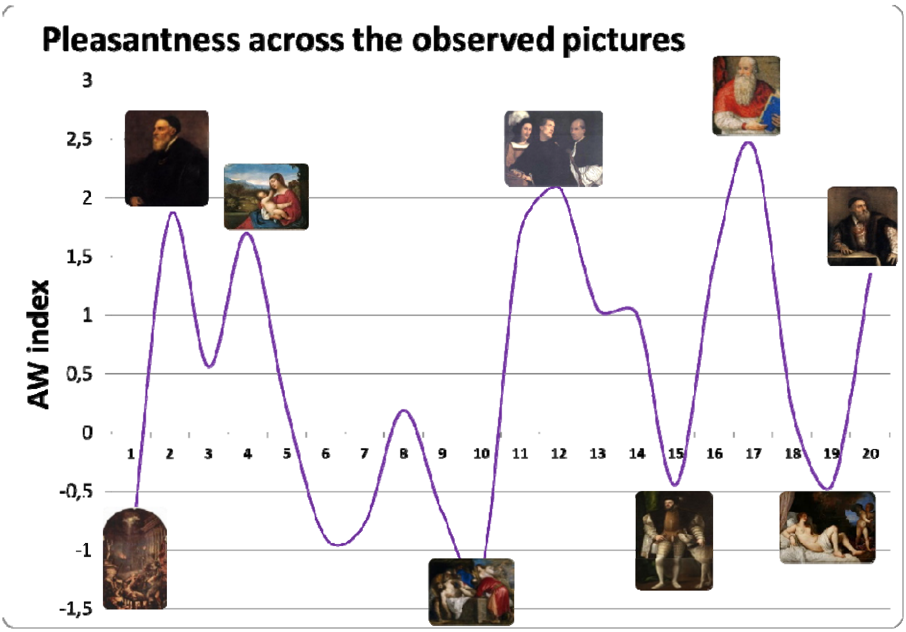


Fig. 4. Variation of the Approach/Withdrawal (AW) index along the gallery visit. The abscisse numbers represent the number of the painting encountered along the experiment as described in the Fig. 1. Ordinate values are related to the z-score computed for the unbalancing of the EEG power spectra over the prefrontal cortex. Maximum value of the AW index is reached for the observation of the painting number 17. The values showed are relative to the grand average of the z-scores values of AW obtained for every subjects involved in the analysis for each painting.

The ANOVA performed on AW index values returned a significant value for the factor CONTENT ($F = 9.95$; $p<0.005$) and for the interaction CONTENT x PAINTING ($F = 2.64$; $p<0.007$), being then the AW values of the Portrait level greater than the Religious level.

The correlation between AW20 and AW60 indexes was high ($r = 0.82$) and statistically significant ($p < 0.0001$).

3.3 Eye Fixations Related to the Paintings Observation

In the first 10 seconds, a higher number of fixations have been produced during the observation of the paintings that the subjects like most when compared to the fixations generated for the observation of paintings that the same subjects did not like. Such difference is statistically significant according to the paired Student's t-test, with a $p < 0.019$. Also, the total time spent in the fixations for the first 10 seconds of these observations is greater for the painting the subjects like most when compared to the paintings the subject did not like. This difference was statistically significant according to the paired Student's t-test, with a $p<0.036$. Interestingly, this statistical difference in the number of fixations between pleasant and less pleasant paintings vanished when the time period allowed for the observation of the painting reached one minute ($p = 0.54$). In this condition, there is no difference between the number of fixations received by the paintings that the subjects like or reject most.

4 Discussion

This paper provided evidences of specific scalp prefrontal activity correlated to the evaluation of a series of real paintings by Titian during a visit in a fine art gallery. Previously, general sensory or motor stimuli induced by the observation of paintings have been investigated [1–5]. In addition, it was also described the involvement of medial orbitofrontal cortex as assessed by hemodynamic measurements during the perception of artistic artifacts [2]. However, in the present study it must be stressed that the estimation of the AW index on the scalp prefrontal areas is not equivalent to the use of source current density methodologies, that could estimate the cortical activity from EEG measurements [14–21].

It might be argued that the estimation of the AW index could be affected by the occurrence of ocular artifacts, since it was clearly allowed to the subjects to move their eyes during the free observation of the paintings. However, it must be considered that the occurrence of the ocular artifacts have a bilateral spread toward the frontal and the central scalp areas. Thus, it is unlikely that such artefacts could play a role in the evaluation of the AW index, which is based on the unbalance between the EEG PSD over the left and right prefrontal scalp areas.

Results obtained suggest that the investigated population shows a higher AW scores for paintings in the Portrait section when compared to the AW scores for paintings listed in the Religious section. This finding is in agreement with the results obtained in a previous similar investigation performed on the paintings of Jan Vermeer [13].

In such investigation, paintings related to portraits received a higher AW score than those related to landscapes. A possible interpretation of such results is related to the importance of face recognition in humans when compared to the other elements of the visual scene represented in the painting.

A specific finding of this study is that the AW index estimated in the analyzed population during the first 20 seconds of the free observations of the paintings is highly correlated with the value of the AW estimated during the second part of the observations ($r=0.82$, $p<0.0001$). This result could be interpreted as that the scalp "prefrontal" aesthetic evaluation of the paintings by the subjects occurred within the first 20 seconds of the observation of the paintings and did not change in the successive 40 seconds.

It might be argued that EEG autocorrelation could be responsible for the re-occurrence of the same EEG patterns at the base of the similarity of the AW20 and AW60 indexes related to the free observation of the paintings. On the other hand, such hypothesis could be challenged at the light of the fact that the AW index is estimated with respect to a baseline. Thus, it will be unlikely that a particular unbalance of EEG activity over the prefrontal electrodes due to the EEG autocorrelation would be maintained above the chance level when compared to its occurrence during the baseline period. If an autocorrelation would be at the base of the phenomena observed, it would be present also in the baseline before the task was performed.

The evidences returned from the eye-tracking part of the study presented a statistically significant higher number of eye fixations for the paintings the subjects liked most when compared to the number of eye fixations for the painting the subjects did not like. This result held only during the first 10 seconds of observation of the paintings. In fact, the number of fixations along the entire duration of the task were similar between the paintings tested. In addition, also the total time spent on fixations were higher in the first 10 seconds of free observation for the paintings the subjects liked most than for the paintings the subjects did not like.

Thus, it could be hypothesized that a high number of fixations and an increased time spent on such fixations characterizes the scanning pattern for paintings that subjects liked most with respect to the case in which the subjects observed paintings they did not like.

It may be argued that in free viewing experiments fixations could be affected by head movements. This fact could be a source of a major confound effect in the analysis of the present eye tracking data. However, it must be stressed that it was asked to the subject to stay calm and avoiding head movement during the time spent on the observation of the painting.

Taken together, these results suggest that the "internal" appreciation or rejection of the paintings was a relatively quick process in the investigated sample. In fact, such process appears to be supported by a high eye fixation process in the first 10 seconds of the observation of the paintings, also sustained by the formation of a consistent unbalanced scalp prefrontal activity in the first 20 seconds of the free view of the painting.

Such flow of events could be tested in a successive experiments related on how and when we forming inside our head the judgment about the beauty perceived by a fine arts painting.

Acknowledgments. The present work was supported by BrainSigns srl. The grants provided to FB by the Italian Minister of University and Education under the PRIN 2012 and that provided by the Minister of Foreign Affairs for the bilateral relation between Italy and China for the project "Neuropredictor" are also gratefully acknowledged.

References

1. Zeki, S.: Art and the Brain. J. Conscious. Stud. **6**, 76–96 (1999)
2. Di Dio, C., Gallese, V.: Neuroaesthetics a review. Current Opinion in Neurobiology **19**, 682–687 (2009)
3. Cela-Conde, C.J., et al.: Sex-related similarities and differences in the neural correlates of beauty. Proc. Natl. Acad. Sci. **106**, 3847–3852 (2009)
4. Lengger, P.G., et al.: Functional neuroanatomy of the perception of modern art: A DC-EEG study on the influence of stylistic information on aesthetic experience. Brain Res. **1158**, 93–102 (2007)
5. Umilta, M.A., et al.: Abstract art and cortical motor activation: an EEG study. Front. Hum. Neurosci. **6**, 311 (2012)
6. Makeig, S., Gramann, K., Jung, T.P., Sejnowski, T.J., Poizner, H.: Linking brain, mind and behavior. International Journal of Psychophysiology **73**(2), 95–100 (2009)
7. Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A.M., Kliegl, R.: Coregistration of eye movements and EEG in natural reading: analyses and review. Journal of Experimental Psychology General **140**(4), 552 (2011)
8. Harmon-Jones, E., Gable, P.A., Peterson, C.K.: The role of asymmetric frontal cortical activity in emotion-related phenomena: a review and update. Biological Psychology **84**(3), 451–462 (2010)
9. Davidson, R.J.: Cerebral asymmetry, emotion, and affective style. In: Davidson, R.J., Hugdahl, K. (eds.) Brain Asymmetry, pp. 361–387. MIT Press, Cambridge (1995)
10. Demaree, H.A., et al.: Brain lateralization of emotional processing: Historical roots and a future incorporating "dominance". Behavioral and Cognitive Neuroscience Reviews **4**, 3–20 (2005)
11. Klimesch, W.: EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res. Rev. **29**(2), 169–195 (1999)
12. Lehmann, D., Skrandies, W.: Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr. Clin. Neurophysiol. **48**, 609–621 (1980)
13. Babiloni, F., Cherubino, P., Graziani, I., Trettel, A., Infarinato, F., Picconi, D., Borghini, G., Maglione, A.G., Mattia, D., Vecchiato, G.: Neuroelectric brain imaging during a real visit of a fine arts gallery: a neuroaesthetic study of XVII century dutch painters. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., pp. 6179–6182 (2013)
14. Werkle-Bergner, M., et al.: Cortical EEG correlates of successful memory encoding: Implications for lifespan comparisons. Neuroscience and Biobehavioral Reviews **30**, 839–854 (2006)

15. Davidson, R.J.: What does the prefrontal cortex "do" in affect: perspectives on frontal EEG asymmetry research. Biol. Psychol. **67**(1–2), 219–233 (2004)
16. Babiloni, F., Carducci, F., Cincotti, F., Del Gratta, C., Roberti, G.M., Romani, G.L., Rossini, P.M., Babiloni, C.: Integration of high resolution EEG and functional magnetic resonance in the study of human movement-related potentials. Methods Inf. Med. **39**(2), 179–182 (2000)
17. Mattia, D., Babiloni, F., Romigi, A., Cincotti, F., Bianchi, L., Sperli, F., Placidi, F., Bozzao, A., Giacomini, P., Floris, R., Marciani, M.G.: Quantitative EEG and dynamic susceptibility contrast MRI in Alzheimer's disease: a correlative study. Clinical neurophysiology **114**(7), 1210–1216 (2003)
18. Babiloni, C., Babiloni, F., Carducci, F., Cincotti, F., Rosciarelli, F., Rossini, P.M., Arendt-Nielsen, L., Chen, A.: Mapping of early and late human somatosensory evoked brain potentials to phasic galvanic painful stimulation. Human brain mapping **12**(3), 168–179 (2001)
19. Babiloni, C., Babiloni, F., Carducci, F., Cappa, S., Cincotti, F., Del Percio, C., Miniussi, C., Moretti, D.V., Pasqualetti, P., Rossi, S.: Human cortical EEG rhythms during long-term episodic memory task. A high-resolution EEG study of the HERA model. Neuroimage **21**(4), 1576–1584 (2004)
20. Pichiorri, F., De Vico Fallani, F., Cincotti, F., Babiloni, F., Molinari, M., Kleih, S.C., Neuper, C., Kübler, A., Mattia, D.: Sensorimotor rhythm-based brain–computer interface training: the impact on motor cortical responsiveness. Journal of neural engineering **8**(2), 025020 (2011)
21. Babiloni, C., Babiloni, F., Carducci, F., Cincotti, F., Rosciarelli, F., Arendt-Nielsen, L., Chen, A., Rossini, P.M.: Human brain oscillatory activity phase-locked to painful electrical stimulations: A multi-channel EEG study. Human brain mapping **15**(2), 112–123 (2002)

Symbiotic Adaptive Interfaces: A Case Study Using BrainX^3

Ryszard Cetnarski[1], Alberto Betella[1], Andrea Miotto[2], Riccardo Zucca[1], Xerxes D. Arsiwalla[1], Pedro Omedas[1], Jonathan Freeman[2], and Paul F.M.J. Verschure[1,3](✉)

[1] SPECS, N-RAS, DTIC, Universitat Pompeu Fabra, Barcelona, Spain
paul.verschure@upf.edu
http://specs.upf.edu
[2] Department of Psychology, Goldsmiths - University of London, London, UK
[3] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Abstract. Modern symbiotic and adaptive HCI paradigms include real-time algorithms capable of inferring user's states. We present an experimental interface which aims to understand the information processing capacity of a human user and use this information to improve interaction in a database exploration task. We collected the electrodermal activity and pupillometry signals in tasks of increasing difficulty and used their features to infer whether the subject was performing the task correctly or not. By combining principal component analysis and logistic regression methods, we successfully inferred the accuracy of users' responses from the signal after the response was made. This study provides a quantitative framework for modelling user internal states and evaluates it in a practical human computer interaction task.

Keywords: Human-computer-interaction · Adaptive interfaces · Affective computing · EDA · Pupillometry

1 Introduction

1.1 Using Physiological Signals to Infer Implicit User States

Symbiotic and adaptive interfaces open new channels of communication between the human and the machine [4,11]. The principal aim of such interfaces is to decode the internal states of the human user and use this information to improve the interaction [18,20]. So far, the traditional interfaces are only capable of reacting to explicit user input which puts a limit on the possible understanding such system can have of its user. Between humans, on the other hand, explicit communication constitutes only a small fraction of the actual message [14]. Being able to decode affective and cognitive states of another person (i.e. mentalizing) is what makes humans effective at communication and collaboration [9]. Such states are not declared explicitly in human-human interaction and need to be decoded from plenty of noisy sources. We aim at developing a symbiotic and

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 33–44, 2015.
DOI: 10.1007/978-3-319-24917-9_4

adaptive interface which is capable of inferring such implicit internal states in order to adjust the visualisation and interaction parameters of a neuroscience application, BrainX^3 [3]. BrainX^3 was developed as a part of the CEEDs project which aims at developing an immersive mixed reality framework to support the exploration of neuroscientific data [13]. The purpose of this interface is to infer the information processing capacity (i.e., cognitive load) of a human user and act accordingly to optimize the context in which the database is explored and analysed.

Among the many individual internal states, cognitive load reflects the information processing capacity of a person [17]. Other internal states, such as the affective state of arousal, report (although not primarily) the degree to which a person is capable of paying attention to a relevant task [8]. These two internal states affect the autonomic nervous system (ANS) responses [7], and can be inferred from the pupillometric and electrodermal signals [2,6]. Pupil size was shown to increase in cognitively demanding tasks (such as arithmetic or memory tasks) [10] while electrodermal activity (EDA) is widely used to measure the arousal levels in response to stimuli (for instance, startle responses to fearful stimuli) [5]. EDA is a more selective measure of internal states, reporting only the affective ones, whereas the pupil size changes in response to both cognitive and affective aspects of a task [19].

Using the pupillometry and EDA measurements we investigated the relationship between user's internal states and performance during the exploration of a neuroscience database. Characterizing such relationship allows to adapt the interface, for example reducing the display complexity in order to adjust to the user information processing capacity. Such task can be compared to what a tutor does when he notices that a student is not coping (or coping too well) with the given difficulty of the assigned task. Given such analogy we named our system the Sentient Agent, as it should monitor the internal states of the user, interpret them and adapt the task accordingly. Here we report the result of an experiment in which we tested two fundamental requirements of a symbiotic system. First, whether the physiological signals accurately describe the internal states of the participants and second, whether the signals relate to user behavior. The tonic and event related signal analysis will address these two questions respectively.

1.2 BrainX^3 Application

The context and setup of the present experiment is the BrainX^3 application (figure 1), a visualization and analysis tool aimed to assist researchers in the development of hypotheses on regularities and principles of brain organization and function [3]. BrainX^3 is supported by the XIM-engine, a control architecture that integrates and controls different sensors and effectors in the immersive mixed reality infrastructure called eXperience Induction Machine (XIM) [16]. The data set presented in BrainX^3 is the human connectome, a large network of nodes and edges representing the structural connectivity of the human brain, which can be analysed using different graph theoretical measures to describe and infer structural properties of the brain [1]. Connectome provides important information not only on the whole brain description level. Sometimes the researcher needs to know if particular anatomical connectivity exists, i.e. is node 'A' connected directly to node 'B' or are there some intermediaries? The anatomical map of nodes and edges can be revealing about the effective functional properties of different brain regions and visual inspection provides the first step towards more quantitative analysis the application provides. Given the nature of the tasks a researcher is likely to encounter, we designed the experiment to be visually identical to the human connectome display. Therefore, for the experimental setup we designed an artificial network where the user has to navigate from one end to another while understanding the structural organization of the data set. Such setup was chosen to create a controlled experiment in the context of big data exploration, using the BrainX^3 application.

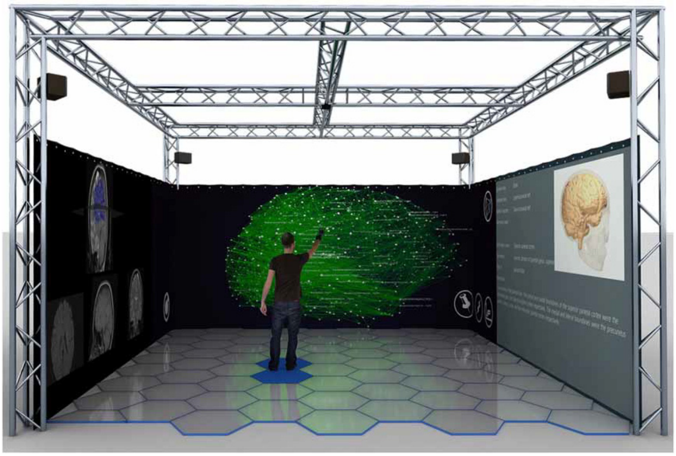


Fig. 1. Computer rendering of BrainX^3 within the eXperience Induction Machine (XIM). XIM is a cave system equipped with a range of physiological and tracking sensors. The BrainX^3 display, including the 3D connectome network are projected on the frontal screen. The user can navigate and interact with the model with hand gestures or a tablet interface. This application is used to investigate the relation between the structural organization of the brain and it's functional properties [1].

2 Materials and Methods

2.1 Sensors

The pupil signal was recorded with a wearable eye-tracker from Pupil Labs at a sampling rate of 30 Hz (Pupil Labs UG, Berlin, Germany). The EDA was sampled at 100 Hz using a custom made, wearable, wireless glove sensor. The signals were collected and time-adjusted with the experimental application using the SSI Framework [21]. The gestures and position of the user were captured by a Kinect2 (Microsoft, Redmond, WA, USA). The interaction in the mixed reality setup was controlled by the XIM-engine [16], whereas the visualization of the virtual reality (VR) environment was developed using the Unity engine (Unity Technologies, San Francisco, CA, USA).

2.2 Experimental Procedure

The experimental task required the participants to navigate from one side of a network (identified as the start node) to the opposite one (end node). A synthetic network was constructed from a repeating pattern of layers (figure 2), such that each node was connected with two edges to the nodes inside the same layer and with one edge to the next layer of nodes. The participant had to move from node to node, selecting the edge to move by next. Thus, at each node, participants could decide between three paths two take, one (correct) which was leading towards the other end of the network and two (incorrect) leading to the same current layer.

Participants were instructed to reach the end of the network selecting the shortest path, thus trying to minimize the number of incorrect choices. Subjects were informed about the structure of the network and how they should decide the path to take. After each step (i.e. decision) subjects were provided with feedback on form of a score on a colored bar and a distinct sound.

After a training period aimed to let the participant familiarize with the task, Gaussian noise was added to the 3D coordinates of all the nodes making the structure more difficult to perceive. Three levels of noise were added resulting in 3 difficulty levels of the task. Each participant completed all difficulty levels in a random order. Each level was composed of 12 layers and it took on average 14.48 minutes to complete them all.

3 Results

3.1 Participants Selection

51 participants recruited from the University campus participated to the study (24 females, mean age 22.1±4.2 SD). Participants were paid 10 euros after the experiment and signed an informed consent form. 5 participants were excluded from the analysis of the pupil signal and other 14 subjects for the EDA analysis. Subjects were excluded when more then 30% of the data was missing.

3.2 Signal Pre-processing

Both signals were first reconstructed through interpolation (using numpy.interp) of the corrupted segments with the surrounding signal, and then detrended applying a Hodrick-Prescott filter [15]. The EDA signal was not low pass filtered due to the many averaging steps necessary to obtain the final time-series. Comparing the results to the procedure including low-passing did not produce

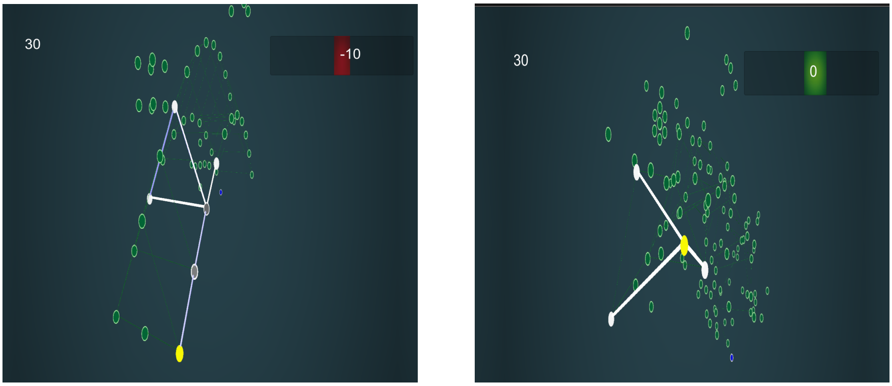


Fig. 2. (Left side) The structure of the artificial network. Each layer is triangular. Each node is connected to the two other nodes inside the same triangle (i.e. layer) and one outgoing connection to the next layer. The task of the participant is to move to the next triangle and avoid moving around the same one. By always selecting the path connecting to the next layer subject will move in the shortest path from one side of the network to the end. (Right side) The same experimental network modified with noise, making the underlying structure of the network more difficult to perceive. Gaussian noise was added to the 3D position of each node and the connections were maintained as in the figure on the left. The participant still has to follow the shortest path but in this case the decision can be difficult. The highlighted paths represent the options available to the subject, one of which is correct and two are incorrect.

any significant differences in the outcome. The aligned responses of each subject were finally normalized and averaged separately for correct and incorrect events. All analysis was done using python with numpy and scipy libraries.

3.3 Tonic Measure

To analyze the tonic properties [12] of the signals a time window starting one second before and ending 5 seconds after the decision time was removed from the signal. The remaining time series were averaged to produce the tonic measure for comparison between the easy, medium and hard task difficulty conditions. This method produced a repeated measure of tonic EDA and pupil signals. Such measure should reflect the overall internal state changes for each participant, which are not related to a particular decision, but to the overall difficulty of the task.

A repeated measures ANOVA with a Greenhouse-Geisser correction determined that mean tonic measure of the pupil signal did not statistically significantly differed between difficulty levels ($F_{(1.989,87.509)} = 2.226$, $p<0.114$) (figure 3). The mean value was 75.76 (10.91 SD) in the easy, 76.32 (11.37 SD) in the medium and 76.31 (11.07 SD) in the hard difficulty level.

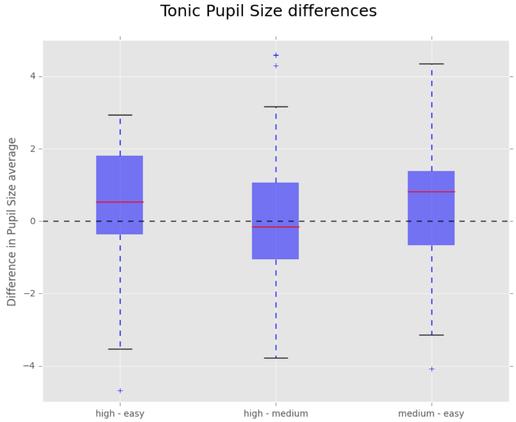


Fig. 3. The tonic activity of the pupil signal during different task difficulties. The box plots represent within subject differences between easy, medium and high difficulties. The horizontal line is marking zero level, which is a point of equality between the compared tonic activities.

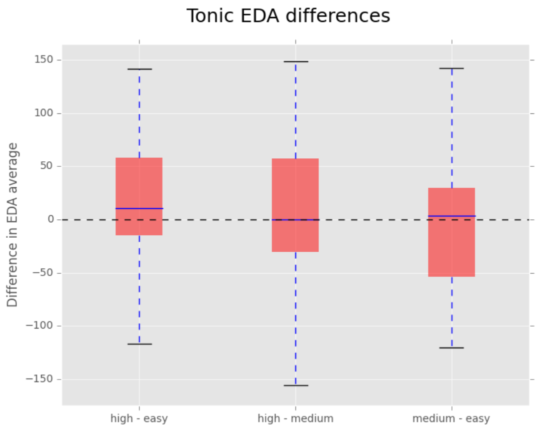


Fig. 4. The tonic EDA measure during different task difficulties. The box plots represent within subject differences between easy, medium and high difficulties. The horizontal line is marking zero level, which is a point of equality between the compared tonic activities.

There was no statistically significant difference in the EDA tonic measure between the difficulty levels of the task, $\tilde{\chi}^2(2) = .438$, $p = 0.804$) (figure 4). Median scores for the easy, medium and hard difficulties were 760 (617 to 951), 742 (631 to 988) and 746 (640 to 1025), respectively.

3.4 Event-Related Measure

For the analysis of the event related responses (i.e. phasic responses) [12] a 5 seconds signal slice was selected centered on the decision time. The aligned signal slices of each subject were then normalized and averaged separating the correct and incorrect decisions. The event related measure was taken by comparing averaged correct and incorrect signal slices for each participant. The standard deviation thus represents the between-subject variability.

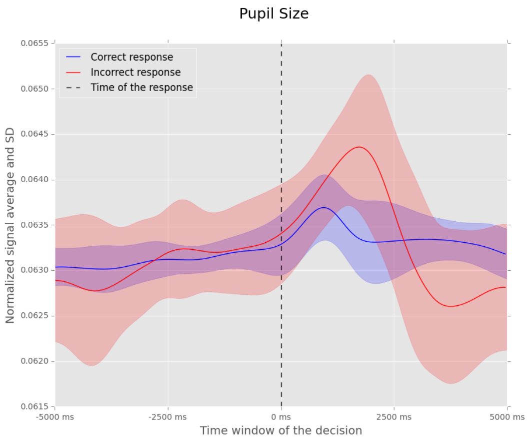


Fig. 5. Time course of the pupil size around the time of the decision. Blue and red lines represent the mean signal value and shaded regions represent the standard deviation range. The peak in the signal after the decision is the subject reaction to the feedback.

The time course of the pupil size (figure 5) show three peaks which can be used to distinguish correct from incorrect decisions. First peak is the negative deviation (downwards) of the incorrect time course and it occurs around 4 seconds before the decision. The second peak is the positive deviation (upwards) about 2 seconds after the decision. The last, third peak occurs about 4 seconds after the decision. The difference between the correct and incorrect time courses is quantitatively compared in the following section (figure 6).

For both the pupil and EDA time courses our aim was to use the signal in order to infer whether the user made a correct or incorrect decision. In order to compare the whole time series we used a dimensionality reduction method, the principal component analysis (PCA), which can be used to find the axis with the largest variance in the transformed signal. We can then use the value of each signal projected into this axis (i.e., the principal component) as the predictor variable in the logistic regression. Logistic regression is useful for predicting a binary outcome, in our case a correct or incorrect decision from a set of continuous variables provided by the PCA transformation.

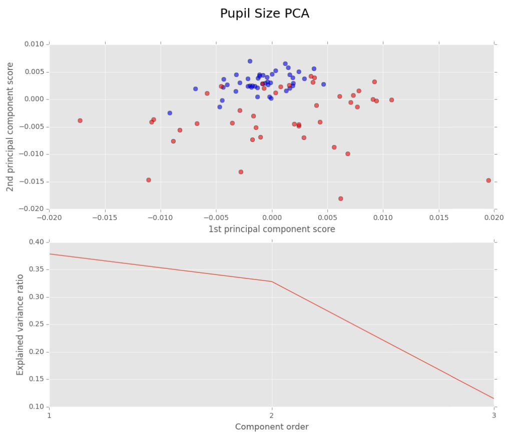


Fig. 6. Principal component analysis of the event-related pupil signal slices. The top panel represents the signals reduced to the first two principal components. The blue points represent correct signal slices and red points represent the incorrect ones. The bottom panel represents the proportion of variance covered by the component.

A logistic regression analysis was conducted to predict the accuracy of the subject decision using PCA components of the pupil signal as predictors (figure 6). Testing the full model against a constant only model was statistically significant, indicating that at least some of the predictors reliably distinguished between correct and incorrect decisions ($\tilde{\chi}^2$= 64.990, p<.001 with df = 3). Nagelkerke's R2 of .696 indicated a moderately strong relationship between prediction and observed outcome. Prediction success overall was 81.8% (86.7% for correct and 76.7% for incorrect decisions). The Wald criterion demonstrated that both the first and second principal component made a significant contribution to prediction (p<0.05 for both components). EXP(B) value indicates that when the second component is raised by one unit the odds ratio is 2 times as large and therefore the subject is twice more likely to make a correct decision. On the other hand, increasing the first component by one unit decreased the odds that the subject will make a correct decision by 30%.

The time course of the EDA (figure 7) shows much higher variance in case of incorrect decisions (red line) than correct decisions (blue line). The average time course varies mostly about 3 seconds after the decision. In contrast to the pupil signal the EDA shows no statistically significant deviation between correct and incorrect time courses prior to the decision (see above).

The logistic regression model using the PCA transformed EDA to predict the accuracy of the subject decision was statistically significant ($\tilde{\chi}^2$ = 8.938, p<.05 with df = 3) (figure 8). The model explained 17.6% (Nagelkerke's R2) of the variance in the decision accuracy and correctly classified 61.9% of the cases (78.1% for correct and 45.2% for incorrect decisions). The Wald criterion demonstrated that the first principal component made a significant contribution

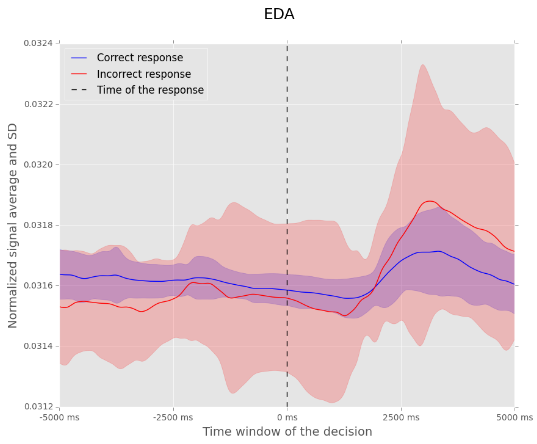


Fig. 7. Time course of the EDA signal around the time of the decision. Blue and red lines represent the mean signal value and shaded regions represent the standard deviation range. The peak in the signal after the decision is the subject reaction to the feedback.

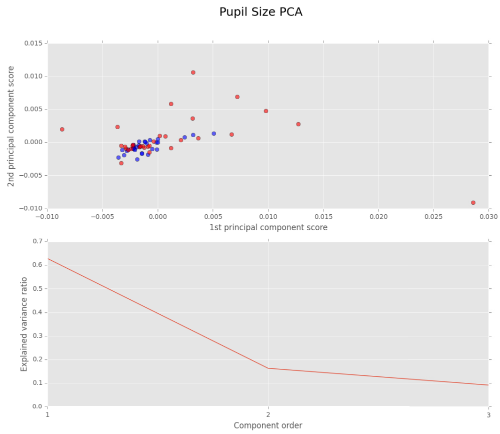


Fig. 8. Principal component analysis of the event-related EDA signal slices. The top panel represents the signals reduced to the first two principal components. The blue points represent correct signal slices and red points represent the incorrect ones. The bottom panel represents the proportion of variance covered by the component.

to prediction (p=0.05). EXP(B) value indicates that when the second component is raised by one unit the odds ratio is 0.8 times smaller and therefore the subject is 20% less likely to make a correct decision. Increasing the second component by one unit also decreased the odds that the subject will make a correct decision by 25% but was on the border of not making a significant contribution to the model (Wald test p=0.07).

We repeated the same procedure using only the time course of pupil signal 2.5 seconds prior to the decision. The model was statistically significant ($\tilde{\chi}^2 = 7.764$, p = .05 with df = 3) and explained 11.3% (Nagelkerke's R2) of the variance in the decision accuracy and correctly classified 62.5% of the cases (73.4% for correct and 51.2% for incorrect decisions).

4 Discussion

Symbiotic interfaces require decoding of the human internal states which are relevant for the goals in the HCI tasks. We have shown, how using psychophysiological measures of pupil and EDA signals such decoding is possible. From the collected signals we could classify different user responses in a task relevant for neuroscientific database exploration. Looking at the pupil time-series (figure 5) we identified three peaks in the negative average. First occurs about 4 seconds before the decision, the second about 2 seconds after and the last one about 4 seconds after the decision. These peaks are the parts of the signal where the largest difference between correct and incorrect decision can be observed and are represented in the first three principal components of the signal, used to classify the signal in the logistic regression model.

The strongest difference between correct and incorrect averages is visible after the decision was made, which is not the information of primary importance to a truly adaptive system. This visible difference is most likely attributed to the difference in reaction the subjects had to the positive (sound and earning points) versus negative feedback (distinct sound and loosing points). To recognize such response the adaptive system has to know a priori what decisions are correct and incorrect, like in the presented experiment, thus making the human effectively redundant for the correct completion of the tasks. However, applying the same analysis (PCA and logistic regression) on the signal slice before the decision yielded promising results.

Being able to predict whether a human is about to make a correct or incorrect decision prior to it actually being made is a necessary condition for a truly adaptive system. Being capable to do so means that the system does not have to know a priori what are the correct decisions, but can rely on a user's model which is predictive of the user's performance or behavior. The model was statistically significant (p = .05) and explained 11.3% of the variance in the decision accuracy and correctly classified 62.5% of the cases (73.4% for correct and 51.2% for incorrect decisions). However, given the low explanatory power of the model for classyfing the incorrect decisions it could not yet be employed for a practical adaptive interface.

The tonic measure of pupil size and of EDA could not distinguish between the conditions on statistically significant grounds. Although the test of the tonic pupil measure was on the border of statistical null effect (p =0.116) we could see that the trend was not linear. The medium task difficulty resulted in a higher tonic activity than both easy and hard difficulties (figure 3), the high - medium difference is the only one with median smaller than 0. We would expect the pupil size to linearly increase with the task difficulty, but this was not the case. In figure 4 we reported that the tonic EDA is little modulated by the task difficulty. The affective reactions which the EDA measure is sensitive too might not occur in the cognitive tasks with little affective components. Altogether these results suggest that physiological signals are not viable to describe the general state of the user, i.e. in the absence of any specific stimulus in the context of the BrainX^3 application.

Interestingly, the pupil signal was a better predictor for correct and incorrect choices. This finding is rather surprising since we attributed the post-decision peak in the signal as subject response to the feedback. Differences between negative and positive feedback relies more in the affective than cognitive domain and it should be more visible in the EDA signal, which was not the case. This result makes our explanation for the observed peak uncertain.

In conclusion, it was proven difficult to predict the accuracy of the response using the physiological signals before the user's choice, but possible to do so afterwards. As a follow up, it will be interesting to investigate what happens without providing any feedback to the user.

Overall the sensing architecture of the XIM engine is capable of decoding user states, however the challenge of symbiotic adaptive interfaces is to develop real time algorithms capable of classifying user states in the same way it is possible in offline analysis.

Overall these results suggest that the more specific the interaction is defined the greater the predictive power of the psychophysiological signals. We ground this conclusion in finding observable differences in the event related properties of the signals but not in the tonic ones. Therefore the path for the symbiotic interfaces seems tied to the development of concrete and practical applications, rather then a general solution for many HCI tasks. BrainX^3 is an application that provides the context for developing a symbiotic interface aiding neuroscience research but has yet to be tested on a real neuroscientific task.

References

1. Arsiwalla, X.D., Zucca, R., Betella, A., Martinez, E., Dalmazzo, D., Omedas, P., Deco, G., Verschure, P.F.: Network dynamics with Brain X3: a large-scale simulation of the human brain network with real-time interaction. Frontiers in Neuroinformatics **9** (2015)
2. Beatty, J., Lucero-Wagoner, B.: The pupillary system. Handbook of Psychophysiology **2**, 142–162 (2000)
3. Betella, A., Cetnarski, R., Zucca, R., Arsiwalla, X.D., Martínez, E., Omedas, P., Mura, A., Verschure, P.F.: Brain X3: embodied exploration of neural data. In: Proceedings of the 2014 Virtual Reality International Conference, p. 37. ACM (2014)
4. Betella, A., Martínez, E., Zucca, R., Arsiwalla, X.D., Omedas, P., Wierenga, S., Mura, A., Wagner, J., Lingenfelser, F., André, E., et al.: Advanced interfaces to stem the data deluge in mixed reality: placing human (un) consciousness in the loop. In: ACM SIGGRAPH 2013 Posters, p. 68. ACM (2013)

5. Boucsein, W.: Electrodermal activity. Springer Science & Business Media (2012)
6. Clark, L., Crooks, B., Clarke, R., Aitken, M.R., Dunn, B.D.: Physiological responses to near-miss outcomes and personal control during simulated gambling. Journal of Gambling Studies **28**(1), 123–137 (2012)
7. Ekman, P., Levenson, R.W., Friesen, W.V.: Autonomic nervous system activity distinguishes among emotions. Science **221**(4616), 1208–1210 (1983)
8. Fichtenholtz, H.M., Dean, H.L., Dillon, D.G., Yamasaki, H., McCarthy, G., LaBar, K.S.: Emotion-attention network interactions during a visual oddball task. Cognitive Brain Research **20**(1), 67–80 (2004)
9. Frith, C.D., Frith, U.: The neural basis of mentalizing. Neuron **50**(4), 531–534 (2006)
10. Granholm, E., Asarnow, R.F., Sarkin, A.J., Dykes, K.L.: Pupillary responses index cognitive resource limitations. Psychophysiology **33**(4), 457–461 (1996)
11. Jacucci, G., Spagnolli, A., Freeman, J., Gamberini, L.: Symbiotic interaction: a critical definition and comparison to other human-computer paradigms. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 3–20. Springer, Heidelberg (2014)
12. Lanatà, A., Valenza, G., Scilingo, E.P.: A novel EDA glove based on textile-integrated electrodes for affective computing. Medical & Biological Engineering & Computing **50**(11), 1163–1172 (2012)
13. Lessiter, J., Miotto, A., Freeman, J., Verschure, P., Bernardet, U.: CEEDs: unleashing the power of the subconscious. Procedia Computer Science **7**, 214–215 (2011)
14. McDonald, S., Flanagan, S.: Social perception deficits after traumatic brain injury: interaction between emotion recognition, mentalizing ability, and social communication. Neuropsychology **18**(3), 572 (2004)
15. McKinney, W., Perktold, J., Seabold, S.: Time Series Analysis in Python with statsmodels. Jarrodmillman.Com, pp. 96–102, July 2011
16. Omedas, P., Betella, A., Zucca, R., Arsiwalla, X.D., Pacheco, D., Wagner, J., Lingenfelser, F., Andre, E., Mazzei, D., Lanatà, A., et al.: Xim-engine: a software framework to support the development of interactive applications that uses conscious and unconscious reactions in immersive mixed reality. In: Proceedings of the 2014 Virtual Reality International Conference, p. 26. ACM (2014)
17. Paas, F., Tuovinen, J.E., Tabbers, H., Van Gerven, P.W.: Cognitive load measurement as a means to advance cognitive load theory. Educational Psychologist **38**(1), 63–71 (2003)
18. Pantic, M., Pentland, A., Nijholt, A., Huang, T.S.: Human computing and machine understanding of human behavior: a survey. In: Huang, T.S., Nijholt, A., Pantic, M., Pentland, A. (eds.) ICMI/IJCAI Workshops 2007. LNCS (LNAI), vol. 4451, pp. 47–71. Springer, Heidelberg (2007)
19. Setz, C., Arnrich, B., Schumm, J., La Marca, R., Troster, G., Ehlert, U.: Discriminating stress from cognitive load using a wearable EDA device. IEEE Transactions on Information Technology in Biomedicine **14**(2), 410–417 (2010)
20. Wagner, J., Lingenfelser, F., André, E., Mazzei, D., Tognetti, A., Lanatà, A., De Rossi, D., Betella, A., Zucca, R., Omedas, P., et al.: A sensing architecture for empathetic data systems. In: Proceedings of the 4th Augmented Human International Conference, pp. 96–99. ACM (2013)
21. Wagner, J., Lingenfelser, F., Baur, T., Damian, I., Kistler, F., André, E.: The social signal interpretation (ssi) framework: multimodal signal processing and recognition in real-time. In: Proceedings of the 21st ACM International Conference on Multimedia, pp. 831–834. ACM (2013)

On the Use of Cognitive Neurometric Indexes in Aeronautic and Air Traffic Management Environments

Gianluca Di Flumeri[1,2], Gianluca Borghini[3], Pietro Aricò[3], Alfredo Colosimo[1], Simone Pozzi[4], Stefano Bonelli[4], Alessia Golfetti[4], Wanzeng Kong[5], and Fabio Babiloni[2,6](✉)

[1] Departtment of Anatomy, Histology, Forensic Medicine and Orthopedics, Sapienza University of Rome, Rome, Italy
[2] BrainSigns srl, Rome, Italy
fabio.babiloni@uniroma1.it
[3] Departtment of Physiology and Pharmacology, Sapienza University of Rome, Rome, Italy
[4] Deep Blue srl, Rome, Italy
[5] Department of Computer Science, Hangzhou Dianzi University, Hangzhou, China
[6] Department of Molecular Medicine, University of Rome Sapienza, Rome, Italy

Abstract. In this paper the use of neurophysiological indexes for an objective evaluation of mental workload, during an ecological Air Traffic Management (ATM) task, has been proposed.

Six professional Air Traffic Controllers from the Italian ENAV (Società Nazionale per l'Assistenza al Volo) have been involved in this study. They had to perform an ecological Air Traffic Management task by using the eDEP software, a validated simulation platform developed by EUROCONTROL. In order to simulate a realistic situation, the task was developed with a continuously varying difficulty level, i.e. starting form an easy level, then increasing up to a harder one and finishing with an easy one again. During the whole task for each subject the electroencephalographic (EEG) signals were recorded in order to compute the neurophysiological workload index, and at the same time the subjective perception of the mental workload by using the Instantaneous Self-Assessment (ISA) technique. Thus, the EEG-based workload index, estimated by means of machine learning approach, by one side, and the subjective assessed workload index by the other side, have been compared in terms of correlation and difficulty levels discrimination. By the results it emerged: i) a high positive and significant correlation between the two measures, and ii) a significantly discriminability of the task different difficulty levels by using the EEG-based workload indexes, according to the ISA results.

In conclusion, this study validated the EEG-based mental workload index as an efficient objective evaluation method of the cognitive resources demand in a real operative scenario, and moreover as an index able to monitor its variations.

Keywords: EEG · EOG · Machine learning · Mental workload · Self-assessment · ATM · ATCO · eDEP

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 45–56, 2015.
DOI: 10.1007/978-3-319-24917-9_5

1 Introduction

In specific working environments where safety is paramount big issue, the human factor could be the risk reason less controllable and, at the same time, the main cause of danger. This is often because of an underestimation of the actual mental workload of the operator. In fact, as cognitive workload increases, maintaining task performance within an acceptable range becomes harder. High cognitive workload may demand more cognitive resources than those available in the human brain, resulting into performance degradation and errors commission [1]. The use of objective measures of mental workload based on biomarkers has been proposed for the evaluation of different systems design to allocate the workload, to minimize errors due to overloads or to intervene on the systems in real-time before the operators performing critical tasks become overloaded [2]. For example, few studies investigated neurophysiological indexes about the user states in safety-critical applications, such as driving, industrial environments or security surveillance. With respect to driving assistance applications, recent studies have explored the use of psychophysiological measures in a driving simulation for assessing driving performance and inattentiveness, as well as for robust detection of user intention before the braking onset [3–8].

In this regard, another example of operative environment where lack of performance or overloads may be fatal is the aviation context. Nowadays, the 80% of airplane incidents is still due to human - factors and, as the air - traffic keeps growing exponentially, the impact of new tools able to assess the interaction human – machine, in terms of cognitive resources, is becoming very important. In fact, there are evidences that the failure to perceive correctly the mental demands of a flight task, has been a causative factor in several aircraft accidents. This is true also for other operators critically involved in the air traffic managing (i.e. Air Traffic Control Officers, ATCOs). Both pilots and ATCOs categories of workers have to generate a continuous high quality performance with potential catastrophic results in occasion of error occurrence.

Focusing on the ATCOs, they have to perform a variety of tasks, including monitoring air traffic, anticipating loss of separation between aircraft, and intervening to resolve conflicts and minimize disruption to flow (for an extensive compilation of the tasks and goals of *en-route* control, see [9]). The ATCO's behavior could be measured through several human factor tools, such as the explicit measurement of errors performed during the task, or by using questionnaires related to the perception of the severity of the task executed and so forth, such as for instance the NASA-TLX or the SWAT questionnaires. Each of these methods has pros and cons, but there is not a standard one generally accepted [10], therefore the need of an objective measure becomes more important. Moreover, for their inherently subjective nature, none allows to have an objective and reliable measure of the actual cognitive demand in a real environment. Instead of only measuring secondary physiological effects, the EEG methods will offer a direct insight into the operator's state in complement to the common physiological measurements, as discussed above. There are many evidences that have underlined the correlation between the increase of the cognitive effort and the decision making in a strategy selection process during a complex task and the increase of the Electroencephalogram (EEG) *Power Spectral Density* (PSD) in the theta frequency band [4–7 Hz] over the frontal and occipital brain areas. In addition, it was also noted a corresponding decrease of the EEG PSD in the alpha frequency band [8–12 Hz] over the centro-parietal and parietal brain areas [3–6].

In a previous work [11, 12], it has been defined an algorithm able to evaluate the mental workload of novice ATCOs by using neurophysiological signals, during the execution of ATM task under different difficulty levels. Each difficulty level has been maintained constant for several minutes in order to keep the experimentation as controlled as possible. The results showed that the neurophysiological measure was able to evaluate the mental workload of the operator for each difficulty level.

On the basis of the previous results, the aim of this work was to test the reliability of the algorithm also during more ecological settings, where the difficulty of the task changes continuously. In this way, we have tested if the algorithm was able to track the fluctuation of the operators' mental workload within the operative task. In order to validate the results, the neurophysiological measure has been compared with the subjective measure of the mental workload, collected by the Instantaneous Self-Assessment (ISA) technique.

2 Methodology

2.1 Experimental Protocol and Task

Six professional ATCOs (49 ± 3.2 years) from ENAV S.p.A. (Società Nazionale per l'Assistenza al Volo, Italy) have been involved in this experimentation, in particular they have been asked to manage air-traffic under two different difficulty levels (EASY and HARD), using the ATM simulator eDEP (*Early Demonstration & Evaluation Platform*).

The eDEP software has been developed by EUROCONTROL, with the aim to produce a low-cost-lightweight, web-enabled ATM simulator platform, offering an ideal environment for research and advanced concept projects to rapid prototype applications [13]. A specific experimental protocol has been defined with the aim to highlight the investigated cognitive phenomena, that is the mental workload experienced by the subjects during the execution of the task.

In Fig. 1, a picture of the experimental setting during the task. The air-traffic task lasted about 37 minutes, during which the task difficulty varied between the two levels (EASY and HARD).

Since the eDEP software simulates a real scenario, the difficulty during the whole task varied continuously, thus, there were not constant difficulty conditions, but a

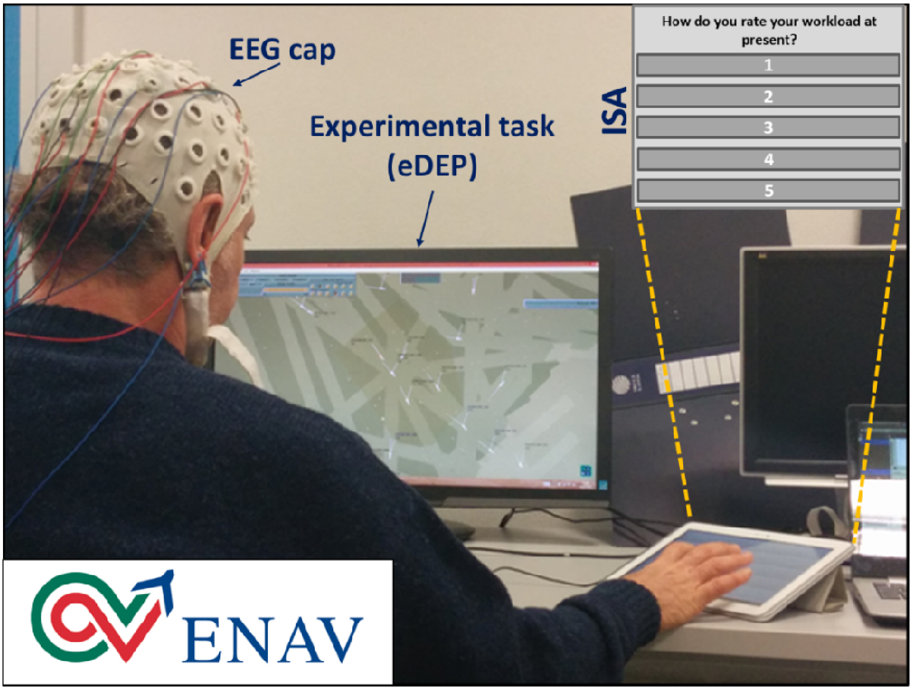


Fig. 1. Picture of the experimental setting during the task: on the right the tablet on which the ISA is presented; on the right at the top a zoom of the ISA interface.

difficulty profile designed as a "reverse – U". In other words, in the first 16 minutes ATCOs had to manage an EASY air traffic condition, in the following 16 minutes an HARD air traffic load and, in the last 5 minutes again an EASY air traffic condition (Fig. 2).

2.2 Instantaneous Self-Assessment (ISA)

Simultaneously to the execution of eDEP, ATCOs have been asked to fill the *Instantaneous Self-Assessment* (ISA).

In particular, the ISA [14] is a technique that has been developed to provide immediate subjective ratings of workload demands, from 1 (very easy) to 5 (very difficult), during the performance of a primary task, in our case, an air traffic management (ATM) task.

The ISA scale has been presented to the ATCOs every 3 minutes in the form of a colour-coded keypad on a tablet screen (10 inchs) positioned just below the main monitor (Fig 1). The keypad flashed and sounded when the workload rating was required, and the participants simply pushed the button that corresponds to their workload perception.

The ISA technique allowed a profile of the operator's workload throughout the eDEP task. The appeal of the ISA technique lies in its low resource usage and its low intrusiveness, that is, it ensured no interference with the main task.

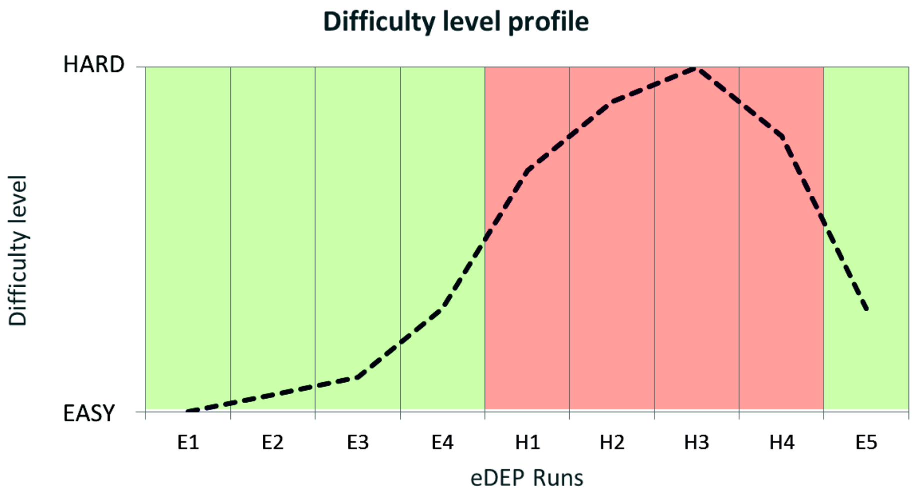


Fig. 2. Profile of the difficulty level, varying during the experimental task on the eDEP platform.

2.3 Neurophysiological Signals Acquisition and Analysis

Neurophysiological signals have been recorded by the digital monitoring *BEmicro* system (EBNeuro system). The 13 EEG channels (FPz, F3, Fz, F4, AF3, AF4, P3, Pz, P4, POz, O1, Oz and O2) and the EOG channel have been collected simultaneously with a sampling frequency of 256 (Hz). All the EEG electrodes have been referenced to both the earlobes, and the impedances of the electrodes have been kept below 10 (kΩ). The bipolar electrodes for the EOG have been positioned vertically above and below the left eye. The acquired EEG signals have been digitally band-pass filtered by a 4th order Butterworth filter (low-pass filter cut-off frequency: 30 (Hz), high-pass filter cut-off frequency: 1 (Hz)). The EOG signal has been used to remove eyes-blink artifacts from the EEG data by using the *Gratton* method [15]. For other sources of artifacts, specific procedures of the *EEGLAB* toolbox, based on threshold methods have been used [16].

After the artifact rejection, the EEG signals have been segmented in epochs of 2 seconds, 0.125 (ms) shifted. The PSD has then been estimated, for each epoch and for each EEG channel, by using the *Fast Fourier Transform* (FFT) in the EEG frequency bands, defined for each subject by the estimation of the *Individual Alpha Frequency* (IAF) value [17], correlated with the mental workload variations, therefore the theta [IAF-6 ÷ IAF-2] (Hz) and alpha [IAF-2 ÷ IAF+2] (Hz) bands. Furthermore, the PSD has been calculated using a Hanning window of the same length of the considered epoch (2 seconds length, is that 0.5 (Hz) of frequency resolution. Thus, with this frequency resolution, and considering the investigated frequency range equal to [IAF-6 ÷ IAF+2] (Hz), there was 17 PSD values for each channel.

2.4 EEG-Based Workload Index

A *Stepwise Linear Discriminant Analysis* (SWLDA, [11–12]) has been used to select the most relevant spectral features, within a features domain consisted of 221 values (13 ch * 17 PSD values), to discriminate the mental workload of the subjects within the different experimental conditions (EASY and HARD). In particular, the performed SWLDA used $\alpha_{ENTER} = .05$ and $\alpha_{REMOVE} = .1$, as probabilistic criterion for including and excluding features of the SWLDA itself. Once identified such spectral features, the SWLDA assigns to each one specific weights ($w_{i\ train}$), plus a bias (b_{train}), such that the SWLDA discriminant function ($y_{train}(t)$) takes the value *1* in the hardest condition and *0* in the easiest one. This step represents the *training phase* of the classifier. Later on, the weights and the bias determined during the training phase have been used to calculate the linear discriminant function ($y_{test}(t)$) over the testing dataset (*testing phase*). Finally, a moving average of 8 seconds (8MA) has been applied to the $y_{test}(t)$ function in order to smooth it out by reducing the variance of the measures, and we defined it as *EEG-based workload index* (W_{EEG}).

Here below are reported the training SWLDA discriminant function (1, where $f_{i\ train}(t)$ represents the PSD matrix of the training dataset at the time sample t, and of the i^{th} feature), the testing one (2, where $f_{i\ test}(t)$ is as $f_{i\ train}(t)$ but related to the testing dataset) and the equation of the *EEG-based workload index*, W_{EEG} (3).

$$y_{train}(t) = \sum_i w_{i\ train} \cdot f_{i\ train}(t) + b_{train} \tag{1}$$

$$y_{test}(t) = \sum_i w_{i\ train} \cdot f_{i\ test}(t) + b_{train} \tag{2}$$

$$W_{EEG} = 8MA\ (y_{test}(t)) \tag{3}$$

In order to have a more accurate resolution in terms of task difficulty variation, the dataset related to each subject has been segmented in 9 parts of 4 minutes each, so that we gathered 4 EASY runs (E1, E2, E3, E4), 4 HARD (H1, H2, H3 and H4) runs and another EASY run (E5). At this point, for each subject we have used the algorithm described above to train the classifier with one couple of EASY and HARD runs and to test it over the remaining eDEP difficulty conditions of the same subject. In order to appropriately choose this couple of conditions, we considered that i) the Controllers could need few minutes (i.e. E1 and E2) to become confident with the eDEP interface, so that we have used the *E3* as easy condition in the training dataset; ii) for the hard condition we have taken into account that the eDEP scenario's profile has been designed as a "reverse – U", so that we expected that the hardest condition would be in the middle. Therefore, the *H3* run has been chosen as hard condition in the training dataset.

2.5 Performed Analyses

Two kinds of statistical analysis have been performed in this study. In the first one, we estimated the Pearson's correlation coefficient between the ISA scores and the W_{EEG} measures for each run (i.e. E1, E2, E4, H1, H2, H4, E5). The E3 and H3 runs have not been considered in the analysis because they have been used for training the classifier.

In the second analysis, we tested both if the ISA and the W_{EEG} scores were statistically different over the two difficulty levels (EASY and HARD). In this way, we averaged both the EASY and HARD runs, and we performed two one-tailed student's t-tests ($\alpha = .05$) between the two classes, one for the ISA scores and one for the W_{EEG} indexes.

Before every statistical analysis, we used a *z-score* [18] correction of data for normalize the different behaviors of the subjects. In particular, we calculated this score by using the mean and the standard deviations of the related values (ISA, W_{EEG} scores) over the different runs (i.e. E1-E5, H1-H4).

3 Results

In Fig. 3, the ISA scores have been reported for each of the experimental tasks (i.e. E1-E4, H1-H4, E5). It could be appreciated as the individual perceptions of the tasks difficulties as proposed to the ATCos are in line with the intended simulations. In fact, the harder conditions (H2-H4) were perceived as more engaging cerebrally when compared to the easy ones (E1-E5). The proposed H1 task was perceived as not difficult from the ATCos involved with respect to the experimenter's judgment.

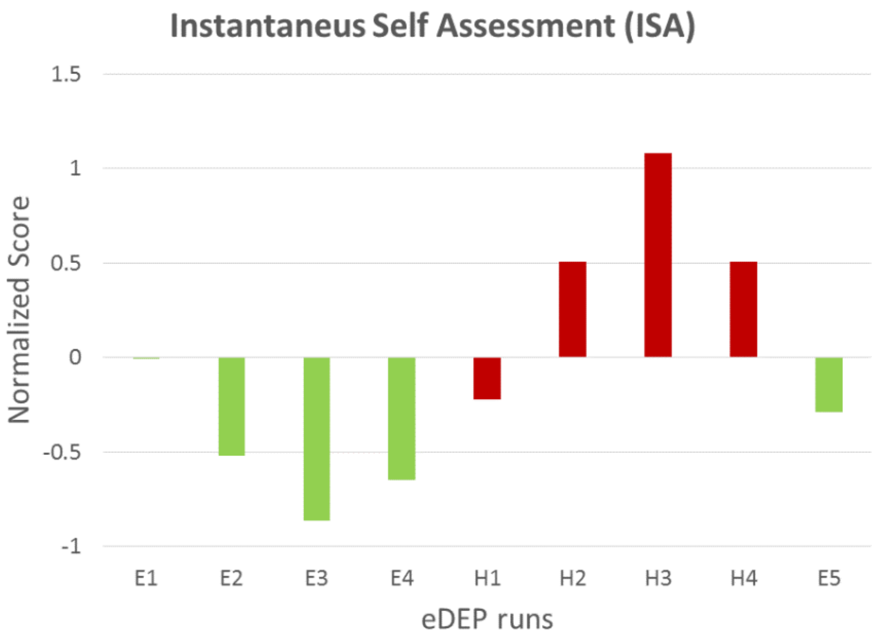


Fig. 3. ISA normalized score in the different eDEP difficulty conditions (E: EASY, H:HARD) of the ATCOs from ENAV.

Furthermore, in the Fig. 4 we showed the same representation for the EEG workload score, in which the E3 and the H3 tasks have not been reported, as stated before, since they have been used for training the classifier employed. It is possible to note that the red columns associated with the estimation of the EEG-based workload score are higher when compared to the green colums, associated to the estimation of the same score for the easier conditions. Note as the easy condition after the occurrence of a difficult sequence (E5, that arrives after the hard condition H4) was estimated of low engage from the personal judgments of the ATCOs while it was estimated still engaging above the average by using the EEG based workload index. This could suggest a major sensitivity of the EEG index to the cumulative effect of mental fatigue when compared to the personal judgment that at this moment appears to be based on external characteristic of the air traffic proposed.

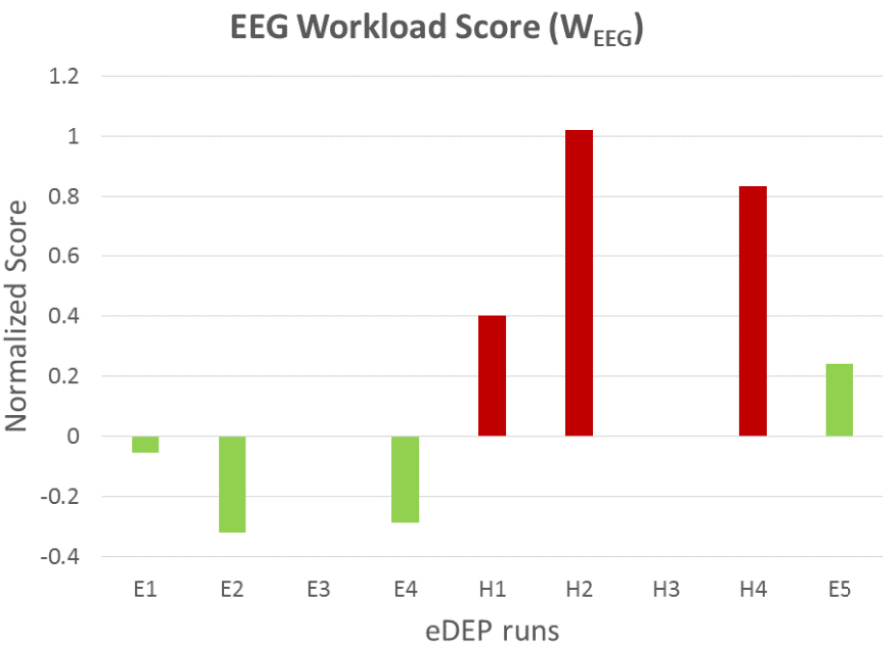


Fig. 4. Workload normalized score in the different eDEP difficulty conditions (E: EASY, H: HARD) estimated by the EEG signals of the ATCOs from ENAV.

It appears that there is a quite strong correlation with the detection of the difficult working states for EEG and ISA based indexes. In order to quantify such correlation, for the runs in which both the ISA and the W_{EEG} scores (i.e. E1 – E2 – E4 – H1 – H2 – H4 – E5) have been available, we performed a Pearson's correlation analysis. Results showed a high positive significant correlation between the two indexes trends ($R = 0.9$; $p = 0.006$), as highlighted in Fig. 5.

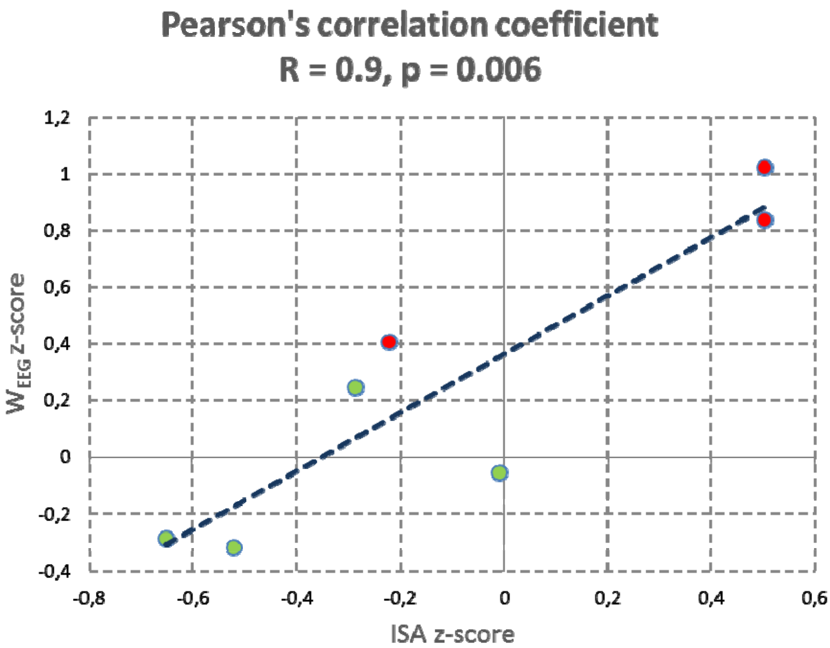


Fig. 5. Scatter plot between the normalized ISA score and the normalized W_{EEG} workload indexes. The tendency line shows the positive correlation between the brain workload perceived by the subject (ISA) and the neurophysiological measure of it (W_{EEG}).

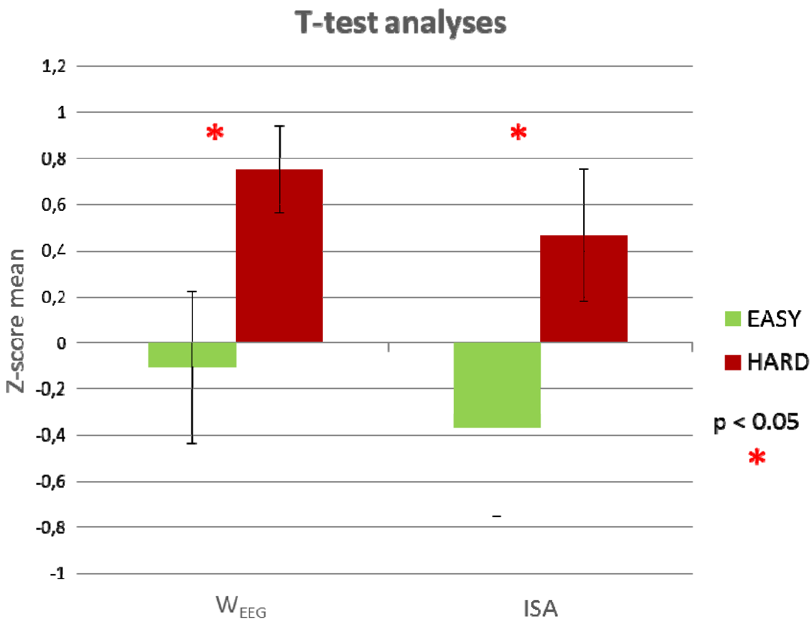


Fig. 6. The difficulty levels (EASY and HARD) of the task are significantly discriminated by both the indexes, the subjective (ISA) and the neurophysiological one (W_{EEG}).

The t-tests showed that the ISA scores related to the EASY and HARD conditions were significantly different ($p < .05$). Consistently, the overall EEG workload index calculated over the HARD conditions was significantly higher than the index calculated over the EASY tasks ($p < .05$). Figure 6 shows the results of the application of such t-tests on the different experimental conditions analyzed. The red columns are associated to the values of the analyzed indexes related to the hard working conditions while the green columns are associated to the values of the indexes related to the easy working conditions. It could be appreciated as both the use of EEG workload index as well as ISA are able to significantly distinguish the easy and the hard working conditions.

4 Discussion

Professional ATCOs have been involved in this study, where a neurophysiological workload measure (W_{EEG}) has been tested while the ATM operators performed an ecological air traffic control task. The ATCos have not been trained to use the eDEP platform before the experiments and, even if eDEP is a professional ATM simulator, during the first parts of the task (EASY1 and EASY2) they needed information and instructions to learn how to use correctly its interface. This aspect has been confirmed both by the ISA (Fig. 3) and by the EEG workload index (Fig. 4). In fact, during the E1 and E2 runs subjects showed higher workload perception (ISA) and physiological increment (W_{EEG}) of the workload than during the next EASY runs. Furthermore, as the eDEP scenario's profile has been designed as a "reverse – U", both the ISA score and the mental workload (W_{EEG}) index showed the same shape, confirmed by a high and significant correlation index ($R = 0.9$; $p = 0.006$).

In conclusion, both the workload perception (ISA) and the neurophysiological (W_{EEG}) measures showed a significant discriminability ($p<.05$) between the difficulty levels (EASY and HARD).

Our previous studies [5, 6, 11, 12, 19] showed the possibility to track the mental workload of the user even online, during simulated tasks in laboratory settings. In those studies, the difficulty of the task has been maintained constant for each experimental condition.

The results of the actual study confirmed that the neurophysiological workload measure can be used as a reliable index of the mental workload experienced by an operator also in ecological working scenario, where the difficulty of the task has not a discrete, but a continuous, profile. With the aim of confirming these results, further experiments will be performed over a bigger experimental sample size of ATCos, and probably with a greater resolution in terms of difficulty levels, in order to ensure that the estimated index is actually related to the experienced brain workload.

5 Conclusions

An algorithm able to track the mental workload of the user by using its brain activity, while performing an ecological operative task has been proposed in this study.

Results showed that the neurophysiological workload index (W_{EEG}) i) showed a high significant correlation with the perceived workload (ISA) and ii) was able to discriminate significantly two different difficulty levels, according to the ATCOs self-assessment.

We can then conclude that neurophysiological measures could provide objective evaluation of cognitive phenomena, e.g. the mental workload, both in real-time (on-line) and in ecological environments. In fact, questionnaires or rating scales might not fit real operative settings, where the operators (e.g. ATCOs) have to be focused exclusively on the task and they could not pay attention to secondary task(s), with the aim to provide data about their cognitive state, probably increasing the final task demand and operating in dangerous condition (under or over-load zone).

Acknowledgments. The present work was supported by the European Commission by FP7 project "ALICIA" and co-financed by Euro Control acting on behalf of the SESAR joint undertaking as part of work package E in the SESAR program and NINA project. The grants provided to FB by the Italian Minister of University and Education under the PRIN 2012 and that provided by the Minister of Foreign Affairs for the bilateral relation between Italy and China for the project "Neuropredictor" are gratefully acknowledged.

References

1. Norman, D.A., Bobrow, D.G.: On the analysis of performance operating characteristics. Psychological Review **83**(6), 508–510 (1976)
2. Byrne, E.A., Parasuraman, R.: Psychophysiology and adaptive automation. Biol Psychol. **42**(3), 249–268 (1996)
3. Welke, S., Jurgensohn, T., Roetting, M.: Single-trial detection of cognitive processes for increasing traffic safety. In: Proceedings of the 21st (Esv) International Technical Conference on the Enhanced Safety of Vehicles, Held June 2009, Stuttgart, Germany (2009). http://trid.trb.org/view.aspx?id=1100165. Accessed November 10, 2011
4. Venthur, B., Blankertz, B., Gugler, M.F., Curio, G.: Novel applications of BCI technology: psychophysiological optimization of working conditions in industry. In: 2010 IEEE International Conference on Systems Man and Cybernetics (SMC), pp. 417–421. IEEE (2010)
5. Borghini, G., Arico, P., Astolfi, L., Toppi, J., Cincotti, F., Mattia, D., et al.: Frontal EEG theta changes assess the training improvements of novices in flight simulation tasks. In: Conf Proc IEEE Eng Med Biol Soc. 2013, pp. 6619–6622 (2013)
6. Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., Babiloni, F.: Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neurosci Biobehav Rev., October 30, 2012
7. Haufe, S., Kim, J., Kim, I., Sonnleitner, A., Schrauf, M., Curio, G., Blankertz, B.: Electrophysiology-based detection of emergency braking intention in real-world driving. J Neural Eng **11**(5), 056011 (2014)
8. Haufe, S., Treder, M.S., Gugler, M.F., Sagebaum, M., Curio, G., Blankertz, B.: EEG potentials predict upcoming emergency brakings during simulated driving. J Neural Eng **8**, 056001 (2011)
9. Rodgers, M.D., Drechsler, G.K.: Conversion of the CTA, Inc., en route operations concept database into a formal sentence outline job task taxonomy. Washington, DC: Federal Aviation Administration Office of Aviation Medicine (1993)

10. Rubio, S., Díaz, E., Martín, J., Puente, J.M.: Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. Applied Psychology. **53**(1), 61–86 (2004)
11. Aricò, P., Borghini, G., Graziani, I., Imbert, J.-P., Granger, G., Benhacene, R., et al.: Air-Traffic-Controllers (ATCO) - Neurophysiological analysis of training and workload ATCO. IJASM **12**, 18–35 (2015)
12. Arico, P., Borghini, G., Graziani, I., Taya, F., Sun, Y., Bezerianos, A., et al.: Towards a multimodal bioelectrical framework for the online mental workload evaluation. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3001–3004 (2014)
13. Owen, S., Stainton, C.: eDEP Technical Document – CWP Performance report. 2006 Jan. Report No.: Reference GL/eDEP/EEC/TRSC60-2004/WP6.4
14. Kirwan, B., Scaife, R., Kennedy, R.: Investigating complexity factors in UK air traffic management. Human Factors and Aerospace Safety 1(2) (2001). http://trid.trb.org/view.aspx?id=717648. Accessed April 28, 2014
15. Gratton, G., Coles, M.G., Donchin, E.: A new method for off-line removal of ocular artifact. Electroencephalogr Clin Neurophysiol. **55**(4), 468–484 (1983)
16. Delorme, A., Makeig, S.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. **134**(1), 9–21 (2004)
17. Klimesch, W.: EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res Brain Res Rev. **29**(2–3), 169–195 (1999)
18. Zhang, J.-H., Chung, T.D.Y., Oldenburg, K.R.: A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. J Biomol Screen. **4**(2), 67–73 (1999)
19. Borghini, G., Aricò, P., Graziani, I., Salinari, S., Sun, Y., Taya, F., et al.: Quantitative Assessment of the Training Improvement in a Motor-Cognitive Task by Using EEG, ECG and EOG Signals. Brain Topogr. January 22, 2015

A Closed-Loop Perspective on Symbiotic Human-Computer Interaction

Stephen Fairclough(✉)

Liverpool John Moores University, Liverpool, UK
s.fairclough@ljmu.ac.uk

Abstract. This paper is concerned with how people interact with an emergent form of technology that is capable of both monitoring and affecting the psychology and behaviour of the user. The current relationship between people and computer is characterised as asymmetrical and static. The closed-loop dynamic of physiological computing systems is used as an example of a symmetrical and symbiotic HCI, where the central nervous system of the user and an adaptive software controller are engaged in constant dialogue. This emergent technology offers several benefits such as: intelligent adaptation, a capacity to learn and an ability to personalise software to the individual. This paper argues that such benefits can only be obtained at the cost of a strategic reconfiguration of the relationship between people and technology - specifically users must cede a degree of control over their interaction with technology in order to create an interaction that is active, dynamic and capable of responding in a stochastic fashion. The capacity of the system to successfully translate human goals and values into adaptive responses that are appropriate and effective at the interface represents a particular challenge. It is concluded that technology can develop lifelike qualities (e.g. complexity, sentience, freedom) through sustained and symbiotic interaction with human beings. However, there are a number of risks associated with this strategy as interaction with this category of technology can subvert skills, self-knowledge and the autonomy of human user.

Keywords: Symbiosis · Physiological computing · Intelligent adaptation

1 Introduction

The last three decades have seen huge innovation with respect to how we interact with computers. Communication via command lines was succeeded by WIMP interfaces and natural modes of communication via gestures and speeches are currently common features of desktop technology. Brain-computer interfaces represent the next frontier in human-computer interaction (HCI), where the neurological foundation of perception and action are utilised directly as a form of input control. Despite advances with respect to the available forms of input control, the basic communication dynamic of the human-computer dyad remains curiously fixed - the human 'speaks' and the computer 'listens and obeys.' Technology inhabits the passive role of slave-system that responds rigidly to a steady stream of directives from a human master, who directs actions towards a desired goal.

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 57–67, 2015.
DOI: 10.1007/978-3-319-24917-9_6

The distinction between the active role of the user and the passive function of the machine is starkly defined by the rigid turn-taking structure of contemporary HCI. This flow of information between person and machine has been depicted as two monologues rather than a genuine dialogue [1]. The way in which people interact with technology has also been described as asymmetrical with respect to the flow of information [2]. In other words, the person is free to interrogate the operational state of the computer (e.g. memory usage, Wi-Fi speed etc.) whereas the latter remains essentially blind to the psychological status of its user. By contrast, when technologies communicate with one another, information exchange can be symmetrical because each entity may freely probe and cross-examine all operational aspects of the other. The asymmetry that characterises interaction between humans and computers is distinguished by the absence of awareness on the part of the machine, which relegates a technological agent to the role of a passive and inert participant. In the absence of any ability to perceive or interpret the inner world of the user, the computer has minimal capacity for inference, anticipation, learning or any other quality that would liberate technology from its role as a slave-system.

The evolution of symmetrical forms of HCI are key to the creation of 'smart' technologies, which possess autonomy and intelligent adaptation [1]. This development should be considered within a general context of symbiosis between people and technology. Symbiosis may be described simply as two unlike organisms "living together" [3] in a relationship that may be mutualistic (i.e. both parties benefit), commensalistic (i.e. one benefits but the other is neither harmed or helped), or parasitic (i.e. one benefits with harm inflicted on the other).

If we define technology in the broadest sense, from the humble pencil to a nuclear power station [4], there are obvious benefits of technological forms for humanity as a species. Technology extends and augments our human limitations, a shovel allows the person to dig more effectively and efficiently, the motor car offers greater speed of transportation than travelling by foot [5]. Binoculars, telescopes and microscopes extend the range of visual perception and create a flexible, orthotic range [6] for human senses that greatly exceeds our "natural" limitations. The emergence of mobile devices combined with Internet connectivity and enhanced data storage augment our finite cognitive capabilities with respect to the storage and retrieval of information [7]. All these enhancements are achieved by "redistributing" task or information-processing demands between the human being and technological aids. It has been argued that the human brain has two important qualities that forge and fortify reliance on technology [8]. The brain is opportunistic in that it seeks to invent technological tools wherever there is potential for a significant improvement of efficiency and effectiveness. The brain is also a malleable organ, capable of co-opting technological tools seamlessly into existing behaviour and representations of self - and then creating a second and even third layers of tools to further bolster our human efficiency and effectiveness [5].

The relationship between symbiotic species may be described as obligate or facultative [9]. The former describes a state of co-dependence where each entity depends entirely upon the other for its continued survival. A facultative relationship represents those instances where two species can but not obliged to live together in order to survive. Whilst humans are currently the primary creators of technology, it would be a mistake to regard our relationship with technology as anything but an obligate form of mutualism. Individuals may attempt to (unsuccessfully) relinquish technological tools (see [5] Ch. 10), but technology is so entwined with human existence that any attempt to live without technological aids would force the human recipient to endure the kind of harsh living conditions that characterised feudal life 800 years ago [6]. It is also doubtful whether humans would be even capable of eradicating technology from our world if one considers the logistic barriers to that ill-advised endeavour [5]. Hence, we find ourselves in the contradictory position of being both master and slave to technology [5]. Rather than bemoaning our collective dependency on gadgets and computers, perhaps the most realistic course of action is to embrace this obligate relationship to further exploit human symbiosis with machines, as we have already been doing for several centuries. In the words of Hancock [6]: "Our ecology is technology. If we are to achieve our individual and collective goals, it will be through technology" (p. 66).

Our relationship with technology as a species is constructed upon an obligate form of symbiosis where humans rely on machines to extend our senses and capabilities - and technologies depend on human need and ingenuity in order to provide them with form and function. Despite this inter-dependence, the way in which we interact with machines remains asymmetrical with autonomy within HCI residing purely with the human user. This paper will outline the potential of physiological computing to both facilitate symmetrical forms of HCI and enhance our symbiotic relationship with technological systems. If technology can develop in this direction, the relationship between users and machines evolves towards a close, collaborative interaction that has profound implications for future technologies and its human users.

2 A Closed-Loop Perspective on Human-Machine Symbiosis

Human-machine symbiosis can describe the relationship between machine and person that occurs within a shared space or task [10]. A recent review defined human-machine symbiosis in terms of a computer that was capable of both monitoring and affecting the cognitions, emotions and behaviours of the user [11]. This description is identical to the closed-loop logic of physiological computing systems [12, 13] where signals from the brain and body of the user are converted to control inputs in order to facilitate intelligent adaptation at the interface. Physiological computing systems are constructed around a biocybernetic loop [14] where data from brain activity and the autonomic nervous system are collected, analysed and classified for input into an adaptive controller, which triggers actions at the interface.

2.1 Monitoring the User

Data from the brain and body are particularly appropriate for monitoring the psychological state of the user; in addition, these data have the advantages of being: quantifiable, continuously available, sensitive to unconscious activity and implicit, i.e. no overt response is required from the user [15]. In the case of physiological computing, the dynamic state of the user is inferred on the basis of spontaneous activity from the brain and the body [13, 16]. Analyses of these data yield a digital and quantified representation of the user state, which is made constantly available to the system. It is important to note that this representation of the user state is achieved via analogy as opposed to a literal re-representation of embodied experience [17]. The first step towards human-computer symbiosis is a simplification and quantification of embodied human experience into sparse information patterns that are digestible and reconcilable with a closed-loop mechanism of control and communication [18]. This act of abstraction is necessary in order to integrate the dynamic psychological state of the user within a cybernetic control loop.

There is a peculiar duality to this digital representation of self that acts as a point of origin within the biocybernetic loop. Whilst data from the brain and body are not a literal representation of the self or experience, they are derived from activity within the central nervous system and evoke both a degree of identification and biophilia [19], i.e. a preference for living systems. On the other hand, this quantified representation of self simultaneously evokes a technophilic proclivity for tools and technologies [5] and a reflexive perspective on self, i.e. the person becomes "an observing system observing itself observing" [17] (p. 144). By endowing a symbiotic computing system with the capacity to both monitor and represent the user, the loop creates a contradictory entity that (from a human perspective) is both self and other - the data are representative of the self but viewed from the objective perspective of another. It is important that users are fully informed in this respect. In other words, the measures upon which the quantification of state ought to be clearly defined and the user deserves a degree of education about the sensitivity and fallibility of this process. The user should understand that the process of measurement is neither perfectly sensitive nor absolutely representative due to the inherent limitations of measuring brain and body outside of the laboratory. This is important because users should not harbour unrealistic expectations about the fidelity of this representation or degree of personal insight that may be obtained via interaction with a biocybernetic system.

The capacity to monitor the user is the first challenge for symmetrical HCI, the next question is how the closed-loop mechanism should work with that user representation in order to create intelligent adaptation at the interface.

2.2 The Machine with an Agenda

The adaptive controller is the core element within the biocybernetic loop. This component receives information about the state of the user and translates these data into a range of appropriate responses at the interface. The adaptive controller encompasses a set of rules to describe how target state *a* is linked to an adaptive response *x* at the interface; for fuller technical description, see [16].

Aside from its technical substance, the adaptive component represents the means by which the system exerts a specific influence on the state or behaviour of the user. A number of biocybernetic loops have been created to serve different application domains, from mental workload classification [20], affective computing [21] and entertainment [22] to attention training [23]. In each case, the closed-loop model requires a target state to be defined and adaptations at the interface are designed to either induce/sustain a 'desirable' target state or reduce/ameliorate any target state deemed to be "undesirable."

For mental workload monitoring, the loop is designed to sustain a moderate level of mental workload and to avoid instances of high workload in order to preserve performance and safety. An affective computing system may be designed to detect a negative emotional state, such as frustration, and to trigger adaptive responses at the interfaces designed to reduce this emotion. An adaptive computer game would adjust gaming parameters in real-time to avoid the player becoming bored or disengaged. The definition of a psychological state to be achieved or avoided is common theme to all closed loop systems, and is especially relevant to symbiotic systems.

The closed loop system is governed by goal-directed logic. Unlike the inert and passive technology of today, this symmetrical interaction is characterised by a degree of agency on the part of the machine and a requirement for the human to cede a degree of control to the system. A user can decide whether or not to engage with the technology, but once the interaction has been initiated, the system can respond in a stochastic (as opposed to a deterministic) fashion. This is a small but significant shift in the relationship between people and computers.

Given that symmetrical HCI requires the human to relinquish a degree of control over the interaction, it is important to define the agenda of the machine to be effective, reliable and not lead to unforeseen circumstances. The introduction of agency or intentionality on the part of a machine shifts attention from the 'how' to the 'why' of technology because "the quintessential bottom line is that technology must be used to enfranchise not to enslave." [6] (p. 60). A closed loop system with intentionality must be used to materialise human goals and human values [24].

The formulation of human values within the closed-loop system remains a significant challenge. Illich [24] forwarded the case for convivial tools as technologies that create an opportunity for users to enhance and enrich the contribution of autonomous individuals. But how to recast this vague notion of conviviality within the precise semantics that are required by an adaptive controller within closed-loop control? In the first instance, a directive to promote engagement during an adaptive game may have unintended negative side effects for the player, e.g. spend too long playing the game, suffer from fatigue and sleeplessness. Even if these caveats are captured within the rules of the system, there are other hurdles to be faced with respect to materialisation of goals and values. Precise definition of goals and values may differ enormously between different members of the user population. In addition, there may be a number of stakeholders aside from the user who are directly or indirectly affected by the directives of the system, e.g. user's line manager & colleagues, user's family, system designer, corporation who supplied technology etc. There is also the potential for ambiguity or conflict because the definition of a goal for the loop may differ at the levels of individual, society and nation [6]. For example, a closed-loop system designed to improve productivity in a company could enfranchise the board of directors whilst enslaving their employees. It may be unrealistic to expect technology to encompass convivial goals per se, but rather we should seek to build conviviality into technological tools by carefully defining the context and operating conditions under which technology is used [5].

The use of technology to explicitly enshrine and define our human values presents a number of significant challenges, as well as considerable opportunities to use technology as a vehicle to enshrine and develop a humanist agenda - in the words of Arthur [4] “we trust in nature but we hope in technology” (p. 246).

3 First- and Second-Order Adaptation

The biocybernetic loop encompasses a process of monitoring the user and translating those data into intelligent adaptation at the interface. This procedure requires a set of rules whereby target state *a* triggers adaptive response *x,* however, this relationship is not an exclusive and there may be a range of potential responses that are appropriate once a specific target state has been recognised by the system. A detection of frustration could trigger an offer of help or the suggestion of a rest break or an alteration of current music to a calming playlist. The rules that translate detection into an adaptive response may draw from a repertoire of possibilities, all of which could conceivably result in a desired effect on the user. In addition, some users may favour certain categories of adaptive response from the repertoire over others.

It is the convention to think of closed-loop systems in terms of one discrete cycle of monitoring and adaptation. In this case, a single cycle may describe how the detection of frustration is translated into the appearance of help information at the interface. This is a first-order process of adaptation wherein the loop detects and responds to a target state in the short-term. Once this adaptation has been activated, it is possible for the system to detect those changes in user state, which occur as a direct consequence of that adaptive response. If help is offered in order to alleviate frustration, the continual process of monitoring will indicate whether this response successfully achieved its goal. If no such change occurs, or if frustration actually increased, the adaptive controller must select a different response from its repertoire, such as selecting a playlist of calming music. Once the calming music has been activated for a short period, the system can perform a third check to assess whether frustration has been alleviated as expected. This process is called second-order adaptation or reflexive adaptation [25] because the loop monitors the consequences of its own intervention on the state of the user. This second-order level of adaptation fulfills two functions, it is a self-check (that the original adaptive response was effective) and represents an opportunity for a closed-loop system to collate information about user preferences based a long-term process of repeated interaction.

It is easy to understand how this second-order process of adaptation can facilitate machine learning over a sustained period of use. In order for the system to function, it must accumulate a database that describes those adaptive responses found to be effective for a particular user and those that are not. Therefore, the system is installed and initiated with a large number of potential adaptations, and through a process of sustained interaction coupled with second-order processing, all items in the adaptive repertoire are tagged with a value, which directly affects the probability of selection for that specific user. Second-order adaptation describes a generative process of individualisation where software is customised on the basis of its repeated interactions with a particular user. Second-order adaptation also represents a level of human-machine symbiosis where the technology is able to learn about the effects of its own actions.

The evolving lifecycle of this reflexive technology has been described as a process of mutual adaptation with three main phases [25, 26]. The initial encounters between the adaptive system and the user are characterised by a process of *improvisation*. The system responds to the user in a generic fashion using default adaptations with no prior knowledge of individual preferences. Adaptation may be perceived by the user to be erratic and occasionally inappropriate. As the user spends more time interacting with the system, second-order adaptation should improve the timeliness and quality of the responses made by the system. This second phase of *reciprocal coupling* is characterised by enhanced performance as the adaptive repertoire of the system is tailored to the individual. This is the phase wherein the system constructs a stable model of user preferences based on repeated interactions. If we look further ahead in time, in terms of years and decades, it is reasonable to expect that any stable model of preferences will have limited longevity as the user acquires higher levels of skill or habituates to popular responses or experiences cognitive changes due to ageing. The third phase of *co-evolution* describes a process of updating the existing model of user preferences as the system adjusts to long-term changes over several years. This cycle of monitoring, adaptation and reflexive adaptation represents perhaps the ultimate expression of user-centred software design.

A process of reflexive adaptation may also have some bearing on the problem of formalising convivial goals within a technological system described in the previous section. These difficulties were recognised over fifty years ago by Norbert Weiner [27]; his solution was to build cycles of self-correction into the loop by inserting regular interventions from a human arbitrator within the learning process of the cybernetic loop. This strategy was suggested as a safeguard to ensure that the actions of the machine did not significantly depart from the preferences and values of the human being. The capacity of the biocybernetic loop to interact with the human central nervous system continuously and over a sustained period of time captures the essence of this idea - provided that implicit data from the brain and body are sufficiently nuanced to intercede on behalf of the person; however, there are concerns about the test-retest reliability of psychophysiological measures in the field [28]. For this strategy to act as a proxy for the human arbitrator, much depends on the sensitivity and reliability of the data used to represent the user, if these data are inconsistent then the possibilities for machine learning in the long-term are fundamentally compromised.

4 Technology for Life

The development of symmetrical HCI via the biocybernetic loop reconfigures the relationship between people and computers. Our earlier characterisation where the human "speaks" and the computer "listens" remains relevant, but with the additional caveat that the computer can now "speak back." This machine with an agenda is active and dynamic as opposed to the passive and static technologies that we currently use on a daily basis. A nascent form of closed-loop control offers the prospect of smart technology, capable of intelligent adaptation and personalisation, but at the price of subverted human autonomy. This change does not mean simply that the traditional roles of human and machine are recomposed, by converting the user into a pattern of information that is *operated upon* within a closed-loop, the loop obscures the boundary between human and computer. Within this conception, human and machine function as a single "cooperative intelligent entity" [29] - a cybernetic organism that is capable of learning based on previous interaction to create a flexible repertoire of adaptive responses.

We have already described how technology can supplement our human capacities and capabilities. Consider the inverse of that position - how can humans develop the capacities, proficiencies and potential of technology? According to Kelly [5], the developmental trajectory of technology is characterised by universal tendencies towards: complexity, diversity, freedom, mutualism, sentience and evolvability. These inclinations are accelerated by the concepts described in this paper. The closed-loop logic of symmetrical HCI requires the additional complexity of monitoring and representing the human user. The capacity of the loop to facilitate learning in the longer-term creates the potential for greater diversity within the same piece of software, i.e. software co-evolves with the individual user, begetting a generative process where different patterns of development are possible within the same technology. The loop is a machine with an agenda and this agenda imbues technology with the freedom to make mistakes and to learn from those mistakes in order to make better choices in future. The loop is a human-machine hybrid that deepens the degree of cooperation, dependency and mutualism between person and computer. The process of second-order adaptation permits technology to reflect on the effects of its own actions, thus creating a rudimentary form of sentience. Most importantly, the process of monitoring and adaptation allows technology to develop advanced capabilities by learning directly from repeated interaction with human users. Several authors have described a process of bootstrapping [5, 8] whereby humans supplement their skills and capabilities via technology, we may now contemplate a future where closed-loop technology uses sustained interaction with people as an engine to boost capabilities and accelerate its own evolutionary development.

One hopes that such exciting and provocative developments occur in a convivial spirit, thus maximising the potential and possibilities for all human life. However, living so closely with technology has the potential to create several significant problems for our species. There is the obvious issue of control or rather uncontrollability when a person submits to interaction with technology within a closed-loop. By relinquishing total control over technology, there is the potential to undermine human agency; in the words of Wiener [18]: "When human atoms are knit into an organisation in which they are used, not in their full rights as responsible human beings, but as cogs and levers and rods, it matters little that their raw material is flesh and blood" (p. 185). There is also the problem of data privacy, intrusion and misrepresentation via the process of monitoring within the loop [30]. It has already been emphasised that representation of self within the loop is an analogous creation rather than a literal re-representation of thoughts, moods and experiences. The act of interacting with this analogous representation, which is both self and other, has the potential to simultaneously alienate the individual and could even create feelings of disembodiment [8]. Like all systems that automate or semi-automate, symmetrical HCI has the potential to de-skill the individual [31], whether that person is driving a car or playing a computer game.

The long-term relationship between humanity and technology has been characterised as an infinite game [5] and the purpose of an infinite game is not to win but to keep playing. The burgeoning complexity of our relationship with machines emphasises how any attempt to sustain human beings in the sovereign position of a master who retains ultimate control over his technological creation are doomed to failure [6]. We must explore new trajectories of interaction with technology, which maximises opportunities for both humans and machines as a single intelligent cooperative entity.

5 Summary

Our historical relationship with technology has been characterised by the use of tools being used to extend human capabilities and capacities. We are currently entering a period where symmetrical HCI via physiological computing will lead to greater mutualism between people and computers. It is argued that emerging technology will demonstrate greater intelligence during interactions with people by monitoring and affecting user psychology. In addition, these 'smart' technologies will be capable of anticipating the needs of the individual and personalising responses; they will respond in an active and stochastic pattern. In order to reap these benefits, humans must submit themselves to implicit monitoring by technology, allow complex and embodied internal states to be reduced to sparse, analogous representations, and cede a degree of control to the computer.

The challenge for designers of this emergent technology is to enable this transition in a convivial fashion to:

1. Ensure that human user can disable the adaptive process at any time
2. Ensure that human user can manually edit (i.e. enable/disable) the repertoire of adaptive responses
3. To carefully formulate adaptive responses from the system that are compatible with the goals and values of the user
4. To use second-order monitoring to ensure that adaptive responses are desirable from the perspective of the user
5. Educate users with respect to the internal logic of the system in order for engender trust in the technology via enhanced understanding [32]

If these compromises can be made in a convivial fashion, machines can be permitted to learn from regular interaction with the individual in order to customise responses to the preferences of the individual. The creation of an intelligent, cooperative entity, which arises from close coupling between human or machine, will increase benefits and opportunities for both parties.

References

1. Norman, D.A.: The Design of Future Things. Basic Books, New York (2007)
2. Hettinger, L.J., et al.: Neuroadaptive technologies: applying neuroergonomics to the design of advanced interfaces. Theoretical Issues in Ergonomic Science **4**(1–2), 220–237 (2003)
3. Martin, B.D., Schwab, E.: Current usage of symbiosis and associated terminology. International Journal of Biology **5**(1), 32–42 (2013)
4. Arthur, W.B.: The Nature of Technology: What It Is and How It Evolves. Penguin (2009)
5. Kelly, K.: What Technology Wants. Penguin (2010)
6. Hancock, P.A.: Mind, Machine and Morality: Towards a Philosophy of Human-Technology Symbiosis. Ashgate (2009)
7. Clark, A., Chalmers, D.J.: The extended mind. Analysis **58**, 10–28 (1998)
8. Clark, A.: Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence. Oxford University Press (2003)
9. Douglas, A.E.: The Symbiotic Habit. Princeton University Press, New Jersey (2010)
10. Licklider, J.C.R.: Man-Computer Symbiosis. IRE Transactions on Human Factors in Electronics **HFE-1**, 4–11 (1960)
11. Jacucci, G., Spagnolli, A., Freeman, J., Gamberini, L.: Symbiotic interaction: a critical definition and comparison to other human-computer paradigms. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 3–20. Springer, Heidelberg (2014)
12. Allanson, J., Fairclough, S.H.: A research agenda for physiological computing. Interacting with Computers **16**, 857–878 (2004)
13. Fairclough, S.H.: Fundamentals of Physiological Computing. Interacting with Computers **21**, 133–145 (2009)
14. Pope, A.T., Bogart, E.H., Bartolome, D.S.: Biocybernetic system evaluates indices of operator engagement in automated task. Biological Psychology **40**, 187–195 (1995)
15. Byrne, E., Parasuraman, R.: Psychophysiology and adaptive automation. Biological Psychology **42**, 249–268 (1996)
16. Fairclough, S.H., Gilleade, K.: Construction of the biocybernetic loop: a case study. In: 14th ACM International Conference on Multimodal Interaction, pp. 571–578. ACM, Santa Monica (2012)
17. Haynes, N.K.: How We Became Post Human: Virtual Bodies in Cybernetics, Literature and Informatics. University of Chicago Press (1999)
18. Wiener, N.: The Human Use Of Human Beings: Cybernetics & Society. Da Capo Press, Boston (1954)
19. Wilson, E.O.: Biophilia. Harvard University Press, Cambridge (1984)
20. Wilson, G.F.: Pilot workload, operator functional state and adaptive aiding. In: Hockey, G.R.J., Gaillard, A.W.K., Burov, O. (eds.) Operator Functional State: The Assessment and Prediction of Human Performance Degradation in Complex Tasks, pp. 194–203. ISO Press, Amsterdam (2003)
21. Kapoor, A., Burleson, W., Picard, R.W.: Automatic prediction of frustration. International Journal of Human-Computer Studies **65**, 724–736 (2007)
22. Dekker, A., Champion, E.: Please biofeed the zombies: enhancing the gameplay and display of a horror game using biofeedback. In: DiGRA (2007)
23. Mishra, J., Gazzaley, A.: Closed-loop cognition: the next frontier arrives. Trends Cogn. Sci. **19**(5), 242–243 (2015)
24. Illich, I.: Tools For Conviviality. Harper and Row, New York (1973)

25. Serbedzija, N., Fairclough, S.H.: Reflective pervasive systems. ACM Transactions on Autonomous and Adaptive Systems **7**(1) (2012)
26. Serbedzija, N., Fairclough, S.H.: Biocybernetic loops: from awareness to evolution. In: IEEE Congress on Evolutionary Computation, Trondheim, Norway (2009)
27. Wiener, N.: God and Golem Inc. MIT Press, Cambridge (1964)
28. Tomarken, A.J.: A psychometric perspective on psychophysiological measures. Psychological Assessment **7**(3), 387–395 (1995)
29. Rahimi, M., Hancock, P.A.: Optimization of hybrid production systems: the integration of robots into human-occupied work environments. In: Brown, O., Hendricks, H.W. (eds.) Human Factors in Organizational Design and Management – II, pp. 39–54. Elsevier, North-Holland (1986)
30. Fairclough, S.H.: Physiological data should remain confidential. Nature **505**, 263 (2014)
31. Parasuraman, R., Riley, V.: Humans and automation: use, misuse, disuse, abuse. Human Factors **39**, 230–253 (1997)
32. Miller, C.A.: Trust in adaptive automation: the role of etiquette in tuning trust via analogic and affective methods. In: First International Conference on Augmented Cognition, Las Vegas, NV (2005)

Developing a Symbiotic System for Scientific Information Seeking: The MindSee Project

Luciano Gamberini[1,5], Anna Spagnolli[1,5], Benjamin Blankertz[2], Samuel Kaski[3], Jonathan Freeman[4], Laura Acqualagna[2], Oswald Barral[6], Maura Bellio[5], Luca Chech[1], Manuel Eugster[3], Eva Ferrari[4], Paolo Negri[1], Valeria Orso[1], Patrik Pluchino[5], Filippo Minelle[5], Bariş Serim[6], Markus Wenzel[2], and Giulio Jacucci[3,6](✉)

[1] Department of General Psychology, University of Padua, Padua, Italy
[2] Neurotechnology Group, Technische Universität Berlin, Berlin, Germany
[3] Helsinki Institute for Information Technology (HIIT), Aalto University, Espoo, Finland
[4] Department of Psychology, I2 Media Research, Goldsmiths University of London, London, UK
[5] Human Inspired Technology Research Centre (HIT), University of Padua, Padua, Italy
[6] Department of Computer Science, Helsinki Institute for Information Technology (HIIT), University of Helsinki, Helsinki, Finland
giulio.jacucci@helsinki.fi

Abstract. This paper describes an approach for improving the current systems supporting the exploration and research of scientific literature, which generally adopt a query-based information-seeking paradigm. Our approach is to use a symbiotic system paradigm, exploiting central and peripheral physiological data along with eye-tracking data to adapt to users' ongoing subjective relevance and satisfaction with search results. The system described, along with the interdisciplinary theoretical work underpinning it, could serve as a stepping stone for the development and diffusion of next-generation symbiotic systems, enabling a productive interdependence between humans and machines. After introducing the concept and evidence informing the development of symbiotic systems over a wide range of application domains, we describe the rationale of the MindSee project, emphasizing its BCI component and pinpointing the criteria around which users' evaluations can gravitate. We conclude by summarizing the main contribution that MindSee is expected to make.

Keywords: Symbiotic system · Implicit measures · BCI · Information seeking · User experience

1 Symbiotic Systems Today

The last decade has witnessed a radical change in the structure and complexity of the software used in everyday life. An ever-growing number of systems have

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 68–80, 2015.
DOI: 10.1007/978-3-319-24917-9_7

started to do more than passively responding to users' input, and we have tacitly come to expect an increasing amount of proactivity on the systems' part. When we shop online, we want the website or the search engine to recommend articles based on our or other people's purchase history; when we open a music service, we expect it to suggest other artists to listen to based on our preferred genres. We want these systems to know what we are trying to accomplish in a specific context, and we want them to help us, support us and give us recommendations. In short, we want systems to understand us. This expectation is mirrored, in the scientific literature, by the appearance in the vocabulary of human-computer interaction scientists, software developers and information technology (IT) practitioners in general of terms such as recommender engines, adaptive technologies and symbiotic systems. The practice of representing the interaction between humans and computers in terms of symbiosis dates back to the 1960s. Licklider borrowed this notion from biology [1], where it refers to an intimate association between two different organisms characterized by a constant and mutual cooperative behavior. In the IT domain, instead, a symbiotic system indicates a dyad composed by the user and a given technology, which is capable of adjusting rapidly (possibly in real time) to the user's characteristics, in order to improve the quality and efficiency of the interaction. As Jacucci, Spagnolli, Freeman and Gamberini [2] have stated, this involves the "combination of computation, sensing technology, and interaction design to realize deep perception, awareness, and understanding between humans and computers" (p. 11). An adaptive technology feature is both a hard component and a soft component [3]. The hard component relates to the device (or devices) enabling the collection of users' information (e.g., eye trackers, haptic gloves, speech, gesture, emotion recognition). The soft component, in turn, comprises the software, algorithms, decision heuristics and models that act on the user's information and obtain the desired adaption in the system's interface and features.

1.1 Application Domains

This paradigm has started to appear in several application domains, ranging from learning to mobile Internet, human-robot interaction and information retrieval. An application in education is described by Shute and Zapata-Rivera [3]. The authors illustrate a "four-process adaptive cycle" that starts with collecting information on the student, using, for example, eye-tracking, haptic, speech-capture and hand-gesture-capture devices. The system then generates and updates a user model and relies on machine-learning algorithms to elaborate a representation of the user, inferring the user's attentional state, the difficulty experienced during the execution of a specific task, the user's emotional state, etc. The last two steps consist of selecting and presenting appropriate educational material based on the user model. An example of a symbiotic system in the domain of the mobile Internet, and specifically in location-based services, is provided by OUTMedia, a location-sensitive music discovery application [4]. OUTMedia includes an augmented reality (AR) interface, linking music content to specific points of interest in the city to support a dynamic adaptation of media and environment. A user study showed that the use of OUTMedia promotes serendipity in content discovery. In the field of human-robot interaction, an interesting case is offered by wearable robots (WRs), devices conceived to support a human limb for an unhealthy or healthy user. For instance, the European project BioMot aims to improve adaptations between humans and WRs by developing a symbiotic relation based on a framework that takes into account information regarding the dynamics of human gait and the environment [5]. This paper focuses on information retrieval and data searching in complex scenarios as the application domain for symbiotic systems. Studies encouraging and directing efforts in this vein have already appeared. The relationship between semantic processing of words and pupil dilation is a case in point and has been investigated by [6]. The semantic association between two words displayed in succession on a screen was manipulated in each experimental trial. The results showed that faster pupil dilation was observed in trials in which the words were semantically associated. These findings could be used in the design of information retrieval systems, for instance to use pupil changes to implicitly detect the relevance of the displayed information. [7] explores potential interaction between eye gaze and physiological signals (e.g., EEG, EDR and fEMG) and some aspects of information visualization. [8] proposes an interactive image retrieval system, allowing the user to mark images as relevant or not. This process of selection is based, through an iterative process, on the search intent of the user. Based on this input, the system progressively displays results that are closer to those marked as relevant. Other studies have dealt with physiological computing methods to infer users' internal states, of which users are not totally aware [9]. This strategy is supported by recent studies showing that subliminal stimulation can affect users' decisions in selection behavior [10,11] and in navigation tasks in immersive virtual environments [12]. Thus, authors have considered subliminal stimuli to guide the users during their interactions without burdening their cognitive systems [13]. One kind of implicit information that can be used in a symbiotic system is a signal of brain activity. In the next paragraph, we will describe a brain-computer interface (BCI) and how it can be employed in information retrieval.

1.2 BCI in Everyday Symbiotic Applications

The term brain-computer interface (BCI) was introduced in 1973 [14], identifying a research domain in which techniques to extract specific information from the brain in real time and use it as input in computerized devices are investigated. A recent, widely accepted definition describes a BCI system as one that "measures central nervous system (CNS) activity and converts it into artificial output that replaces, restores, enhances, supplements, or improves natural CNS output and thereby changes the ongoing interactions between the CNS and its external or internal environment"[15]. BCI research in the past was almost exclusively employed in medical contexts for the (partial) restoration and rehabilitation of lost communication functions in paralyzed patients or for motor rehabilitation [16,17]. Clinical applications like these rely on intentional control, because the user needs to acquire specific mental states in order to produce a desired output, such as selecting a letter or moving a prosthesis [18,19]. Novel approaches aim instead to exploit BCI technology's utility for estimating implicit information about the user's internal state [20,21] and to envision applications in domains besides the clinical one. Application domains include product design [22,23], industrial workplaces [24–26], video games [27] and driving. The latter application domain, for instance, regards the prediction of emergency braking based on the ongoing EEG signals of the driver (see [28] for a first investigation with a simulator and [29] for a replication on real roads). More examples and a description of possible future developments can be found in the roadmap of brain/neural computer interaction (BNCI [30]). The actual implementation of BCI technology to infer users' intentions and to use implicit information in real-world applications is still a great challenge, especially when it comes to integrating implicit measures of perception with cognitive and emotional responses. This investigation requires a feasible domain in which it is realistic to explore and demonstrate the power of a symbiotic approach, and information seeking represents a promising domain for this purpose. The goal of a symbiotic system applied to information seeking would be to increase the efficiency of users' activity by utilizing physiology-based predictions (central and peripheral physiology) of their intentions (e.g., of the subjective relevance of informational items that are displayed on the screen). This requires a holistic interpretation of behavioral data, perception processes and cognitive responses to interface features, as well as of emotional responses to the visualized information. In the rest of this paper, we will describe the symbiotic system developed in the MindSee project to support information seeking in complex databases. First, we will provide an overview of the state-of-the-art advancement in information-seeking systems. Against that background, we will describe the symbiotic system for information retrieval developed as part of the MindSee project. We will then explain more specifically the kind of BCI implemented in the MindSee system and the criteria for the user experience evaluation.

2 State-of-the-Art Systems for Information Seeking

Current trends in information exploration aim at providing better support for interactive search systems through query suggestion, recommendations for query formulation, and implicit or explicit relevance feedback (RF); this includes personalization on the one hand and information visualization solutions to represent complex information on the other hand [31]. Information visualization has also recently gained importance in the case of open-ended exploration of information, providing continued support and engagement resources along with 3D graphics and gestural computing [32–34]. The increased complexity of information search interfaces complicates the application of query-and-response paradigms and of relevance feedback. MindSee addresses advancements in both aspects simultaneously. In a simple user interface with a query field and results list, a typical interactive information retrieval problem presumes the user's expression of an information need (the query). Based on this query, the system needs to search a corpus of documents to find those matching the query, to present the results to the user and to support the user's reformulation of the query. However, in many exploratory search situations, it is difficult, if not impossible, to formulate such a query precisely. According to [35], 35% to 50% of information needs are exploratory and spread across several individual queries, aiming to find a variety of relevant information by browsing through a large collection of documents, rather than looking for a known document. A commonly observed search strategy is one in which the information seeker issues a quick, imprecise query in the hope of getting into approximately the right part of the information space, and then locally navigates to obtain the information of interest [36,37]. The information need can also evolve throughout the course of the search as the user gains more information about the topic and navigates to alternative topics or more specific subtopics. Even when the searched item is known from the beginning, instead of jumping directly to the target, users typically navigate "with small, local steps, using their contextual knowledge as a guide" (p. 415) [38]. As a consequence, users are confronted with a substantial cognitive load; they must choose a way to express their evolving information needs, to make sense of a search domain and to position themselves in the information space [39]. One standard solution to improve and direct the search process in an underspecified and uncertain domain is relevance feedback. Relevance feedback provides a controlled and broken down process to alter and improve the initial query step-by-step. It requires users to provide feedback on whether they like or dislike the resulting documents or document features that are initially returned by an information retrieval system for a given query, and to use this information to alter the query itself [40]. Users' feedback can be explicit or implicit. Explicit feedback means feedback in which users actively choose from documents or document features. While highly effective when properly used, the selection of documents or document features can be cognitively demeaning for users. This has been shown to lead users to easily abandon the feedback features and reformulate the queries themselves instead [40]. It has been suggested that for explicit feedback to be effective, it should be a part of a natural interaction with the information retrieval system, or directly accompanied with implicit feedback [41,42]. Implicit feedback is inferred from user behavior, such as which documents users select, the duration of time spent viewing a document or a piece of text within a document, or scrolling actions [43]. Mainly, the research has concentrated on analyzing click data, i.e., the selection of web documents, and using the selection as a part of predictive modeling to improve the web search. However, implicit feedback is highly dependent on the quality of the signal source. In many cases, the feedback is noisy and may even result in prohibitive system behavior. Existing feedback techniques determine content relevance only with respect to the cognitive and situational levels of interaction, failing to acknowledge the importance of intentions, motivations and feelings in cognition and decision making [44–46]. Only recently, researchers have gained interest in more advanced psychophysiological sensor measurements [44], such as skin conductance, gaze patterns [47] and EEG [48,49], as a part of relevance feedback. Recently, research has shown how to detect term relevance from brain signals in an information retrieval scenario [50]. Similarly Barral et al. [51] showed relevance detection using peripheral detection in abstract reading. These studies demonstrate the viability of a symbiotic search scenario, although they highlight a number of open challenges.

3 MindSee Project: Exploiting Symbiotic Systems for Scientific Information Seeking

The main purpose of the MindSee project is to create a symbiotic system for the exploration and retrieval of scientific literature. Its starting platform is SciNet, a cutting-edge retrieval system that includes a database of 50 million documents from prominent scientific databases and offers explicit feedback features (e.g., manual modification via mouse of the spatial organization of the displayed information). This platform is upgraded in three ways. First, the MindSee project aims to create a coadaptation of user and computing systems that is based on implicit information (e.g., EEG phenomena coupled with peripheral physiological measures). The project intends to build a system that would be able to predict search intentions by integrating, at a multimodal level, neurophysiology, peripheral physiology and behavioral data, and by complementing explicit user commands with these predictions. In order to achieve its purpose, MindSee will extend psychophysical studies on perception and cognition in the context of free-viewing tasks, in order to identify suitable EEG features. Second, compared to a generic digital library, in which the adoption of lengthy result lists makes it difficult to exploit explicit relevance feedback and to assess the relevance of most items at first glance, MindSee's system will adjust the information output spatially and aesthetically, according to each result's relevance to the current search - for instance, by changing the complexity level of the visualization or by highlighting items that are under early attentional processing. Third, the system presents an affective component of symbiotic interaction to guarantee a superior user experience. This is achieved by quantifying the level of engagement, pleasure, stress and user emotions related to the interaction through a combination of EEG and peripheral signals (EDA and fEMG), and using these measures to adapt the graphic characteristics of the interface (for instance, assigning different colors to items).

3.1 BCI in MindSee

MindSee will capitalize on recent advances in BCI based on physiological signals combined with machine-learning approaches. The symbiosis in the MindSee system will be mostly based on real-time analysis of the brain's electrical activity (electroencefalography, or EEG). Brain signals that are linked to an item that is currently under process could provide information about perceptual elaboration and, most importantly, cognitive processing. Since the cognitive elaboration could reflect the relevance of a specific item for the user, users' searching intent could be inferred by monitoring specific brain signals [52,53] and specifically the EEG signals related to the focused items (duration of the event-related attenuation of the alpha rhythm, see [24]). In addition, a range of eye and pupil measures (eye tracking, or ET, and pupillometry) can be employed to determine which item the user is currently inspecting and to relate the user's perceptual and cognitive processes to a specific item or set of items. For instance, the pupil diameter could be useful to disclose increments in the user's cognitive load [54–56] or to investigate the relevance of a search result [57]. In addition, since the responses of the pupil to specific stimuli/events provide a continuous measure of cognitive processing, despite the fact that the user is unaware of such variations, pupillometry could help detect the level of preconscious processing [58]. Pupil dilation is a simple measure to acquire; the collection procedure is totally noninvasive and is characterized by a short temporal delay, which is useful since the final MindSee system should function almost simultaneously with the ongoing user's perceptual and cognitive processing and should adapt without delays. The MindSee system will also take advantage of peripheral physiological measures (electrodermal activity, or EDA, and facial electromiography, or fEMG) in order to collect information about the arousal level and emotional state of the user.

3.2 User Experience in Symbiotic Systems

One of the aims of the MindSee project is to determine the proper metrics for assessing the usability of a symbiotic system. Symbiotic systems can be thought of as consisting of two main components: one reacting to direct users' inputs (i.e., their explicit behavior), and one interpreting and responding to users' affective and cognitive states through the online analysis of their neurophysiological data [59]. In other words, not only does the system react to explicit behavioral input, but it also collects users' physiological indices as an implicit input to infer users' cognitive and affective states [60]. The final goal is to better support users' activity without placing a burden on their cognitive resources [61–63]. Symbiotic systems can include wearable components ensuring a greater cohesion between the user and the environment and allowing the user to (implicitly) act on an interface, even while engaged with other activities or devices [62,64]. When evaluating the user experience, the peculiarities of symbiotic systems make it necessary to adjust some common metrics and to add some specific ones. First, the scenario that users have in front of them is continuously evolving, as a consequence of the system's response to their cognitive and affective states. Therefore some users' intermediated goals might change as the activity proceeds, even though the overall goal remains the same; this can make it inappropriate to use a score system based on predefined activity steps when evaluating accuracy [65] or efficiency. Second, slowdowns or interruptions of an action due to a crisis in the interpretation of the system's functionalities [66] might be recognized by the system, which would adjust the information presented accordingly; therefore, the usability test should evaluate not only the occurrence of a breakdown in the interaction with the system, but also the circumstances under which the system adaptation is helpful in resolving the breakdown. Third, relevant subjective dimensions to consider when evaluating symbiotic systems are satisfaction, acceptance, perceived sense of control, perceived utility, credibility and comfort [60], which can be measured with self-reported techniques. In particular, satisfaction with the system's ability to effectively adapt to the user's state needs to be evaluated. However, since the user is not aware of the implicit data used by the system and is probably only partially aware of the adaptation process, questionnaires can only inquire about the overall impression of being relieved in their activity and of being well understood by the system [60]. They need to be accompanied by other kinds of data to measure the actual cognitive load (e.g., pupil dilation) or stress (electrodermal activity, or EDA), or rely on the validation design in order to evaluate the gain in performance due to the use of symbiotic features. One way to make this evaluation would be to compare the user experience when such features are active and when they are not. Finally, the system needs to be evaluated in the way in which it fulfills ethics and data protection requirements.

4 Conclusions

We started this paper by arguing that symbiotic systems - as they have been recently defined by Jacucci et al. [2] (symbiotic) - help meet current users' expectation that systems can understand their needs and adapt accordingly. We also argued that information seeking is an ideal domain where symbiotic applications can be tested in everyday life activities, since they allow one to explore the integrated adaptation to the users' implicit and explicit cognitive, perceptual and emotional state. The MindSee project takes up this opportunity and upgrades an existing search tool called SciNet with symbiotic features based mainly on BCI components, but also on other kinds of implicit user input (e.g., behavioral). With respect to the state of the art in information seeking, then, the symbiotic strategy attempted in the MindSee project provides an advanced solution for searching in an unfamiliar or underspecified domain by implementing implicit feedback, allowing query refinement without asking the user to explicitly mark the first query results. The rationale consists of relating a specific item or set of items displayed as a search result with the user's perceptual and cognitive processes, namely the duration of the event-related attenuation of the alpha rhythm in EEG signal, pupil diameter, electrodermal activity and facial electromyography. Since the system's ultimate ambition is to improve performance by reducing users' cognitive load, the actual transparency, naturalness and helpfulness of the system must be evaluated with an appropriate combination of self-reported and observational measures. Similarly, the acceptance of a system that uses implicit data and an adaptation process of which the user is mostly unaware must be measured.

Acknowledgments. This research was supported by the European Project MIND-SEE, Symbiotic Mind Computer Interaction for Information Seeking (Number: 611570; Call: FP7-ICT-2013-10 Information and Communication Technologies).

References

1. Licklider, J.C.R.: Man-computer symbiosis. IRE Trans. on Hum. Factors Electron. **1**, 4–11 (1960)
2. Jacucci, G., Spagnolli, A., Freeman, J., Gamberini, L.: Symbiotic interaction: a critical definition and comparison to other human-computer paradigms. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 3–20. Springer, Heidelberg (2014)
3. Shute, V.J., Zapata-Rivera, D.: Adaptive technologies. In: Handbook of Research on Educational Communications and Technology, pp. 277–294 (2008)
4. Åman, P., Liikkanen, L.A., Jacucci, G., Hinkka, A.: OUTMedia – symbiotic service for music discovery in urban augmented reality. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 61–71. Springer, Heidelberg (2014)
5. Moreno, J.C., et al.: Symbiotic wearable robotic exoskeletons: the concept of the biomot project. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 72–83. Springer, Heidelberg (2014)
6. Pluchino, P., Gamberini, L., Barral, O., Minelle, F.: How semantic processing of words evokes changes in pupil. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 99–112. Springer, Heidelberg (2014)
7. Serim, B.: Querying and display of information: symbiosis in exploratory search interaction scenarios. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 115–120. Springer, Heidelberg (2014)
8. Hore, S., Tyrvainen, L., Pyykko, J., Glowacka, D.: A reinforcement learning approach to query-less image retrieval. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 121–126. Springer, Heidelberg (2014)
9. Barral, O., Jacucci, G.: Applying physiological computing methods to study psychological, affective and motivational relevance. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 35–46. Springer, Heidelberg (2014)
10. Aranyi, G., Kouider, S., Lindsay, A., Prins, H., Ahmed, I., Jacucci, G., Negri, P., Gamberini, L., Pizzi, D., Cavazza, M.: Subliminal cueing of selection behavior in a virtual environment. Presence: Teleoperators Virtual Environ. **23**(1), 33–50 (2014)
11. Barral, O., et al.: Covert persuasive technologies: bringing subliminal cues to human-computer interaction. In: Spagnolli, A., Chittaro, L., Gamberini, L. (eds.) PERSUASIVE 2014. LNCS, vol. 8462, pp. 1–12. Springer, Heidelberg (2014)
12. Cetnarski, R., Betella, A., Prins, H., Kouider, S., Verschure, P.F.: Subliminal Response Priming in mixed reality: The ecological validity of a classic paradigm of perception. Presence: Teleoperators Virtual Environ. **23**(1), 1–17 (2014)
13. Negri, P., Gamberini, L., Cutini, S.: A review of the research on subliminal techniques for implicit interaction in symbiotic systems. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 47–58. Springer, Heidelberg (2014)

14. Vidal, J.J.: Toward direct brain-computer communication. Annu. Rev. Biophys. Bioeng. **2**(1), 157–180 (1973)
15. Wolpaw, J.R., Wolpaw, E.W.: Brain-computer interfaces: something new under the sun. In: Wolpaw, J.R., Wolpaw, E.W. (eds.) Brain-Computer Interfaces: Principles and Practice, pp. 3–12. Oxford University Press, Oxford (2012)
16. Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kübler, A., Perelmouter, J., Taub, E., Flor, H.: A spelling device for the paralysed. Nature **398**(6725), 297–298 (1999)
17. Sellers, E.W., Vaughan, T.M., Wolpaw, J.R.: A brain-computer interface for long-term independent home use. Amyotroph. Lateral Scler. **11**, 449–455 (2010)
18. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain-computer interfaces for communication and control. Clin. Neurophysiol. **113**(6), 767–791 (2002)
19. Pfurtscheller, G., Allison, B.Z., Brunner, C., Bauernfeind, G., Solis-Escalante, T., Scherer, R., Zander, T.O., Mueller-Putz, G., Neuper, C., Birbaumer, N.: The hybrid BCI. Front. Neurosci. **4** (2010)
20. Blankertz, B., Tangermann, M., Vidaurre, C., Fazli, S., Sannelli, C., Haufe, S., Maeder, C., Ramsey, L., Sturm, I., Curio, G., Müller, K.R.: The Berlin brain-computer interface: non-medical uses of BCI technology. Front. Neurosci. **4** (2010)
21. Brunner, C., Birbaumer, N., Blankertz, B., Guger, C., Kübler, A., Mattia, D., Millán, J.D.R., Miralles, F., Nijholt, A., Opisso, E., Müller-Putz, G.R.: BNCI Horizon 2020: towards a roadmap for the BCI community. Brain-Computer Interfaces (ahead-of-print), 1–10 (2015)
22. Acqualagna, L., Bosse, S., Porbadnigk, A.K., Curio, G., Müller, K.R., Wiegand, T., Blankertz, B.: EEG-based classification of video quality perception using steady state visual evoked potentials (SSVEPs). J. Neural Eng. **12**(2), 026012 (2015)
23. Porbadnigk, A.K., Treder, M.S., Blankertz, B., Antons, J.N., Schleicher, R., Möller, S., Curio, G., Müller, K.R.: Single-trial analysis of the neural correlates of speech quality perception. J. Neural Eng. **10**(5), 056003 (2013)
24. Venthur, B., Blankertz, B., Gugler, M.F., Curio, G.: Novel applications of BCI technology: psychophysiological optimization of working conditions in industry. In: Systems, Man and Cybernetics (SMC 2010), pp. 417–421. IEEE Press, New York (2010)
25. Schultze-Kraft, M., Birman, D., Rusconi, M., Allefeld, C., Görgen, K., Dähne, S., Blankertz, B., Haynes, J.D.: Predicting and interrupting movement intentions with a closed loop. In: Presented at the 21st Annual Meeting of the Organization for Human Brain Mapping, Honolulu, USA (2015)
26. Di Flumeri, G., Borghini, G., Aric, P., Colosimo, A., Pozzi, S., Bonelli, S., Golfetti, A., Kong, W., Babiloni, F.: On the use of cognitive neurometric indexes in aeronautic and air traffic management environments. In: Blankertz, B., Jacucci, G., Gamberini, L., Spagnolli, A., Freeman, J. (eds.) Symbiotic 2015. LNCS, vol. 9359. Springer International Publishing, Cham (2015)
27. Bos, D.P.O., Reuderink, B., van de Laar, B., Gürkök, H., Mühl, C., Poel, M., Nijholt, A., Heylen, D.: Brain-computer interfacing and games. In: Tan, D., Nijholt, A. (eds.) Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction, pp. 149–178. Springer, London (2010)
28. Haufe, S., Treder, M.S., Gugler, M.F., Sagebaum, M., Curio, G., Blankertz, B.: EEG potentials predict upcoming emergency brakings during simulated driving. J. Neural Eng. **8**(5), 056001 (2011)

29. Haufe, S., Kim, J.W., Kim, I.H., Sonnleitner, A., Schrauf, M., Curio, G., Blankertz, B.: Electrophysiology-based detection of emergency braking intention in real-world driving. J. Neural Eng. **11**(5), 056011 (2014)
30. Brunner, C., et al.: BNCI horizon 2020 – towards a roadmap for brain/neural computer interaction. In: Stephanidis, C., Antona, M. (eds.) UAHCI 2014, Part I. LNCS, vol. 8513, pp. 475–486. Springer, Heidelberg (2014)
31. White, R.W., Roth, R.A.: Exploratory search: Beyond the Query-response Paradigm. Synthesis Lectures on Information Concepts, Retrieval and Services. Morgan and Claypool, San Francisco (2009)
32. Hinrichs, U., Schmidt, H., Carpendale, S.: EMDialog: Bringing information visualization into the museum. IEEE Trans. Vis. Comput. Graph. **14**(6), 1181–1188 (2008)
33. Jacucci, G., Morrison, A., Richard, G.T., Kleimola, J., Peltonen, P., Parisi, L., Laitinen, T.: Worlds of information: designing for engagement at a public multi-touch display. In: Proceedings of the 28th International Conference on Human Factors in Computing Systems, pp. 2267–2276. ACM (2010)
34. Coutrix, C., Kuikkaniemi, K., Kurvinen, E., Jacucci, G., Avdouevski, I., Mäkelä, R.: Fizzyvis: designing for playful information browsing on a multitouch public display. In: Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces, vol. 27. ACM (2011)
35. Hearst, M.: Search user interfaces. Cambridge University Press, Cambridge (2009)
36. Bates, M.J.: Where should the person stop and the information search interface start? Inf. Process. Manag. **26**(5), 575–591 (1990)
37. Marchionini, G.: Exploratory search: from finding to understanding. Communications of the ACM **49**(4), 41–46 (2006)
38. Teevan, J., Alvarado, C., Ackerman, M.S., Karger, D.R.: The perfect search engine is not enough: a study of orienteering behavior in directed search. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 415–422. ACM (2004)
39. Tombros, A., Sanderson, M.: Advantages of query biased summaries in information retrieval. In: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 2–10. ACM (1998)
40. Manning, C.D., Raghavan, P., Schütze, H.: Introduction to Information Retrieval, pp. 1–6. Cambridge University Press, Cambridge (2008)
41. Limbu, D.K., Connor, A., Pears, R., MacDonell, S.: Contextual relevance feedback in web information retrieval. In: Proceedings of the 1st International Conference on Information Interaction in Context, pp. 138–143. ACM (2006)
42. Puolamäki, K., Salojärvi, J., Savia, E., Simola, J., Kaski, S.: Combining eye movements and collaborative filtering for proactive information retrieval. In: Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 146–153. ACM (2005)
43. Kelly, D., Teevan, J.: Implicit feedback for inferring user preference: a bibliography. ACM SIGIR Forum **37**(2), 18–28 (2003). ACM
44. Arapakis, I., Konstas, I., Jose, J.M.: Using facial expressions and peripheral physiological signals as implicit indicators of topical relevance. In: Proceedings of the 17th ACM International Conference on Multimedia, pp. 461–470. ACM (2009)
45. Damasio, A.: Descartes' error: Emotion, reason and the human brain. Random House, New York (2008)

46. Pfister, H.R., Bhm, G.: The multiplicity of emotions: A framework of emotional functions in decision making. Judgm. Decis. Mak. **3**(1), 5–17 (2008)
47. Ajanki, A., Hardoon, D.R., Kaski, S., Puolamki, K., Shawe-Taylor, J.: Can eyes reveal interest? Implicit queries from gaze patterns. User Model User-Adapt Interact. **19**(4), 307–339 (2009)
48. Gwizdka, J.: Distribution of cognitive load in web search. J. Am. Soc. Inform. Sci. Technol. **61**(11), 2167–2187 (2010)
49. Gwizdka, J., Cole, M.J.: Inferring Cognitive States from Multimodal Measures in Information Science (2011)
50. Eugster, M.J., Ruotsalo, T., Spapé, M.M., Kosunen, I., Barral, O., Ravaja, N., Jacucci, G., Kaski, S.: Predicting term-relevance from brain signals. In: Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 425–434. ACM (2014)
51. Barral, O., Eugster, M.J., Ruotsalo, T., Spapé, M.M., Kosunen, I., Ravaja, N., Kaski, S., Jacucci, G.: Exploring peripheral physiology as a predictor of perceived relevance in information retrieval. In: Proceedings of the 20th International Conference on Intelligent User Interfaces, pp. 389–399. ACM (2015)
52. Treder, M.S., Schmidt, N.M., Blankertz, B.: Gaze independent brain computer interfaces based on covert attention and feature attention. J. Neural Eng. **8**(6) (2011)
53. Schmidt, N.M., Blankertz, B., Treder, M.S.: Online detection of error-related potentials boosts the performance of mental typewriters. BMC Neurosci. **13**(1), 19 (2012)
54. Pomplun, M., Sunkara, S.: Pupil dilation as an indicator of cognitive workload in human-computer interaction. In: Proceedings of HCI International 2003, vol. 3, pp. 542–546. Erlbaum, Mahwah (2003)
55. Iqbal, S.T., Zheng, X.S., Bailey, B.P.: Task-evoked pupillary response to mental workload in human-computer interaction. In: CHI 2004 Extended Abstracts on Human Factors in Computing Systems, pp. 1477–1480. ACM, April 2004
56. Chen, S., Epps, J.: Using task-induced pupil diameter and blink rate to infer cognitive load. Hum-Comput. Int. **29**(4), 390–413 (2014)
57. Oliveira, F.T., Aula, A., Russell, D.M.: Discriminating the relevance of web search results with measures of pupil size. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2209–2212. ACM (2009)
58. Laeng, B., Sirois, S., Gredebck, G.: Pupillometry a window to the preconscious? Perspect. Psychol. Sci. **7**(1), 18–27 (2012)
59. Lessiter, J., Freeman, J., Miotto, A., Ferrari, E.: Ghosts in the machines: towards a taxonomy of human computer interaction. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 21–31. Springer, Heidelberg (2014)
60. Spagnolli, A., Guardigl, E., Orso, V., Varotto, A., Gamberini, L.: Measuring user acceptance of wearable symbiotic devices: validation study across application scenarios. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 87–98. Springer, Heidelberg (2014)
61. Buenaflor, C.B., Kim, H.C.K.: Six Human Factors to Acceptability of Wearable Computers. IJMUE **8**, 103–114 (2013)
62. Barfield, W., Mann, S., Baird, K., Gemperle, F., Kasabach, C., Stivoric, J., Bauer, M., Martin, R., Cho, G.: Computational clothing and accessories. In: Fundamentals of Wearable Computers and Augmented Reality, pp. 471–509. Erlbaum, Mahwah (2001)

63. Duval, S., Hoareau, C., Hashizume, H.: 7 humanistic needs as seeds. In: Wong, W.K., Guo, Z.X. (eds.) Smart Clothing, pp. 153–187. Taylor and Francis Group, Boca Raton (2014)
64. Baber, C.: Wearable computers: a human factors review. Int. J. Hum-Comput. Int. **13**(2), 123–145 (2001)
65. Rubin, J., Chisnell, D.: Handbook of usability testing: how to plan, design and conduct effective tests. John Wiley and Sons, New York (2008)
66. Gamberini, L., Spagnolli, A., Prontu, L., Furlan, S., Martino, F., Solaz, B.R., Alcañiz, M., Lozano, J.A.: How natural is a natural interface? An evaluation procedure based on action breakdowns. Pers. Ubiquit. Comput. **17**(1), 69–79 (2013)

Live Demonstrator of EEG and Eye-Tracking Input for Disambiguation of Image Search Results

Jan-Eike Golenia(✉), Markus Wenzel, and Benjamin Blankertz

Neurotechnology Group, Technische Universität Berlin, Berlin, Germany
jagoleni@uos.de, {markus.wenzel,benjamin.blankertz}@tu-berlin.de

Abstract. When searching images on the web, users are often confronted with irrelevant results due to ambiguous queries. Consider a search term like *'Bill'*: Results will probably consist of multiple images depicting Bill Clinton, Bill Cosby and money bills. Given that the user is only interested in pictures of money bills, most of the results are irrelevant. We built a demo application that exploits EEG and eye-tracking data for the disambiguation of one of two possible interpretations of an ambiguous search term. The demo exhibits the integration of sensor input into a modern web application.

Keywords: Human-computer interaction · Brain-computer interface · EEG · Eye-tracking · Information retrieval · Free viewing · ERP · Multivariate decoding

1 Introduction

Ambiguous search results are a common problem in web search. Although the user may be more or less clear about what she is searching, the used search query is often ambiguous. Just consider possible results for the query *bank*. Is the user looking for the institution, a bankside or a bench ('bank' is the German word for bench)?

The user may give explicit feedback to resolve such an ambiguity. It would be most straightforward to ask the user whether a result is relevant or not. In addition, the user could rephrase the search query, add keywords or answer questions about preferences. Another approach to resolve ambiguities is gathering implicit information about preferences from previous queries, the browsing history or other data sources. These methods are already used by popular search engines. Google search offers auto-completion for queries. While typing, the user can choose among several suggestions. Search engine providers also use a plethora of other data sources for implicit relevance feedback: The IP (e.g. the map displayed in Google Maps is usually the users's current location), the language settings or information from the user's Google account. While collecting explicit feedback can be costly, implicit feedback usually comes at no extra cost for the user and is unobtrusive.

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 81–86, 2015.
DOI: 10.1007/978-3-319-24917-9_8

A new approach to deliver implicit relevance feedback is gathering data from physiological sensors. This field became more and more of interest to research recently. Eye-movements were used for implicit feedback in image ranking and annotation ([5], [6]), EEG for image retrieval ([7]) and EEG coupled computer-vision for rapid image search ([3]). Combined MEG and eye-tracking data were used for decoding image relevance ([8]), and combined EEG and eye-tracking to search for relevant objects in a 3D environment ([9]).

To demonstrate the feasability of online brain-computer interfacing in a practical application, we built a web application that exploits information from EEG and eye-tracking data for an information seeking task. The application mimicks the process of skimming ambiguous image search results. To create the look and feel of an actual image search application, we used modern web technologies (see section 3). The user is asked to choose one of two possible interpretation of an ambiguous search term. Subsequently, the user is presented with a grid of 24 images, that either belong to one or the other subcategory of the search term. The user inspects in free viewing which of the images belongs to the previously chosen subcategory. When finished observing the results, the application predicts which subcategory the user is looking for based on EEG or eye-tracking data. To investigate how EEG and eye-tracking features can be used to build a model that can predict the subcategory of interest, we conducted an experimental study with ten participants using the demo application. The study can be conceived as research on how to conduct the calibration (model training) of a user. You can calibrate and switch to the online phase within the application. In the online phase, the application yields a prediction after every search result screen.

2 Experimental Study

2.1 Experimental Design

Our aim was to design a demo application that resembles the procedure of an image search application. Flickr, Google image search and the like are examples for such an application: The user types a query and matching images are presented in a mosaic on the screen. The user then proceeds to quickly skim the results for relevant images.

Stimuli. We imitated such a result page (we refer to the result page as *search screen*) by using 24 square pictures, arranged in a grid consisting of 4 rows and 6 columns. Non-square pictures were cropped. Each picture was picked randomly from one of the two subcategories with probability $p = 11/24$ and with probability $p = 2/24$ from the noise picture pool. All stimuli were taken from Flickr (www.flickr.com), a service for sharing pictures aimed at amateur and professional photographers. Flickr provides access via an API to large amounts of high quality pictures that are annotated by users (www.flickr.com/services/api/). The shown pictures are all related to an ambiguous search term (e.g. comb) and one of two subcategories (e.g. honey-bee-bees and chicken-rooster-red) or are picked randomly from a pool of diverse pictures (noise pictures). The experiment consisted of 154 search screens in total. A screen displaying the two subcategories for subcategory choice was presented before the search screen. Ambiguities are rarely due to lexical homonymy, but more often due to underspecified search terms: Image search results for the search term *filter* yield pictures of coffee filters, images processed by different digital filters and analogue filter lenses.

Procedure. Participants were instructed to count at a faster pace rather than a slower and more thorough one. With these instructions we wanted to encourage the subjects to quickly skim the results instead of prioritizing the correct accomplishment of the task. The participants were asked to select one of the two subcategories before the search screen was presented. After choosing, they had to count the pictures belonging to the chosen subcategory, while EEG and eye movements were tracked. When finished with counting, participants had to press space and enter the target image count on the following screen. Subsequently, a screen with feedback was displayed to indicate if the count was correct. See figure 1 for a mockup of the course of a single trial.

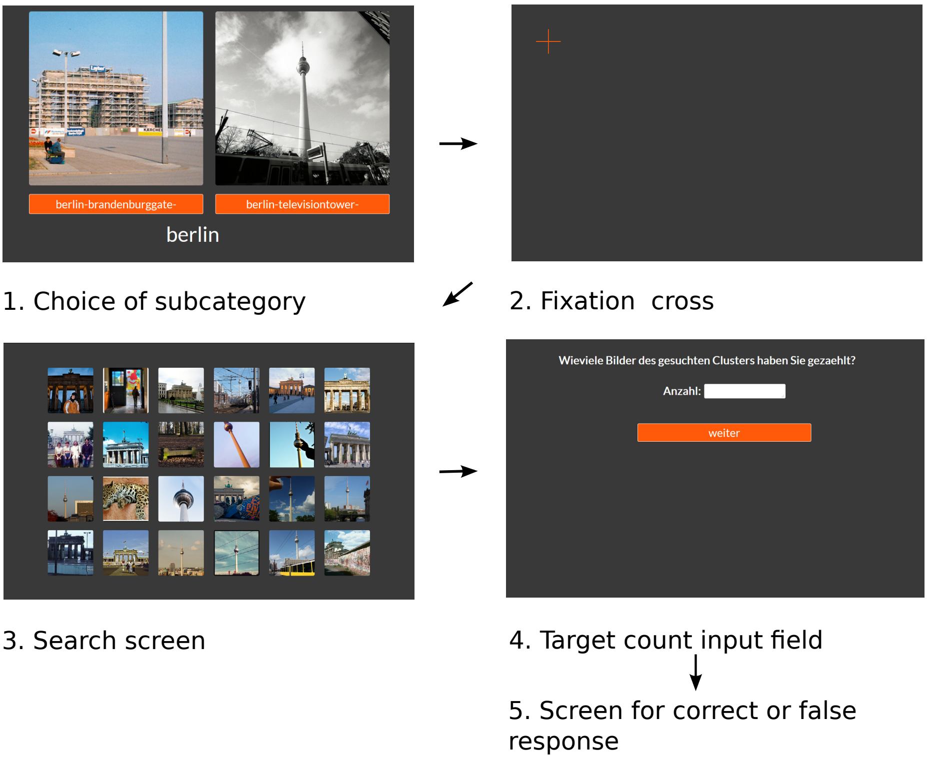


Fig. 1. Mockup of a single trial of the experimental study. The order of screens is enumerated. Screen 1. and 4. were enlarged and cropped for better visibility.

2.2 Analysis

Aim of the analysis was to find eye-tracking and EEG features that can predict the attended target subcategory.

Eye-Tracking. As eye-tracking feature, we used the mean dwell time, the maximal fixation length, the median fixation length and the average fixation count for target and non-target images of each search screen. These four measures were used as features for a LDA classification to predict which of the two subcategories the subject chose to count. The subcategory that the subject attended to and whose images had to be counted are referred to as *target subcategory* and otherwise as *non-target subcategory*. Every subcategory was used as a single data point for training and for each subject a single classifier was trained. We tested the classifier performance by using ten different segmentations of a ten fold cross-validation.

EEG. The EEG data were low-pass filtered with a second order Chebyshev filter (42 Hz passband, 49 Hz stop band), down-sampled to 100 Hz, re-referenced to the linked-mastoids and high-pass filtered with a Butterworth filter at 0.2 Hz.

The continuous multi-channel EEG data were segmented in epochs aligned to the onset of the longest fixation on an image. No baseline was subtracted. Segments aligned to fixations on images are referred to as *target epochs* and *non-target epochs*, depending whether the data were aligned to a fixation on an image belonging to the target or non-target subcategory. We trained a classifier using shrinkage LDA with spatio-temporal features on single epochs (see [1]). The shrinkage parameter was calculated analytically (see [2], [4]). We used the average activity of 50 ms intervals from 0 ms to 1000 ms after fixation onset and 62 channels, resulting in 1240 features. The number of the longest fixations (data points) on either target or non-target images ranged from 3083 to 5408 for a single subject, with slightly unbalanced classes, as target images were fixated more often than non-target images. We tested the classifier performance by using ten different segmentations of a ten fold cross-validation. To classify which subcategory the user attended to, the average of the classifier output over all target and non-target epochs for a single search screen was computed. The classifier output indicates the distance to the decision hyperplane and the sign indicates the class membership. The subcategory with the larger average of outputs was labeled as the target subcategory of a search screen.

2.3 Results and Limitations

Prediction of the chosen subcategory based on EEG and eye-tracking data is possible: Performance of eye-tracking and EEG based predicition is above chance level (50%) for every subject. Mean performance of classification based on eye-tracking features ranged from 67% to 92% correctly classified search screens, and from 62% to 88% correctly classified search screens for EEG features. The EEG based classification is probably confounded by the eye-movements, and results

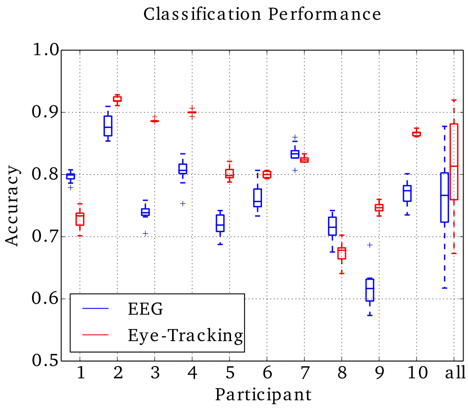


Fig. 2. Every box plot represents ten different segmentations of the data, each data point being the average of a ten fold cross-validation. Measure of performance is accuracy (percentage of correct classifications per search screen).

have to be taken with a grain of salt. This confound will be investigated in the future, as well as whether the two modalities contain complementary information such that data fusion improves the classification performance.

3 Demo Set Up

Hardware. The hardware setup consists of an eye tracker (RED 250, SensoMotoric Instruments, Teltow, Germany; sampling frequency of 250 Hz) attached to a screen (resolution: 1680 x 1050 pixel, size: 47.2 cm x 29.6 cm), and two computers, one for presentation and EEG data acquisition, the other one for eye-tracking data acquisition. Physiological signals are recorded with two amplifiers with 62 active EEG electrodes and one active electrode for electrooculography (EOG) (BrainAmp, ActiCap, BrainProducts, Munich, Germany; sampling frequency of 1000 Hz). Linked mastoids are used for the referencing of the EEG signal.

Software. The presentation of the demo application was implemented using HTML5 for formatting, CSS for styling and JavaScript for client-side code. The demo application is interactive and not a static prearranged sequence of screens. The user can calibrate and validate the eye-tracker inside the browser. Additionally, she can switch the feedback mode on and off whenever she likes. She can also navigate between different menu screens (e.g. a screen for trial selection, a menu for training and selecting a model for feedback) in the browser. The web server handling HTTP requests is an integrated part of Flask, an open source web development framework. Flask also includes a templating language (Jinja2) which combines Python like expressions and HTML to dynamically generate HTML code. For eye-tracking we combined the vendor API (iViewX SDK) with Zero-MQ (a high speed messaging library) to allow for easy cross-platform, cross-language and interprocess communication. We used

Wyrm (github.com/bbci/wyrm) and Mushu for EEG data-analysis and signal acquisition.

4 Conclusion

In this study, we demonstrated that EEG and eye-tracking data can be used to estimate the relevance of images of a user's search. It is possible to resolve image search result ambiguities using this implicit information.

Modern web technologies were used for the implementation of the demo to create the look and feel of an actual web application and to approach the endeavor of combining BCI and web technologies. Technologies that are capable of online prediction, based on eye-tracking and EEG data, and a feedback loop using eye-tracking data were implemented. In combination with the algorithms developed for the prediction of relevance, we set the grounds for a modern web application using implicit relevance prediction.

Acknowledgments. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 611570, and in part by the BMBF (contract 01GQ0850, BFNT).

References

1. Blankertz, B., Lemm, S., Treder, M., Haufe, S., Mller, K.-R.: Single-trial analysis and classification of ERP components a tutorial. NeuroImage **56**, 814–825 (2011)
2. Friedman, J.H.: Regularized discriminant analysis. J. Amer. Statist. Assoc. **84**, 165–175 (1989)
3. Gerson, A.D., Parra, L.D., Sajda, P.: Cortically Coupled Computer-Vision for Rapid Image Search. Neural Systems and Rehabilitation Engineering **14**, 174–179 (2006)
4. Ledoit, O., Wolf, M.: A well-conditioned estimator for large-dimensional covariance matrices. J. of Multivariate Analysis **88**, 365–411 (2004)
5. Hardoon, D.R., Pasupa, K.: Image ranking with implicit feedback from eye movements. In: Proceedings of the 6th Biennial Symposium on Eye Tracking Research & Applications (ETRA), pp. 291–298. ACM, New York (2010)
6. Hajimirza, S.N., Proulx, M.J., Izquierdo, E.: Reading users' minds from their eyes: a method for implicit image annotation. IEEE Transactions on Multimedia, 805–815 (2012). IEEE press, New York
7. Ušćumlić, M., Chavarriaga, R., Milln, J.R.: An Iterative Framework for EEG-based Image Search: Robust Retrieval with Weak Classifiers. PLoS ONE **8**, e72018 (2013)
8. Kauppi, J.-P., Kandemir, M., Saarinen, V.-M., Hirvenkari, L., Parkkonen, L., Klami, A., Hari, R., Kaski, S.: Towards brain-activity-controlled information retrieval: Decoding image relevance from MEG signals. NeuroImage **112**, 288–298 (2015)
9. Jangraw, D.C., Wang, J., Lance, B.J., Chang, S.-F., Sajda, P.: Neurally and ocularly informed graph-based models for searching 3D environments. J. of Neural Engineering **11**, 046003 (2014)

Applying Psychology Research to Shopper Mindsets with Implications for Future Symbiotic Search Systems

Jane Lessiter(✉), Eva Ferrari, Alessia Eletta Coppi, and Jonathan Freeman

Psychology Department, Goldsmiths University of London, London, UK
{J.Lessiter,E.Ferrari,A.Coppi,J.Freeman}@gold.ac.uk

Abstract. Optimising communications to take account of user states is a nascent, huge and growing business opportunity for the retail and advertising worlds. Understanding people's behaviours, thoughts and emotions to different messages in different contexts at different times, can inform the strategic planning and design of systems promoting positive customer experiences and increasing retail sales. Using theory combined with applied insights from our projects, this paper highlights a number of factors (mindset, attention, focus, time pressure and salience) that drive 'search' behaviour in a dynamic retail based environment. The work has implications for developing symbiotic systems.

Keywords: Salience · Perceptual prominence · Open · Closed · Mindset · Symbiosis · Symbiotic · Human-computer · Interaction · Technology-mediated · Affective systems · Attention · Focus · Time pressure · Adaptive · Responsive · Advertising · Retail · Marketing · Shopper · Customer

1 Introduction

Personally adaptive, responsive systems have an increasing presence in the advertising world and are likely to become more symbiotic in the future. Personally incentivised shopping is not new - think store loyalty cards - and even recently, low tech examples still gain media attention. For instance, in the battle of the supermarkets in the UK, Waitrose has recently launched a new scheme for registered loyal shoppers. Loyalty customers are able to choose items on which they receive a 20 per cent discount (Ruddick, 2015). It's an interesting example because it is 'transparent', reliant on full user control and not inferred from behavioural indicators - almost an anti-technology move, particularly where 'trust' in use of personal data and sense of autonomy in decision making still feature high on many customers' ethical agenda.

In the retail world competition to grab customer attention is high, particularly with increased opportunities for engagement and connectivity that new media technologies and digital services afford. Nevertheless, failure to respect customer privacy concerns or getting inferences wrong can damage reputation, particularly given the impact of our social media, and betrays the customer hand that feeds them.

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 87–98, 2015.
DOI: 10.1007/978-3-319-24917-9_9

Research can play a valuable role in understanding how to optimise the effectiveness of advertising by exploring constructs relevant to the shopper ‘mindset’. Drawing from a variety of applied research projects that we have conducted, this paper provides a summary of our current thinking around the retail experience of the current and future consumer and provides a background to our recent experimental work.

This work has implications for the design of future symbiotic systems in the retail space, including systems that support ‘search’ behaviour through the myriad (advertising) messages and information that we seek or are bombarded with on a daily basis.

2 Mindset, Motivation and Volition

2.1 Theory

We assume that retail audiences vary in their ‘mindset’ towards advertisements at any given moment depending on a number of factors. Advertisements can be located at different sites/contexts, which could be indoor, in either public or personal spaces, or outdoor, such as on the High Street. Temporal factors (time of the day, week, month or year) are also important. For instance, consider how ‘porridge’ might be a more or less appealing purchase on a cold winter morning compared with a warm summer evening; or how that luxury item you desire is only attainable after payday or at a special time of the year. Life stage, roles and experiences will also impact on what is considered relevant and meaningful to the individual at different times. We vary in our ‘receptivity’ to a message, defined as “consumers’ conscious and unconscious readiness to accept, process, and respond to brand messaging.” (p.98, Dhar & Kim, 2007). Theoretical perspectives on mindset can be drawn from a range of disciplines such as marketing and consumer psychology, cognitive and social psychology, and the psychology and neurobiology of personality.

Gollwitzer (1990) drawing on the Rubicon model of action, described mindset in terms of the “phase-typical cognitive orientation that promotes task completion” (p.63). In this sense it is part of a motivational process, preparing the individual so that the stimulus presented - in this case, an advertising message - can be analysed resourcefully during the process, leading to successful task completion - product purchase.

Mindset can be viewed as multiple cognitive stages in the identification, pursuit, consummation and evaluation of goal directed action. So how might we understand types of mindset relevant to shopper behaviour?

We consider that every mindset stage during the retail experience is influenced by the relationship between different internal (personal) and external (contextual) conditions. These impact the likelihood of the shopper proceeding efficiently to the next stage toward the goal.

According to Gollwitzer (1990) there are four types of mindset influencing cognitive orientation. The ‘deliberative’ mindset is characterised by making the choice about a goal including realistic prior analysis of the information available and issues. Is the goal desirable and feasible?

Whilst he labels it 'deliberative', suggestive of conscious intention, there are likely unconscious factors at play too, and the properties of stimuli and their relationship to the perceiver are likely key to whether or not a potentially relevant stimulus is noticed, let alone considered or acted upon. A pre-requisite factor to Gollwitzer's 'deliberative' mindset relates to stimulus 'salience': what makes something stand out to one person rather than another, at/in a given time/context rather than another? Is an individual able to identify as well as think about available opportunities?

The second, 'implemental', mindset that Gollwitzer describes is characterised by thoughts about the when, where, and how the goal should be implemented - tasks involved in action initiation. This seems akin to cognitive planning and coordinating/sequencing, important for effective environment-related interaction.

For Gollwitzer, the 'actional' mindset is associated with monitoring the behaviour that is currently in execution towards an identified goal. Finally, the 'evaluative' mindset occurs post action when the individual measures the outcomes reached. Was it worth it? Will it be worth noticing it again in the future?

Whilst these processes can be broken out and studied as conscious reflective subjective experiences, as Gollwitzer has with the implemental and deliberative stages (Gollwitzer, 1990), it is also useful to note that goal directed behaviour is often subtle, fast and some aspects are rather automatic with much less conscious control.

Gollwitzer's ideas also relate to other motivational models of approach behaviour and reward responsivity, describing tasks of motor control, coordination and action relative to the position of the goal object. For instance, neurobiological substrates of these mindsets may have implications for development of future symbiotic systems that are able to identify and optimizing search results based on objective indicators of mindset. There is evidence that part of this motivational process is characterised neurobiologically by activation of the (mesocorticolimbic) dopamine pathway which extends to the prefrontal cortex (e.g., Schultz et al., 1997) - an area involved in reward motivated behaviour (e.g., see review by Balleine & Dickinson, 1998; Frith, Friston, Liddle & Fracowiak, 1992; Bannon & Roth, 1983) - and through regions related to incentive salience (e.g., nucleus accumbens). Berridge (1996) argues that dopamine systems play a primary role in ascribing "incentive salience to selected percepts and representations... which causes it to become attractive and wanted." (p.15).

Ability to satisfy an intention can also be explored in terms of individual differences in sensitivity to 'significant' contextual cues, generated in a complex internal-external interplay. Consider basic drives like hunger and cravings, for instance because of cigarette withdrawal, combined with the availability of conditioned stimuli in the environment (e.g., Powell, Tait & Lessiter, 2002).

Personality traits such as 'impulsivity' are relevant to consumer behaviour in this regard. Such constructs are measurable for instance, with psychometrically validated scales (e.g., EPQ-R: Eysenck & Eysenck, 1991; BIS/BAS scales Carver & White, 1994), and the neurobiological underpinnings and psychophysiological correlates of these traits may have future value in symbiotic systems. The Reinforcement Sensitivity Theory of personality developed through research from the late Jeffrey Gray and his colleagues may have utility in this regard (e.g., see Corr, 2008 for a comprehensive coverage of research exploring the Reinforcement Sensitivity Theory).

Gollwitzer's theory suggests that if shoppers are more concerned with making a purchase and know what they're looking for, they are operating within an action mindset and deliberation is not the primary focus. Gollwitzer found support for his hypothesis that participants placed in deliberative mindsets would show increased receptiveness to the available information compared with those placed in implemental mindsets. This has clear implications for consumer behaviour.

Mindset has also been conceptualised in terms of the predominance of abstract (broad, high level, generic, far time focus) or concrete (narrow, low level, specific, near time focus) cognitive reference frames when consumers are contemplating purchase decisions (Goldsmith, Xu, & Dhar, 2010). The Construal Level Theory (CLT) (e.g., Dhar & Kim, 2007) outlines a decision making stage model (awareness/goal pursuit, consideration-set formation/receptivity, option selection through comparison in context, and post choice affective evaluation) with stages similar to those outlined by Gollwitzer. In this theory, 'progress' or 'commitment' mindsets are also evoked when goal directed behaviour is evaluated. Importantly, CLT extends the notion of temporal distance to embrace 'psychological distance' which accounts for individual variation in perception of temporal proximity to the final goal and subgoals. Dahr & Kim cite research suggesting that an individual's receptivity is optimised under particular combinations of messaging with psychological distance: when psychological distance is decreased, lower level construal claims increase receptivity, and when psychological distance is increased, higher level construal claims increase receptivity. It would be interesting to explore the influence of impulsivity on these relationships.

2.2 Applied Insights: Learning from Retail Customer Experiences in Airports

In one of our projects, a major international airport wanted to explore the types of mindsets of passengers passing through their airport retail spaces with a view to improving the customer experience and increasing sales.

Using observation and interviewing methods, we identified a range of shop design hurdles and barriers to passengers moving with ease through the shopping process. We also identified characteristics important to understanding what types of passengers wanted what products/services and where. For instance, relevance of product availability in different spaces was basic but important. Consider the different needs of host country resident passengers compared with long or short term visitors at arrivals and departures, and those of connecting passengers with variable holding times between flights; the impact of security-related restrictions in the process; and importantly, the fact that the retail element is incidental and not the primary reason for being at an airport. Furthermore, the target sample was also wider than passengers: actual retail footfall comprised a significant number of airport and temporary staff as well as passengers.

In this airport environment, two factors seemed particularly important to driving retail engagement and understanding propensity to respond to advertising/marketing messages.

First, airports can be stressful environments. Indeed research by Credit Card Protection company, CPP, reported in the British Psychological Society news (2011) found that 25 per cent of UK residents found such travel experiences to be as stressful as moving house. In one direction, there is a deadline – 'Boarding, don't miss your flight!' - and in the other, people are wanting to move on, away from the airport and into the host country. There are numerous airport initiatives to improve passenger experience in terms of alleviating stress (e.g., the birdsong installation at Schipol airport: e.g., Winterman, 2013), and research has explored the restorative effects of watching aircraft (Ratcliffe & Freeman, in preparation).

Whilst there are recommended airport arrival times, passengers vary hugely in the times they allow or indeed just find themselves in between their airport arrival and departure. Ultimately there is time pressure to move on, and some people appear more sensitive to this than others. For instance, we found that frequent flyers on business generally allowed very little retail dwell time for themselves. Whilst our research was observational, there is also likely to be a positive correlation between individual sensitivity to anxiety and intended airport dwell time.

Second, it was interesting to watch how some passengers/people in the retail spaces chose to use their available dwell time. Watching them either wander aimlessly along the guided path from security through the glitter of Duty Free, or hurry, searching for travel partners or flight information. A retail experience was high on the agenda for some, part of the 'holiday treat'. For others, it was a definite grab and go: a quick visit to purchase a newspaper and a bottle or water, or grab a few last minute toiletries. Others still lingered and pondered, searching for final gifts for their loved ones. Airport customers varied in their retail 'focus' and our observations (see Figure 1) were consistent with Gollwitzer's notion that deliberative mindsets would show increased receptiveness to the available information compared with those placed in implemental mindsets.

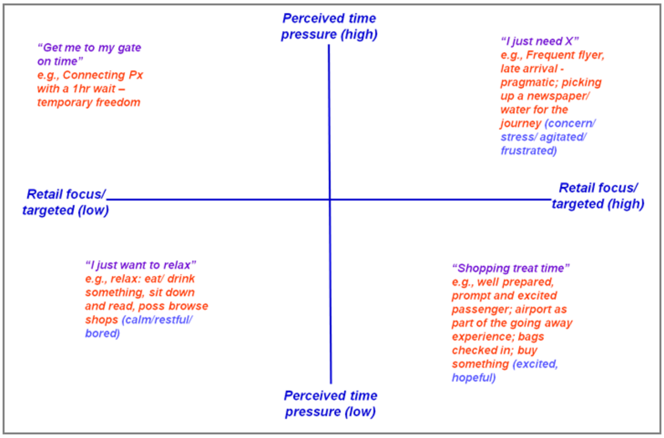


Fig. 1. Time pressure and focus applied to consumers in retail environments

3 Time Pressure

3.1 Theory

Time pressure has been explored in different research areas such as work and stress, decision-making, consumer behaviour, attention and visual search. Time pressure is a subjective judgement, with studies showing mismatches between perceptions of time and actual time (Stroud, 1955).

Hawes (1980) defined time pressure as a complex, subjective experience which varies across people in different ways at different times. Similarly, Iyer (1989) indicated that time pressure is the perceived limitation of time available for a given task. Iyer also noted that time pressure influences the reasoning process given the limitations of our cognitive capacity. Being under time pressure generates difficulties in decision making and research supports this idea, showing that perceived time pressure influences the elaboration of information with repercussions on cognitive processes such as attention and decision-making.

Reutskaja et al. (2011), with an interest in 'cognitive overload' and time pressure on decision making, have used eye tracking and duration of fixations as indicators of speed of search and decision times. They found that participants were able to make judgements in relatively short time periods even with large choice sets, and they made use of different strategies when having to process information within different time periods. This suggests that with increased choice, complexity of stimuli, and little time, consumers must become more selective and brief in what they attend to. Reutskaja et al. suggests that this leaves people open to marketing methods.

Bronner (1982, cited in Iyer, 1989) studied the effect of time pressure on unplanned purchases hypothesising that when the individual is under time pressure the in-store cues to grab attention are suppressed, but when the time pressure is low the opposite effect will occur. This suggests that in a retail context under high time pressure the number of unplanned purchases will be low as will the surfacing of the need evoked from relevant stimuli (such as noticing discounts). Higher physical design salient messaging (media form/presentation) may be required to re-adjust any deficit in attention resulting from a high time pressure mindset.

Other research on unplanned purchases (see Gibridge, Inman, and Stilley, 2013) has found that early in the shopping experience, unplanned purchases suppress other immediate unplanned purchases, suggesting that for high time pressure situations, there may be little opportunity to capitalise on more than one unplanned purchase. However over the course of a trip this pattern reverses which suggests that with longer shopping trips where time pressure is low there are more opportunities for marketers to provoke multiple unplanned purchases.

3.2 Applied Insights: Learning from Customer Experiences of Product Lookalikes

In an example from our applied research, the Intellectual Property Office was keen to understand the impact of product lookalikes and the impact on consumer behaviour and perceptions of consumer harm. Well known brands invest a lot of time and finance into designing their packaging and marketing. Less established competitors can copy some of the physical characteristics of these better known brands, potentially capturing consumer attention because of the similarity and 'fooling' them into a mistaken purchase.

In the first phase, a survey was developed which asked respondents to rate product trios (two own brands relative to each other and to a manufacturer brand) in twelve different product categories. The products, all on the market in 2010, were carefully selected for the study, with advice and feedback about potential examples of lookalikes provided by the British Brands Group and British Retail Consortium. The research team verified and controlled for properties of the product as much as possible in an applied context.

Ratings were given for degree of similarity and also for perceptions of price, quality, suitability for intended use, and value for money. Multiple versions of the survey were developed which varied the position of the manufacturer brand relative to own brands in the trios. The surveys were made available online via surveymonkey.com to UK consumers in early 2012. The phase 1 survey received 1160 product trio question set responses from 330 respondents.

We found that some own brand products whose packaging is perceived by respondents to look more like that of a manufacturer brand for which it could be a substitute, appear to gain significant advantage in being of higher perceived quality (quality, suitability for intended use, expensiveness, value for money) over another own brand product whose packaging looks less like that of a manufacturer brand. The lookalike effects were evident in responses of users and non-users of the relevant product categories, although the effect seemed to be slightly more evident for non-users of a product category.

In a second survey using nationally representative online research panels, respondents from the UK (n=1000), Germany (n=500) and America (n=500) were asked directly about whether they were aware of having purchased a lookalike product, accidentally or deliberately, and whether by doing so they considered they had been advantaged or disadvantaged. A majority of respondents (50+ per cent) reported purchasing lookalike products accidentally at least once. Up to one quarter of the sample had done so a few times. It was interesting to note that as many respondents claimed it had advantaged as disadvantaged them. The majority of each national sub-sample also claimed to purchase lookalikes on purpose, at least once, particularly in Germany. The vast majority, around 60-75 per cent of each national sub sample, reported that doing so was advantageous rather than disadvantageous to them. Participants reported having previous positive experiences with the product which incentivised repeat behaviour, and rated value for money rather than highest quality being a key characteristic in decision making.

The research revealed that results varied by product category suggesting other influencing variables. For instance, some product categories have been subject to stronger innovation (e.g., razors) than others (e.g., vinegar). Further, some high street retailer own brands have become trusted brands in themselves. The research was unable to control for numerous other factors that would clearly influence noticeability of products (e.g., specifics in font size and other physical ‘design salience’ properties) but acknowledged the importance of these influences including shelf placements in terms of prime positioning and how positioning relative to the manufacturer brand may increase/decrease the lookalike effect. The reader is referred to Johnson, Gibson, and Freeman, 2013 (in particular, Chapter 7 and Appendix D) for details of the methodology and statistical analysis for both surveys.

4 Focus and Attention

4.1 Theory

Facoetti and Molteni (2000) describe focus as a process of task-oriented concentrating that supports the selection of relevant from irrelevant stimuli in the field of information. As a consumer, when we know what we want to buy, there is a sharpening of attentional resources.

In early research William James (1890) suggested that visual attention is composed of a focus, a margin and a fringe. Nearly a century later Posner (1980) proposed a metaphor in which attention was compared to an internal spotlight. The author tested the hypothesis by asking participants to click a button when a light appeared in their visual field. The light could occurred in different places of the field and was pre-cued by an arrow (pointing in the direction of the target) or by an illumination of the field and enhanced the efficiency of detection during the task. Research now over many years has explored how subtle or blatant a cue is needed to grab attention. Perhaps unsurprisingly there is a wealth of mixed evidence showing the effectiveness of different cues to influence conscious and unconscious information processing and decision making.

Eriksen and Rohrbaugh (1970) and Eriksen and Hoffman (1972) noted that the attention process is similar to a zoom lens. In this model attention is considered a process with limited capacity and the process of focusing is controlled by attention that works as a zoom lens that is set over a field of stimuli. When the focus is broad the lens has low power and low resolution of the details of the stimuli, but when the resources on focus become narrowed, the stimuli became more detailed if they are detectable in the visual field. The process seems to be highly functional and when the individual needs higher resolution for a task the zoom shifts from broad to narrow to screen out irrelevant stimuli and help focalization on relevant ones.

Nideffer (1976) proposed the attentional and interpersonal style theory to understand and predict conditions under which individuals would be able to perform up to their potential. The theory, relevant to both physical (motor skill) and mental (decision making, problem solving) performance, postulates that focus of attention moves along a width dimension (broad - narrow) and directional dimension (internal - external). These two dimensions intersect and result in four attentional styles (Nideffer, 1976): the broad-external focus is used when reactivity and awareness of the environment is needed; the broad-internal is used to analyse and concentrate or planning; the narrow-internal is used for an inward check up and the narrow external when it’s necessary to perform specific physical and interpersonal tasks.

When workload increases, resources are allocated to elaborate the information, but if the load is too high information is less optimally processed. The higher the cognitive demand the greater the effort required to sustain a level of efficacy (Grier et al. 2003; Kahneman 1973). This suggests that determined shoppers with too much choice are susceptible to making less effective decisions.

Given that much consumer activity is now online, the sense of presence - a sense of being there in a mediated environment; a state that results from attending to and evaluating incoming sensory information (Barfield et al., 1995) - is relevant to this discussion. The manipulation or measurement of presence (e.g., Lessiter, Freeman, Keogh and Davidoff, 2001) may have important applications in online retail space with regard to influencing consumer attention, focus, engagement, and proximity in the context of consumer mindsets. It also raises theoretical questions around manipulations of mindset using task based instructions (for instance, what are the cognitive/affective experiential differences in 'being there' compared with 'doing there'). The relationships between user experience of system usability, presence and different mindsets is worthy of further exploration.

4.2 Applied Insights: Validation of Focus Manipulation

In one of our collaborative projects (see: mindsee.eu), the EEG correlates of high and low focus was explored (manuscript in preparation). Task instructions were used to experimentally manipulate participants' focus. The instructions provided information about the search task with variation in the degree of specificity/ambiguity about the search target. Before running the study, we wanted to conduct a pilot study to validate that the instructions we generated for the task influenced participant mindset, as expected, in terms of subjective focus.

In search processes, we considered that focus would vary in relation to the level of definition of the task at hand, the ambiguity of the search target and the cognitive states associated with these (e.g., uncertainty, confidence). We considered that the individual is in a high focus state when the target of the search task is more defined, more specific and therefore less ambiguous, making the user feel less uncertain and more confident. For instance, if I am in a high focus mindset as a shopper, I know exactly what purchase I want and where to buy it.

In contrast, the user is in a low focus state when the target of the search task is less defined, less specific and therefore more ambiguous, making the user feeling more uncertain and less confident than those in the high focus condition. States of uncertainty are typical of vague searches, such as those performed at the beginning of a search process when the object of the search is not completely defined (Kuhlthau, 2004); the uncertainty is reduced once the object of the search becomes clearer (high focus).

We hypothesised that participants in high and low focus conditions would perceive the task and the instructions differently in terms of: task specificity (narrow vs. broad), cognitive states associated with the task (e.g., confidence, uncertainty), ambiguity (defined vs. undefined target), and task perceived difficulty (high vs. low).

Using a repeated measures design, 12 participants each completed two search tasks (identifying and counting the number of times (a) a particular 'shape' and (b) a particular pattern, was present in the visual task). For each search task, participants were given a different instruction: one aimed to create a high focus mindset, and the other a low focus mindset. Mindset order (low/high focus) was counterbalanced through the sample as was the type of search target (shape/pattern).

At the end of each visual search task, participants were presented with a series of questions about the task (performing it; the instructions; the search target) and asked to indicate, using a 7 point scale, which of a pair of contrasting attributes they most erred towards. As hypothesised, the results revealed significant differences between the mindsets, for instance, compared with high focus, participants in low focus mindsets reported feeling significantly less sure, more hesitant in their task performance, and found the instructions to be more ambiguous. The results indicated that the task instructions effectively manipulated mindset in a valid, predictable way.

5 Other Research Applications

An outdoor advertising agency wanted to further understand their earlier qualitative findings that consumer mindset influences receptivity to outdoor advertising messages. We were keen to explore whether we could effectively manipulate mindset using the dimensions of time pressure and focus to understand what would happen to participants' ability to recognise 'distractors' (which could be advertising cues).

Initially we designed a lab based experimental study to manipulate these constructs on an abstract rather than applied level. The work required clear conceptualisation and an understanding of what we meant by focus and time pressure, and what might be useful to remember/recognise about adverts in a shopping context. This initial study has since lead to a series of research projects in both commercial and academic contexts including the a salience scale (in development), and exploration of the physiological correlates of different mindsets using measures from EEG (the pilot for the task instructions, summarised in 4.2 above) and eyetracking tools. These studies are currently in preparation for publication and will be described in detail elsewhere.

We described above how the construct of focus was operationalised for empirical study. For the time pressure variable, we similarly used task instructions as the manipulation. Participants were told that the task had a fixed limited time period, or not, with respectively appropriate encouragement. Some of the studies we have conducted have evaluated correlates of mindset, and others have explored people's receptivity to distractor cues following the primary task, testing their recognition of the 'incidental' objects/adverts.

In our most recent commissioned project, we are exploring the impact of personal salience, familiarity (product category user/nonuser) and advert design salience (high/low) on a range of subjective recall, recognition and eye tracking measures to understand if we can predict effectiveness of advertising to different targets audiences.

6 Discussion

This paper (supported by the EC through the Mindsee project - G.A. #611570) has highlighted the potential application of a range of psychology related theories and constructs to better understand consumer search behaviour, specifically in retail/advertising contexts. The particular constructs of interest were mindset, time pressure, focus/attention and salience. Following a more academic outline of relevant theories, applied examples from our research commissions and other projects were provided. Whilst the focus here was largely on retail customer experiences, the work also has extended relevance to the development of various future search tools that understand the limitations of human information processing and work to symbiotically support users make more efficient, effective and advantageous personal decisions. This could include symbiotic applications that are sympathetic to the susceptibility to exploitation of our individual mindsets at different times, and that support personal control in decision making, by revealing how we might be being influenced by unwanted marketing.

References

1. Ruddick, G.: Waitrose to let shoppers pick discounts in 'game changing' move. The Telegraph. www.telegraph.co.uk/finance/newsbysector/retailandconsumer/11681363/Waitrose-to-let-shoppers-pick-discounts-in-game-changing-move.html, June 18, 2015
2. Dhar, R., Kim, E.Y.: Seeing the forest or the trees: Implications of construal level theory for consumer choice. Journal of Consumer Psychology **17**(2), 96–100 (2007)
3. Gollwitzer, P.M.: Chapter 2. Action phases and mind-sets. In: Higgins, E.T., Sorrentino, R.M., (eds.) Handbook of Motivation and Cognition. Foundations of Social Behavior, vol. 2. The Guilford Press (1990). https://psych.nyu.edu/gollwitzer/90Goll_ActionPhasesMindSets.pdf
4. Schultz, W., Dayan, P., Montague, P.R.: A neural substrate of prediction and reward. Science **275**, 15931599 (1997)
5. Balleine, B.W., Dickinson, A.: Goal directed instrumental action: Contingency and incentive learning and their cortical substrates. Neuropsychopharmacology **37**, 407–419 (1998)
6. Frith, C.D., Friston, K.J., Liddle, P.F., Fracowiak, R.S.J.: PET imaging and cognition in schiziphrenia. Journal of the Royal Society of Medicine **85**, 222–224 (1992)
7. Bannon, M.J., Roth, R.H.: Pharmacology of mesocortical dopamine neurons. Pharmacological Reviews **35**(1), 53–68 (1983)
8. Berridge, K.C.: Food reward: Brain substrates of wanting and liking. Neuroscience and Biobehavioural Reviews **20**(1), 1–25 (1996)
9. Powell, J., Tait, S., Lessiter, J.: Cigarette smoking and attention to signals of reward and threat in the Stroop paradigm. Addiction **97**(9), 1163–1170 (2002)
10. Eysenck, H.J., Eysenck, S.B.J.: Manual of the Eysenck Personality Scales. Hodder and Stoughton (1991)
11. Carver, C.S., White, T.L.: Behavioural inhibition, behavioural activation, and affective responses to impending reward and punishment: The BIS/BAS scales. Journal of Personality and Social Psychology **67**, 319–333 (1994)
12. Corr, P.J.: The Reinforcement Sensitivity Theory of Personality. Cambridge University Press, Cambridge (2008)

13. Goldsmith, K., Xu, J., Dhar, R.: The power of customers' mindsets. Sloan Management Review **52**(1), 19–20 (2010)
14. British Psychological Society. Why we find airports so stressful, July 01, 2011. http://www.bps.org.uk/news/why-we-find-airports-so-stressful
15. Ratcliffe, E., Freeman, J.: Aircraft views in airport environments: Affective appraisals and relationships with attention restoration (in preparation)
16. Winterman, D.: The surprising uses for birdsong. BBC News Magazine, May 8, 2013. http://www.bbc.co.uk/news/magazine-22298779
17. Stroud, J.: The fine structure of psychological time. In: Quastler, H. (ed.) Information Theory in Psychology: Problems and Methods, pp. 174–207. Free Press, Glencoe (1955)
18. Hawes, D.K.: The time variable in models of consumer behavior. Advances in Consumer Research **7**(1), 442–447 (1980)
19. Reutskaja, E., Nagel, R., Camerer, C., Rangel, A.: Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study. American Economic Review **101**(2), 900–926 (2011). http://www.rnl.caltech.edu/publications/pdf/reutskaja2011.pdf
20. Gibridge, T.J., Inman, J.J., Stilley, K.M.: What determines unplanned purchases? A model including shopper purchase history and within-trip dynamics (2013). https://marketing.wharton.upenn.edu/mktg/assets/File/Dynamic%20Effects%20in%20Unplanned%20Purchase%20Behavior%202-18-13.pdf
21. Johnson, P., Gibson, J., Freeman, J.: The Impact of Lookalikes: Similar packaging and fast-moving consumer goods. Research Report for the Intellectual Property Office (2013). http://webarchive.nationalarchives.gov.uk/20140320154249/http://www.ipo.gov.uk/ipresearch-looklikes-310513.pdf
22. Facoetti, A., Molteni, M.: Is attentional focusing an inhibitory process at distractor location? Cognitive Brain Research **10**, 185–188 (2000). http://decone.psy.unipd.it/De.Co.Ne_LAB_Unipd/A._Facoetti_files/Cogn%20Brain%20Res%20Facoetti%20%26%20Molteni%202001.pdf
23. James, W.: The Principles of Psychology, vol. 1. Henry Holt and Co., New York (1890)
24. Posner, M.: Orienting of attention. Quarterly Journal of Experimental Psychology **32**, 3–25 (1980)
25. Eriksen, C., Rohrbaugh, J.: Some factors determining efficiency of selective attention. The American Journal of Psychology **83**(3), 330–342 (1970)
26. Eriksen, C., Hoffman, J.: Temporal and spatial characteristics of selective encoding from visual displays. Perception and Psychophysics **12**(2B), 201–204 (1972)
27. Nideffer, M.: Test of attentional and interpersonal style. Journal of Personality and Social Psychology **34**(3), 394–404 (1976)
28. Grier, R.A., Warm, J.S., Dember, W.M., Matthews, G., Galinski, T.L., Szalma, J.L., Parasuraman, R.: The Vigilance Decrement Reflects Limitations in Effortful Attention. Not Mindlessness. Human Factors **45**(3), 349–359 (2003)
29. Kahneman, D.: Attention and Effort. Prentice-Hall, Inc., Englewood Cliffs (1973)
30. Barfield, W., Zeltzer, D., Sheridan, T.B., Slater, M.: Presence and performance within virtual environments. In: Barfield, W., Furness, T.A. (eds.) Virtual Environments and Advanced Interface Design, pp. 473–541. Oxford University Press, Oxford (1995)
31. Lessiter, J., Freeman, J., Keogh, E., Davidoff, J.: A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. Presence Teleoperators and Virtual Environments **10**(3), 282–297 (2001)
32. Kuhlthau, C.C.: Seeking Meaning: A Process Approach to Library and Information Services, 2nd edn. Libraries Unlimited, Westport, CT (2004)

Symbiotic Interaction and the Experience of Agency

Hannah Limerick[1(✉)], James W. Moore[1,2], and David Coyle[1,3]

[1] University of Bristol, Bristol, UK
hannah.limerick@bristol.ac.uk
[2] Goldsmiths, University of London, London, UK
j.moore@gold.ac.uk
[3] University College Dublin, Dublin, Ireland
d.coyle@ucd.ie

Abstract. The sense of agency is the experience of initiating our actions in order to control the external environment. This paper explores the notion of the self and agency in the context of symbiotic computing. Symbiotic computer systems use sensors to detect psychophysiological markers and implicitly interpret our intentions in order to enhance our actions in some way. Maintaining agency independence between the user and the system is central to the symbiotic relationship. Also it is ideal to enhance the user's interaction without subducting control and therefore it is pertinent to consider the sense of agency during symbiotic interactions. This paper will theoretically explore the notion of self-agency in a symbiotic setting, drawing on relevant research into the sense of agency in psychology and neuroscience.

Keywords: Sense of agency · Symbiotic systems · Human computer interaction · User experience · Physiological interaction · Agency

1 Introduction

The sense of agency is the experience of initiating our actions in order to control the external environment. This feeling of being the agent of our actions allows us to know "*I did that*" when we've made actions. It has long been recognized that supporting the user's sense of agency is a key principle in the design of user interfaces [1]. A hitherto unexplored challenge to this principle comes from symbiotic interaction, a new and interdisciplinary field of computer science.

Symbiotic interaction comprises of a computer component that has the ability to implicitly (or subliminally) detect (via sensors) the user's psychophysiological state [3]. This information can then be used to "better adapt output regardless of his/her ability to explicitly refine his/her request" [3]. With this, such systems can use psychophysiological cues to enhance the human computer interaction in new and exciting ways. A simple example is a system that uses psychophysiological data to detect when the user is tired and therefore provides assistance

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 99–104, 2015.
DOI: 10.1007/978-3-319-24917-9_10

to perform a task. However whilst symbiotic systems such as these have great promise, there is a risk that the user's sense of agency is diminished, with the computer automatically responding to the user's state to facilitate task completion. This could have serious ramifications given the importance of sense of agency for our interactions with computers. This has been recognized within the field of physiological interaction for example "*use of a physiological computing system may blur the perception of self or act as an unwanted source of interference on self-perception. This 'splitting' of self-perception is certainly plausible but difficult to evaluate or address at the current time*"[2]. This highlights a demand to evaluate and address the impact on sense of agency that may occur with psychophysiological interaction techniques. Here we first discuss theories and measures of sense of agency, drawing on the rapidly growing cognitive neuroscience literature on the topic. We then explore some of the challenges to sense of agency posed by symbiotic interaction and how we can go about addressing them.

2 Theories of Sense of Agency

An important consideration regarding the sense of agency and human-computer-interaction (HCI) is how the experience comes about. Here, the neurocognitive processes underlying the experience provide valuable insight. Multidisciplinary research currently paints an intricate picture, where various agency cues and indicators feed into the experience of agency [6–9]. Agency cues comprise of external situational information surrounding an action. These external cues can modulate beliefs about agency, for example [4] demonstrated that words which served to prime thoughts prior to an action led participants to experience agency for actions that they were forced to make. Moreover internal sensorimotor cues such as the predicted sensory consequences of movement are also thought to be agency cues (e.g. [5]). Distinct modes of interaction provide a vast range of agency cues that also differ in extent. Thus the agency cues surrounding HCI provide varying experiences of agency (for a review, see [11]).

2.1 Methods of Measuring Sense of Agency

A phenomenological distinction has been made between the Judgement of Agency and the Feeling of Agency [12]. The implicit feeling of agency refers to the pre-reflective, low-level feeling of being the agent of an action. The explicit judgement of agency refers to the attribution of agency to oneself or another on a conceptual level. Researchers have developed several ways of measuring the impact that various agency cues have on the sense of agency experimentally. The explicit judgement of agency is typically measured through verbal report by asking participants to rate their feeling of agency during a task or simply state whether they were the agent or not.

Measures have also been developed to probe implicit aspects of sense of agency. Haggard and colleagues [10] found that voluntary actions and their outcomes result in measureable changes in the perceived timings of these events.

Here, the action and the outcome are perceived as closer together in time (Figure 1). In the case of involuntary actions the perceived temporal interval was found to be longer than the actual interval. This temporal phenomenon is known as 'intentional binding', and taken to be an implicit metric for the sense agency.

Intentional binding is assessed through the use of the so-called 'Libet Clock' [13]. This clock is presented on a computer screen. When the trial starts, a clock hand rotates clockwise around the clock at a speed of one rotation every 2.56s. Participants are instructed to make a self-paced action, which causes an outcome after a fixed 250ms time interval. In the original study the action was a button-press and the outcome was an auditory tone. Participants are asked to report where the clock hand is pointing when the critical event occurs. The critical event refers to the action or outcome depending on the trial condition. There are four trial conditions that enable the intentional binding calculation to be made. There are two baseline conditions where the action or outcome occurs in isolation and there is no causal link between the two events. There are also two operant conditions, where there is a causal link between the action and outcome and participants report the time of the action or the time of the resultant tone (depending on the condition). This intentional binding effect has been widely replicated and has led to considerable advances in our understanding of the sense of agency (see [14] for a review of intentional binding). Measurements of the sense of agency provide insight into the amount of control and volition a user feels under different conditions. Intentional binding is particularly useful in HCI settings because it provides a measure of the degree of control experienced by the user. This is not the case with explicit measures requiring a simple 'yes' or 'no' response.

3 The Experience of Agency in a Symbiotic Interaction

The definition of the symbiotic relationship states that it is "characterized by goals and agency independence of humans and computers" [3] so one challenge is to establish agency independence during the interaction. Another key question for symbiotic computing is how the experience of agency is modulated for actions that are assisted/mediated by technology and therefore not fully their own. Here we explore existing research relevant to addressing some of the open questions in symbiotic interaction pertaining to the sense of agency.

Agency Independence. This refers to the user and the system both being separate and independent agents during the interaction. An individual's beliefs regarding the agent of an action is highly influential to an agentic experience. In one study, prior causal beliefs about the agent of an action led participants to experience less implicit sense of agency for self-generated actions that they believed to be caused by another agent [19]. [4] found that a mistaken belief about intentionally causing an action could be induced by simply priming individuals to think about the action just before it occurred. Visual feedback is

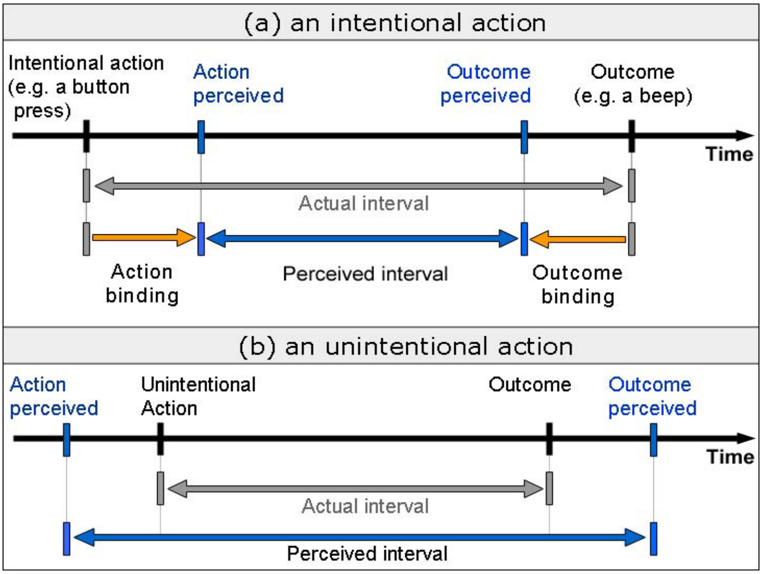


Fig. 1. Intentional Binding. (a) For voluntary actions the perceived time interval between action and outcome is shorter than the actual delay. (b) During involuntary actions the perceived time interval between action and outcome is longer than the actual delay

also an important agency cue which can be manipulated to alter the sense of agency. [17] found that deviations in the visual feedback of a moving curser associated with joystick movement beyond 50° led participants to explicitly attribute their movements to another agent irrespective of their implicit sensorimotor movements. These studies demonstrate that the sense of agency is malleable and that agency cues can counteract the sense of agency. This evidence suggests that in order to establish clear explicit agency independence, cues denoting the agency (or lack of) for the system and the users are crucial. Agency metrics discussed above may be useful in establishing agency independence and deepening the symbiotic relationship.

Assistance. Technological assistance introduces ambiguity to the notion of voluntary or involuntary actions. Whilst this question is philosophical, it is also central to effective interface design. Furthermore, maintaining the user's autonomy is key to other aspects of the interaction, such as motivation in computer games [16]. [18] explored the sense of agency for tasks in a flight control deck with varying degrees of computer automation. Intentional binding measures indicated that increasing automation led participants to feel less sense of agency for the outcome of their actions. Similarly, [15] used the intentional binding metric to investigate the sense of agency and computer assistance during a more familiar task - computer mouse movements. In this task participants used a mouse to select their choice of target on a screen. An algorithm interpreted the intentions of the users and assisted their movement towards the target accordingly. The results indicated that increasing computer assistance led participants to experience less agency for their actions. Moreover the results suggested that there is a point at which assistance could be provided to the user and they would still exhibit intentional binding. Beyond this point of assistance intentional binding breaks down. This presents another opportunity for agency metrics to be used in the development of a symbiotic system in order to calibrate the amount of assistance so that it is explicit to the user and thus establish clear agency independence. Furthermore, this calibration could enable assistance to go unnoticed and implicit if required by the application domain.

Calibration. Neurocognitive research highlights that the sense of agency differs between individuals and therefore sense of agency metrics may also prove useful for calibrating a system for individuals. The sense of agency is thought to be central to mental health disorders which are characterized by delusions of control such as schizophrenia and psychosis [21]. Other factors have been found to modulate the sense of agency, such as cognitive load [22] and physical effort [23]. Therefore depending on the situation people may experience the sense of agency differently. Therefore symbiotic systems designed to assist individuals would benefit from fine-tuning a system based on the agency requirements of the target group. For symbiotic computing a system with a ‘one-size fits all’ strategy may not work effectively. Therefore implicit metrics such as intentional binding may additionally be useful to periodically calibrate a symbiotic system.

4 Conclusion

In this paper we have explored symbiotic computing systems and the experience of agency in the context of assistance and agency independence. This is an important consideration for the implementation of an effective interaction. We conclude that the agency cues provided to the user during the interaction may be crucial in achieving agency independence. Another concept that requires reflection is where actions are being assisted by a symbiotic system. Assistance in HCI settings has been found to reduce the sense of agency [15]. Moreover, we have introduced an implicit metric for the sense of agency - intentional binding - which we propose to be useful addressing some of the challenges in this domain. Intentional binding is based on a temporal phenomenon, which is well known to correlate with volition and control.

References

1. Shneiderman, B., Plaisant, C.: Designing the User Interface: Strategies for Effective Human-Computer Interaction, 4th edn. Pearson Addison-Wesley, Reading (2004)
2. Fairclough, S.H.: Fundamentals of physiological computing. Interacting with Computers **21**(1), 133–145 (2009)

3. Jacucci, G., Spagnolli, A., Freeman, J., Gamberini, L.: Symbiotic interaction: a critical definition and comparison to other human-computer paradigms. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 3–20. Springer, Heidelberg (2014)
4. Wegner, D., Wheatley, T.: Apparent mental causation: Sources of the experience of will. American Psychologist **54**(7), 480 (1999)
5. Blakemore, S., Wolpert, D., Frith, C.: Abnormalities in the awareness of action. Trends Cogn. Sci. **6**, 237–242 (2002)
6. Wegner, D.M., Sparrow, B.: Authorship processing. In: Gazzaniga, M. (ed.) The Cognitive Neurosciences III, pp. 1201–1209. MIT Press, Cambridge (2004)
7. Wegner, D., Sparrow, B., Winerman, L.: Vicarious agency: experiencing control over the movements of others. J. Pers. Soc. Psychol. **86**, 838–848 (2004)
8. Moore, J.W., Wegner, D.M., Haggard, P.: Modulating the sense of agency with external cues. Conscious. Cogn. **18**, 1056–1064 (2009)
9. Moore, J., Fletcher, P.: Sense of agency in health and disease: a review of cue integration approaches. Conscious. Cogn. **21**, 59–68 (2012)
10. Haggard, P., Clark, S., Kalogeras, J.: Voluntary action and conscious awareness. Nat. Neurosci. **5**, 382–385 (2002)
11. Limerick, H., Coyle, D., Moore, J.W.: The experience of agency in human-computer interactions: a review. Front. Hum. Neurosci. **8**, 643 (2014)
12. Synofzik, M., Vosgerau, G., Newen, A.: Beyond the comparator model: a multifactorial two-step account of agency. Conscious. Cogn. **17**, 219–239 (2008)
13. Libet, B., Gleason, C.A., Wright, E.W., Pearl, D.K.: Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). Brain **106**(3), 623–642 (1983)
14. Moore, J.W., Obhi, S.S.: Intentional binding and the sense of agency: a review. Consciousness and Cognition **21**(1), 546–561 (2012)
15. Coyle, D., Moore, J., Kristensson, P.O., Blackwell, A.F., Fletcher, P.C.: I did that! measuring users experience of agency in their own actions. In: ACM Conference on Human Factors in Computing Systems, CHI, pp. 2025–2034 (2012)
16. Ryan, R.M., Rigby, C.S., Przybylski, A.: The motivational pull of video games: A self-determination theory approach. Motivation and Emotion **30**(4), 344–360 (2006)
17. Farrer, C., Bouchereau, M., Jeannerod, M., Franck, N.: Effect of distorted visual feedback on the sense of agency. Behav. Neurol. **19**, 53–57 (2008)
18. Berberian, B., Sarrazin, J.C., Le Blaye, P., Haggard, P.: Automation technology and sense of control: a window on human agency. PLoS One (2012)
19. Desantis, A., Roussel, C., Waszak, F.: On the influence of causal beliefs on the feeling of agency. Conscious. Cogn. **20**, 1211–1220 (2011)
20. Negri, P., Gamberini, L., Cutini, S.: A review of the research on subliminal techniques for implicit interaction in symbiotic systems. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 47–58. Springer, Heidelberg (2014)
21. Frith, C.: The Cognitive Neuropsychology of Schizophrenia. Lawrence Erlbaum Associates, Hove (1992)
22. Hon, N., Poh, J.H., Soon, C.S.: Preoccupied minds feel less control: Sense of agency is modulated by cognitive load. Consciousness and Cognition **22**(2), 556–561 (2013)
23. Demanet, J., Muhle-Karbe, P.S., Lynn, M.T., Blotenberg, I., Brass, M.: Power to the will: how exerting physical effort boosts the sense of agency. Cognition **129**(3), 574–578 (2013)

Toward the Development of a Neuro-Controlled Bidirectional Hand Prosthesis

Silvestro Micera[1], Jacopo Carpaneto[1(✉)], Stanisa Raspopovic[1], Giuseppe Granata[2], Alberto Mazzoni[1], Calogero M. Oddo[1], Christian Cipriani[1], Thomas Stieglitz[3], Matthias Mueller[3], Xavier Navarro[4], Jaume del Valle[4], Hans Scherberger[5], Luigi Raffo[6], Massimo Barbaro[6], Danilo Pani[6], and Paolo M. Rossini[2]

[1] Scuola Superiore Sant'Anna, Pisa, Italy
{s.micera,j.carpaneto}@sssup.it
[2] Catholic University of the Sacred Heart, Rome, Italy
[3] University of Freiburg, Freiburg, Germany
[4] Universitad Autonoma de Barcelona, Barcelona, Spain
[5] German Primate Center, Göttingen, Germany
[6] University of Cagliari, Cagliari, Italy

Abstract. The hand is a powerful tool and its loss causes severe physical and often mental debilitation. Surveys on artificial hands reveal that 30 to 50% amputees do not use their prosthetic hand regularly, due to its low functionality. The fundamental issue is therefore to improve the voluntarily-controlled dexterity to allow amputee to perform tasks that are necessary for activities of daily living and that cannot yet be done with the state-of-the-art artificial limbs. The NEBIAS project, launched at the start of November 2013, aims at developing and clinically evaluating a neuro-controlled upper limb prosthesis intuitively controlled and felt by the amputee as the natural one.

Keywords: Neural engineering · Neural prosthetics · Bionics · Artificial limbs · Neural interfaces

1 Introduction

Amputation is a traumatic event, which changes forever the life of the person who suffers it in a quite dramatic way. The amputee requires an active prosthetic device to perform several activities of daily living and in particular grasping and manipulation functions, which are significantly affected after hand amputations. However, the abandonment rate of currently myoelectric prostheses in favour of body-powered or cosmetic ones is still very high [2]. The main reasons of this tendency have to be searched in the weight, in the limited dexterity of the hand prosthesis, and in the complete absence of sensory feedback due to the lack of rich sensations naturally perceived when grasping an object [2]. Ideal bidirectional

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 105–110, 2015.
DOI: 10.1007/978-3-319-24917-9_11

hand prostheses should involve both a reliable decoding of amputee's intentions and the delivery of sensory feedback through the residual afferent pathways, simultaneously and in real time [6]. Starting from previous encouraging results [9,10] and in the framework of a EU founded research area (i.e., FP7-FET Proactive Evolving Living Technologies), the NEBIAS project (NEurocontrolled BIdirectional Artificial upper limb and hand prosthesiS, www.nebias-project.eu) aims at providing an effective solution to improve the quality of life of people who suffered a hand amputation by developing a novel generation of bionic hand prostheses.

2 Architecture

The final demonstrator of the NEBIAS project will be composed of the following modules (see Figure 1):

- implantable electrodes able to selectively interface the peripheral nervous system;
- embedded electronics for recording, processing, and stimulation wirelessly connected with the electrodes;
- artificial upper limb and dexterous hand prostheses endowed with a neuromorphic tactile and kinaesthetic sensory system;
- decoding and encoding algorithms to develop the bi-directional link between the nervous system and the artificial device.

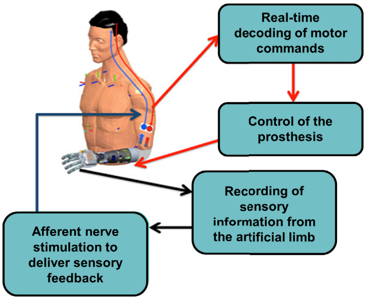


Fig. 1. Conceptual scheme of the NEBIAS neurocontrolled prosthesis.

The development of the neuro-controlled upper limb prosthesis will be achieved by combining microtechnology and material science. It will allow, on one side, recording of the motor-related signals governing the actions of the amputated hand/arm for the motion control of a mechanical prosthesis, and on the other providing sensory feedback from tactile and kinaesthetic sensors through neuromorphic stimulation of the adequate afferent pathway within the residual limb. Moreover, outcomes of this kind of research activity will allow the achievement of increased neuroscientific, clinical and technological knowledge, guidelines for the development of the other bidirectional interfaces and neural prostheses, as well as roadmaps for future development of hybrid bionic systems.

In the next sections, some of the highlights of the NEBIAS project at the stage of its development are presented.

2.1 Implantable Components

Peripheral Neural Interfaces. A novel single sided intrafascicular electrode (NEBIAS1, see Figure 2A) was designed and manufactured during the first year of the project [8]. The design of the NEBIAS1 electrode is inspired by the design of the TIME electrode [13]. All conductive paths and electrodes are fabricated of Platinum/Iridium alloy (Pt/Ir) whereas the substrate is Parylene C.

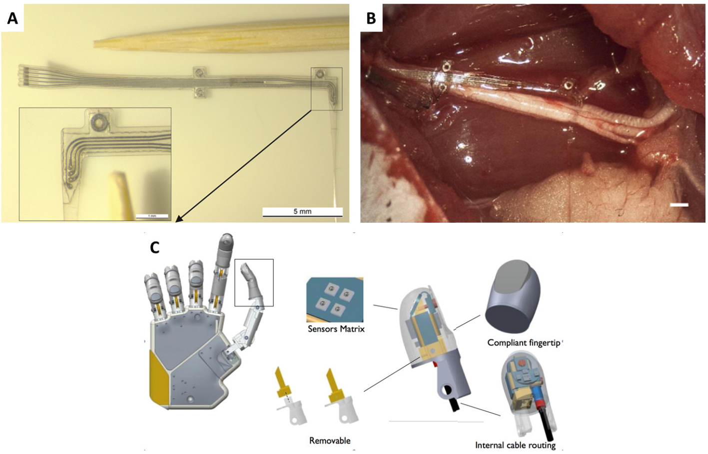


Fig. 2. A: schematic view of a NEBIAS1 electrode; B: a close view of a NEBIAS1 electrode transversally implanted into three fascicles (peroneal, tibial and sural) of the rat sciatic nerve; C: scheme of the main components of the biomimetic fingertip.

While the single conductive paths have a comparable width of 15 μm the pitch between them is about 50 μm. The electrode has a total of 5 contacts. One is a 1 mm*0.1 mm ground contact that is placed outside of the nerve, the other four are 80 m in diameter with a spacing of 200 μm. Total horizontal length of the electrode is 22 mm which allows for interconnection via microflexing as well as sewing the electrode to the nerve. For sewing of the electrode three metal strengthened openings with a diameter of 250 μm are placed alongside the electrode.

NEBIAS1 electrodes were implanted in the sciatic nerve of rats (see Figure 2B) and tested by means of functional, electrophysiological, and histological tests during 3 months follow-up [5]. In parallel, experiments with rhesus macaques are currently performed in order to develop implantation techniques of neural interfaces in the upper limb, to train animals in grasping and perceptual discrimination tasks [11], and to test decoding and encoding algorithms (see next subsections).

Embedded Electronics. An embedded electronic device which will be the bridge between the peripheral neural interface and the artificial limb has been designed an it is currently under a phase of testing [3]. The embedded electronics incorporates (i) a low-noise, low-power front-end for accurate recording of the neural signals captured by the electrodes and safe injection of feedback stimuli and (ii) an embedded digital signal processing architecture able to implement decoding algorithms in real-time and to control the front-end unit.

Some of the developed processing algorithms (see next subsection) have been ported onto an off-the-shelf low-power hybrid multicore DSP, redefining them in order to achieve high performance on the chosen platform. For the most complex algorithm, a deep analysis has been carried out to estimate the final performance in terms of both effectiveness (classification results) and efficiency (latency model, power analysis). The development of a custom architecture implemented on Field Programmable Gate Arrays (FPGA), as a proof of concept of a prospective VLSI implementation, has been also pursued [4].

2.2 Artificial Hand

A novel biomechatronic prosthetic finger endowed with neuromorphic tactile sensors was developed (see Figure 2C) [7]. The designed prosthetic finger integrates the following features of the human finger, with a multi-level biomimetism: (i) the kinematics of the finger and the shape of the fingertip, which can be represented as the aspects reproducing the bone and the muscles of the biological finger; (ii) the deformation properties of the finger skin, reproduced with custom artificial materials.

2.3 Algorithms

Decoding Algorithms. Two approaches have been investigated for the extraction of information(e.g., amputee's intentions) from neural signals recorded from residual peripheral nerves [6]. The first approach could be used in case of neural signals recorded with multi-contact cuff electrode or in case of cumulative signals recorded with intraneural electrodes. It is consists of (i) pre-processing to eliminate the EMG low band and amplifiers high band noise, feature extraction, and classification (by means of Support Vector Machines) [1]. The second approach could be used in case of neural signals with spikes recorded with intraneural electrodes. It consists of pre-processing and wavelet denoising, spike detection and sorting, feature extraction, and classification [10].

Neuromorphic Coding of Sensory Information. The objective of the neuromorphic coding of sensory information is the deliver of neuromorphic inputs into the nervous system to transmit information about the tactile stimuli [12]. This is needed to provide sensory feedback during movement of the hand neuroprosthesis. To this aim the following research actions are currently pursued: to improve the understanding of the encoding of tactile stimuli features in spike trains at the peripheral level, to develop a system able to create proper spiking outputs mimicking the firing dynamics of primate mechanoreceptors, and to make sure that discrimination abilities are robust with respect to variations of the mechanical dynamics of tactile experience (i.e., change in contact force or exploration velocity).

Preliminary experiments were carried out microstimulating the median nerve of 6 healthy subjects via percutaneous needle electrodes. EEG was also recorded in order to correlate the stimulation with electrophysiological measurements at central nervous system level. Moreover, in order to deliver neuromorphic inputs to the nervous system a preliminary but systematic evaluation of the neuromorphic sensor able to convert mechanical tactile stimulation into spiking patterns reproducing those of human mechanoreceptors was carried out.

3 Experimental Validation with Humans

The creation of a symbiotic relationship between an amputee (and his nervous system) and an artificial hand has been recently investigated by members of the NEBIAS project in a case-study [9]. The use of intrafascicular electrodes to link the sensory information from the sensors embedded in the artificial hand with the brain allowed a blindfolded and acoustically isolated amputee to feel the stiffness and the shape of three different objects grasped with the prosthesis [9].

After two years of the project, the first demonstrator of the NEBIAS neuro-controlled prosthesis (with all the elements portable, wearable or implanted) will be tested in one or more transradial amputees enrolled for the study. The patients will undergo a complete clinical and functional neuroimaging study to verify the safety of the prosthetic system and its role in modifying cortical reorganizations following limb amputation and lack of sensory feedback. The feeling of naturalness and effectiveness of the bidirectional neuro-control of the NEBIAS demonstrator will be assessed by means of SHAP and "box and block" tests [14]. Moreover, a measure of the afferent neural feedback ability to promote the embodiment of the prosthesis in upper limb amputees will be evaluated using the rubber hand illusion test. A clinical follow-up evaluation will be also performed 2 months after the electrode removal.

4 Conclusions

NEBIAS is a highly innovative, interdisciplinary project, combining forefront research from information technologies, smart biosensors, control theory, neuroscience, material sciences, embedded electronics, and robotics to solve a major social problem: the development of a prosthetic hand displaying all the basic features of a real human hand. The successful realisation of this highly visionary project requires crossing the boundaries of distinct scientific fields, merging forefront expertise of the consortium to improve quality of life of amputees.

This short overview has presented some of the highlights of the NEBIAS project at the stage of its development.

References

1. Badia, J., Raspopovic, S., Carpaneto, J., Micera, S., Navarro, X.: Spatial and Functional Selectivity of Peripheral Nerve Signal Recording With the Transversal Intrafascicular Multichannel Electrode (TIME). IEEE Trans. Neural Syst. Rehabil. Eng. (in press) (2015)
2. Biddiss, E., Beaton, D., Chau, T.: Consumer design priorities for upper limb prosthetics. Disabil. Rehabil. Assist. Technol. **2**(6), 346–57 (2007)
3. Carboni, C., Bisoni, L., Carta, N., Barbaro, M.: Compact, multi-channel, electronic interface for PNS recording and stimulation. In: Proc. Int. Conf. Biomed. Eng. (BioMed), Zurich (2014)
4. Carta, N., Meloni, P., Tuveri, G., Pani, D., Raffo, L.: A Custom MPSoC Architecture With Integrated Power Management for Real-Time Neural Signal Decoding. IEEE J. Trans. Emerg. Sel. Topics Circuits Syst. **4**(2), 230–241 (2014)
5. del Valle, J., de la Oliva, N., Mueller, M., Stieglitz, T., Navarro, X.: Biocompatibility evaluation of parylene C and polyimide as substrates for peripheral nerve interfaces. In: Proc. 7th Conf. IEEE/EMBS Neural Engineering (NER), Montpellier (2015)
6. Micera, S., Carpaneto, J., Raspopovic, S.: Control of hand prostheses using peripheral information. IEEE Rev. Biomed. Eng. **3**, 48–68 (2010)
7. Montagnani, F., Controzzi, M., Cipriani, C.: Is it Finger or Wrist Dexterity That is Missing in Current Hand Prostheses? IEEE Trans. Neural Syst. Rehabil. Eng. **23**(4), 600–609 (2015)
8. Mueller, M., Ulloa Suarez, M.A., Schuettler, M., Stieglitz, T.: Development of a single-sided parylene C based intrafascicular multichannel electrode for peripheral nerves. In: Proc. 7th Conf. IEEE/EMBS Neural Engineering (NER), Montpellier (2015)
9. Raspopovic, S., Capogrosso, M., et al.: Restoring natural sensory feedback in real-time bidirectional hand prostheses. Sci. Transl. Med. **6**(222), 222ra19 (2014)
10. Rossini, P.M., Micera, S., et al.: Double nerve intraneural interface implant on a human amputee for robotic hand control. Clin. Neurophysiol. **121**(5), 777–783 (2010)
11. Schaffelhofer, S., Agudelo-Toro, A., Scherberger, H.: Decoding a wide range of hand configurations from macaque motor, premotor, and parietal cortices. J. Neurosci. **35**(3), 1068–1081 (2015)
12. Spigler, G., Oddo, C.M., Carrozza, M.C.: Soft-neuromorphic artificial touch for applications in neuro-robotics. In: Proc. 4th IEEE RAS/EMBS (BioRob), Rome (2012)
13. Stieglitz, T., Boretius, T., Navarro, X., et al.: Development of a neurotechnological system for relieving phantom limb pain using transverse intrafascicular electrodes (TIME). Biomed. Tech. (Berl) **57**(6), 457–465 (2012)
14. Quispe, A.H., Ben Amor, H., Christensen, H.: Supplement to: a taxonomy of benchmark tasks for bimanual manipulators. In: Proc. Intl. Conf. Robotics Research (ISRR), Sestri Levante (2015)

Comparing Input Sensors in an Immersive Mixed-Reality Environment for Human-Computer Symbiosis

Paolo Negri[1(✉)], Pedro Omedas[2], Luca Chech[1], Patrik Pluchino[3], Filippo Minelle[3], Paul F.M.J. Verschure[2,4], Giulio Jacucci[5], Jonathan Freeman[6], Anna Spagnolli[1,3], and Luciano Gamberini[1,3]

[1] Department Of General Psychology, University of Padua, Padua, Italy
{Paolonegri0,luca.chech}@gmail.com,
{anna.spagnolli,luciano.gamberini}@unipd.it
[2] Synthetic Perceptive Emotive and Cognitive Systems Lab, Center of Autonomous Systems and Neurorobotics, Universitat Pompeu Fabra, Barcelona, Spain
pedro.omedas@upf.edu
[3] Human Inspired Technology Research Centre, University of Padua, Padua, Italy
patrik.pluchino@unipd.it, minelle.filippo@gmail.com
[4] Instituciò Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
paul.verschure@upf.edu
[5] Helsinki Institute for Information Technology HIIT, Department of Computer Science, University of Helsinki, Helsinki, Finland
giulio.jacucci@helsinki.fi
[6] Psychology Department, Goldsmiths University of London, London, UK
J.Freeman@gold.ac.uk

Abstract. With the emergence of interest in the human-computer symbiosis also rises the need to find input systems adequate to this paradigm. The present research aims to compare three different input systems during the interaction with virtual objects in a wide, immersive mixed-reality environment. The common interaction via keyboard and mouse is compared with two types of interaction mediated by gestural inputs. Specifically, we compared the performance and the user experience of participants interacting with a virtual 3D model of a human brain either with keyboard/mouse, or with two different motion sensing devices to input commands: the Microsoft Kinect360 and the KinectOne. The results seem to suggest that, although participants showed a better efficiency using the keyboard/mouse, in a high immersive environment, a input system that exploits gestures and body movements without requiring the use of any physical artifact, seems to be the preferred one to use.

Keywords: Symbiotic system · Mixed-reality · Natural interface

1 Introduction

In the last few years, in the Human-Computer Interaction (HCI) field, a concept has gained attention, that is, the human-computer symbiosis.

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 111–125, 2015.
DOI: 10.1007/978-3-319-24917-9_12

The idea of human-computer symbiosis is not novel, indeed it dates back to 1960 when Licklider proposed the analogy to the symbiosis occurring in the biological world [1]. Recently, Jacucci and collaborators [2] proposed a detailed definition based on the recent developments in computing science. According to these authors a symbiotic interaction "can be achieved by combining computation, sensing technology, and interaction design to achieve deep perception, awareness, and understanding between humans and computers." [2]. Other authors [3] recently proposed a taxonomy of human-computer interaction for the purpose of point out a range of interaction uses for human-computer symbiosis.

Over the last few years several efforts to explore a symbiotic relationship between human and computer have been made in different domains. In the field of human-robot interaction efforts focused to improve the compliance between wearable robots and user [4] and to improve the efficiency when swarms of partially autonomous and biologically inspired robots collaborate with human [5]. In the field of mobile internet, aiming at supporting a symbiotic relationship between media content and physical places, we can find examples like OUTMedia [6], a location-based music discover application. In the field of information retrieval in complex scenarios, a recent study [7] investigated the possibility of using the pupil behavior of users as implicit input to improve the interaction between information retrieval systems and users. In another study [8], authors proposed a new type of interactive image retrieval system more heavily based on user's research intent. On the technical side, recently a research group proposed [9] a novel tactile glove concept that delivers tactile feedback to the user while he moves in the space of interaction. Other works proposed novel design solutions for future searching interfaces [10] and information retrieval systems [11]. Recent research has tackled the topic concerning the use of physiological measures to infer user's cognitive and affective state [12] that the system could take into account to adapt in real time the interaction. Interestingly, other studies [13,14,15,16] investigated the possibility to introduce subliminal stimuli in interfaces to guide the user when he is struggling with the interaction without requiring him cognitive efforts. Finally, other authors [17] have engaged in developing questionnaires to measure the user experience (UX) related to the use of wearable devices for symbiotic systems.

A relevant aspect to the human-computer symbiosis is the reciprocity and collaborative use of resources for both computers and humans [2]. To this end, and aiming to create a reciprocal, deeper understanding between humans and computers, the system needs to be informed about intentions and internal states of humans by means of a confluence of both explicit (e.g. symbolic communication like written words, body gesture) and implicit (e.g. psychophysiology) signals [2]. On the one hand, the computer can understand the human intentions by constantly monitoring both explicit behaviors and implicit signals that might be recognized with sensing technologies. On the other hand, the computer provides feedback to the human by means of both ordinary (e.g. symbolic, explicit communication) and implicit (e.g. subliminal) outputs [13,14,15,16] establishing a wider communication loop.

The confluence of inputs is a relevant question for the human-computer symbiosis. In this regard, a recent study [18] showed that, interacting with a large dataset, participants exhibited better learning performances when they used a wide and immersive environment which exploits body movements and natural gestures than when they interacted with the same dataset using a normal desktop PC. These findings suggested that the embodied interaction is advantageous for learning when interacting with large datasets.

Given this, the present research aims to take a step forward by comparing the performance and the UX related to three types of input system during the interaction with virtual objects in an immersive environment. The common interaction via keyboard and mouse is compared with two types of embodied interaction in a mixed-reality environment created with the eXperience Induction Machine (XIM) [19]. The XIM is an immersive space consisting of effectors (projectors and loudspeakers) and sensors (see the section “Setting and equipment” for a detailed description) conceived with the aim of studying the human-artifact interaction in condition of good ecological validity [19] and inspired by the ADA project, a large-scale public exhibit for the Swiss Expo.02 national exhibition [20].

As for the embodied interaction, two motion-sensing input devices were selected, the widely used Microsoft Kinect360, and the next version released by Microsoft, the KinectOne. The interaction and visualization in XIM is controlled by the XIM-engine [21], which is in charge on interpreting the input of sensors and change the visualization accordingly.

The study evaluates the interface of the BrainX^3 [22,23] application, a tool designed for exploration of neuroscience datasets that has been develop using the XIM framework. Therefore, the present study compares the performance and the UX inside the XIM using the BrainX^3 interface, when participants utilized for interacting respectively the keyboard and the mouse, or the Kinect360 or the KinectOne as input systems. Noteworthy, we conceived the interaction by means of the Kinect360 requiring the use of a mouse to be kept in the right hand while the interaction by means of the KinectOne did not require it.

Participants were asked to perform ten tasks within the XIM using one input device and then to repeat the task series with each input device. As for the performance, we measured it in terms of time required to complete each task, and we hypothesized that participants would perform better with the keyboard/mouse as input system due to the fact that they are very accustomed to use them. As for the UX, we measured it by means of two questionnaires and we hypothesized that participants would express better evaluations of the interaction via the Kinects than with the keyboard/mouse. This hypothesis roots in the idea that gestural input systems, because they are natural systems, reduce the abstraction of input actions and also the users' effort needed to input commands and express preferences. In fact, gestural inputs (e.g. take a step forward to zoom-in on a visual content) are designed to easily suggest their meaning, that is the action they allow to perform on the content. Therefore, the user is not asked to learn an abstract association between a particular action (e.g. pressing a specific keyboard key) and a particular consequent response by the system. Finally, we hypothesized that the KinectOne would be preferred over the Kinect360. This latter hypothesis is based on the fact that, in our paradigm the Kinect360 requires the use of a mouse, while the KinectOne does not. Therefore, we speculated that participants would experience the use of the KinectOne as more natural and simple.

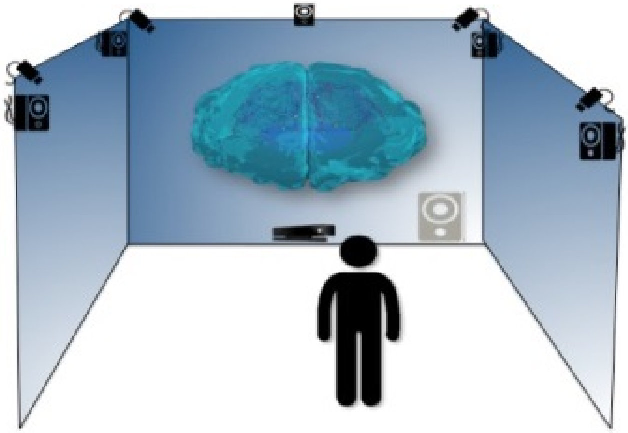


Fig. 1. The eXperience Induction Machine (XIM).

2 Method

2.1 Participants and Design

Twelve undergraduate/graduate students of the University of Padua were recruited for this study (9 males and 3 females; mean age = 23.91; SD = 2.43; range 20-28). All participants had normal or corrected-to-normal vision and gave their informed consent. Participants were asked to perform a series of tasks within the XIM. The input system was varied in a within-participant design, namely, all the participants used all three input systems. The order of the input systems was counterbalanced across participants.

2.2 Setting and Equipment

During the experiment, the participant was the only person that remained within the XIM while the researchers monitored the experimental session through four video-cameras that were mounted inside the XIM (Figure 1). The experimenters instructed the participant through pre-recorded vocal commands.

In one condition the participant could interact with the system via wireless keyboard and mouse seated in a chair and with a table, while in the other two conditions they interacted with the system respectively via a Kinect360 plus a wireless mouse, and via a KinectOne standing in front of the screen.

When using the keyboard/mouse, participants could control the cursor on the screen by means of the mouse. Moreover, users had to press the left button of the mouse in order to select a cerebral area or a button of the interface (the software will be presented in the next section).

When using the two Kinects, participants could control the cursor on the screen by moving the right hand with the arm stretched. The only difference between the two Kinects concerned the way in which the selection command was implemented. In fact, with the Kinect360 participants had to press the left button of the mouse they kept in their right hand, instead with the KinectOne they simply had to close and then open their right hand.

Another relevant difference in the way participants could interact with the interface concerned the zooming actions. When using the keyboard/mouse, participants could perform the zoom in and the zoom out by pressing two keys, instead when using the two Kinects they had to step forward or backward respectively.

Two PCs were utilized in the experiment. Three projectors, which allowed the interface to be displayed on the interior panels of the XIM, were connected to a display machine. The Kinect360 and the KinectOne were connected to a sensor machine via USB cables. Participants were video-recorded during the experiment (i.e. four video-cameras recorded the participants while they were interacting with the interface).

2.3 The Interface

The interface used was designed for interacting with neuroscience data displayed by the BrainX[3] application, which consisted of three different parts, that is, a frontal panel, a right one and a left one. Figure 2 represents the panel in which the connectome, that is the 3D model of the human brain, was shown. In the lower part of this screen, five buttons were presented. The reset button was always visible in the bottom left corner. By clicking on this button the model of the brain went back to the starting position. Whenever the cursor was placed over any area of the brain three buttons appeared. The remove button was used to cancel an action performed on a cerebral area; the bookmark button was used to highlight an area; the inject activity button was used to visualize the neural activity of an area. Finally, the complexity button was always shown in the lower right. By clicking on this button it was possible to increase or decrease the complexity level of the brain representation. See [23] for a complete description of the system interface.

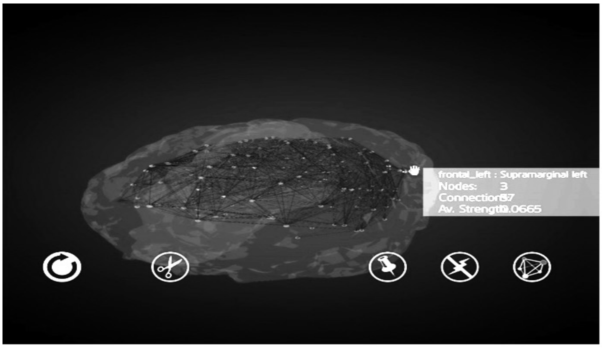


Fig. 2. The interface.

The left screen offered multiple representations of the area currently examined along the three different planes of view (sagittal, coronary and transverse), while the right screen represented the information about the area of the brain on which the cursor was located. However, the task devised for this study did not require the participants to pay attention to the lateral screens.

2.4 Tasks

Participants were asked to perform 10 tasks within BrainX^3 inside the XIM: (1) Multiple pointing. Participants had to place the cursor above three areas highlighted by the experimenter with a laser pointer in a predefined, fixed sequence. (2) Using the bookmark button. Participants had to select a specific cerebral area, grab the corresponding circle, drag it on top of the Bookmark button and then release it. (3) Performing a horizontal leftward rotation. Participants had to horizontally rotate the brain until a cerebral area (highlighted by the experimenter), which was initially located laterally to the right, came to be positioned in the central area of the frontal panel. (4) Performing a horizontal rightward rotation. Participants had to horizontally rotate the brain in order to place a cerebral area (highlighted by the experimenter) which was initially located laterally to the left, in the central area of the frontal panel. (5) Performing a vertical downward rotation. Participants had to vertically rotate the brain in order to place a cerebral area (highlighted by the experimenter) which was initially located in a rostral position within the brain, in the central area of the frontal panel. (6) Performing a vertical upward rotation. Participants had to vertically rotate the brain in order to place a cerebral area (highlighted by the experimenter) located in an inferior position within the brain, in the central area of the frontal panel. (7) Zooming in. Participants had to increase the size of the brain till it reached its maximum. (8) Zooming out. Participants had to reduce the size of the brain till it reached its minimum. (9) Using the inject activity button. Participants had to activate a cerebral area by grabbing the corresponding circle and dropping it on the Inject Activity button. (10) Using the remove button. Participants had to remove a cerebral area by grabbing the corresponding circle and dropping it on the Remove button.

The tasks differed in complexity based on the number of movements that needed to be performed to complete them.

2.5 Measures

Task-Related Experience. A six-item questionnaire was administered at the end of each task. These items aimed at measuring the UX for each combination of input device and task. Participants were asked to express their agreement on a 5-point Likert scale with regard to the following statements: (1) The execution of the commands is easy; (2) The execution of the commands is pleasant; (3) The meaning of the command is intuitively associated with its function; (4) The system responds promptly to the command; (5) The execution of the command is complicated; (6) The cursor/brain on the display moves smoothly.

Device-Related Experience. Seven questions were asked to the participants at the end of the experiment. These questions evaluated the general UX by comparing the three input systems: (1) Which system is easier to use? (2) Which system is more pleasant to use? (3) Which movement was the most difficult to perform with the Kinect360? (4) Which movement was the most difficult to perform with the KinectOne? (5) Which movement was the most difficult to perform with the Keyboard? (6) Which system seemed smoother? (7) What would you change in general?

Task Execution Time. The time to complete each task was measured and considered as a measure of performance.

Video Recording. The whole experimental session was recorded in order to perform video analysis of the participants while executing the tasks with each input system.

2.6 Procedure

At the beginning of the experiment the participants filled in the informed consent that contained a release note for the video-recorded material. Then, they entered the XIM. The researchers monitored the experiment from outside the XIM.

The participants were informed that they would have to perform various tasks which required them to interact with a 3D model of the human brain inside the XIM by using three different input systems. They were instructed to reach the starting position that could be either a desktop at the center of the XIM (in the condition with the wireless keyboard and mouse) or a white line on the floor at the center of the XIM (in the conditions with Kinect360 and KinectOne). The order of conditions (i.e., the order of use of the input systems) was counterbalanced. No time limit was set to complete the task. At the end of each task, the participants could take a short break.

For each task, pre-recorded instructions were presented to the participants, and, when needed, one of the experimenter was in charge of showing the area on which the task had to be performed, by using a laser pointer. Participants could ask to repeat the task instructions. The instructions pertaining to the same task could be slightly different in accordance to the input system.

After all tasks were completed with one device, the participants started anew with the subsequent input device. Therefore the same task series was repeated three times.

At the end of each task, participants were presented with the 6-item task-related questionnaire. At the end of the whole experiment participants were presented with the device-related questionnaire which required them to compare the three conditions.

The testing session lasted about one hour.

3 Results

3.1 Task Execution Time

As for the task execution time, in order to see if the input system utilized had an effect on the performance, a repeated-measures ANOVA with device (three levels) and task (ten levels) as within-participants factors was performed. The analysis revealed a main effect of the task, $F(9,99) = 11.63$, $p < .001$, $\eta^2_p = .51$, indicating that, regardless of the device, the ten tasks differed in the time needed to accomplish them, showing that the task series included tasks of varying difficulty.

Table 1. Marginal Means, Standard Errors and *p*-value of Time (in seconds) on Task, by Type of Sensor. * refers to the comparison between Keyboard/Mouse and Kinect360, ** between Keyboard/Mouse and KinectOne, and *** between Kinect360 and KinectOne. In the *p*-value column only statistically significant comparisons are reported.

Task	Keyboard/Mouse (n=12)	Kinect 360 (n=12)	KinectOne (n=12)	*p*-value
	M (SE)	*M (SE)*	*M (SE)*	
Multiple pointing	6.37 (0.31)	16.28 (1.21)	17.94 (3.12)	* < .001 ** = .009
Bookmark button	4.97 (0.20)	12.54 (1.74)	10.11 (0.67)	* = .003 ** < .001
Horizontal leftward rotation	9.74 (1.29)	11.67 (1.53)	11.37 (1.78)	-
Horizontal rightward rotation	7.91 (0.90)	12.10 (2.79)	11.23 (2.04)	-
Vertical downward rotation	3.82 (0.37)	10.87 (1.20)	15.36 (2.79)	* < .001 ** = .004
Vertical upward rotation	3.70 (0.39)	12.33 (1.04)	9.30 (1.35)	* < .001 ** = .001
Zoom-in	11.39 (0.17)	25.39 (2.56)	12.95 (0.43)	* = .001 ** = .034 *** = .001
Zoom-out	10.69 (0.29)	15.21 (0.84)	27.12 (3.14)	* = .002 ** = .001 *** = .003
Inject activity button	4.89 (0.16)	11.12 (0.84)	11.90 (2.79)	* < .001
Remove button	4.31 (0.11)	10.44 (1.02)	8.78 (0.68)	* < .001 ** < .001

The analysis revealed also a main effect of the device, $F(2,22) = 72.99$, $p < .001$, $\eta^2_p = .869$, indicating that, regardless of the specific task, the mean time needed by the participants in order to accomplish a task differed with the three input devices. The pairwise comparisons showed that when using the keyboard/mouse ($M = 6.78$, $SD = 1.03$) participants completed their tasks more quickly than with both the Kinect360 ($M = 13.80$, $SD = 1.96$; $p < .001$, $d = 4.48$) and the KinectOne ($M = 13.60$, $SD = 2.88$; $p < .001$, $d = 3.15$). No significant differences emerged between the Kinect360 and the KinectOne.

In addition, a two-way interaction between device and task emerged, $F(18,198) = 5.94$, $p < .001$, $\eta^2_p = .35$, indicating that the three devices performed differently with regard to the execution time depending on the task. We therefore compared performance task by task, finding that where there was a difference it was in favor of the keyboard/mouse condition. Between the two Kinect conditions there was generally no difference except in the zoom-in task where the KinectOne was faster than the Kinect360, and in the zoom-out task where the opposite result was found. The pairwise comparisons are summarized in Table 1.

3.2 Video Analysis

A video analysis was performed on the recorded experimental sessions to identify action breakdowns, namely observable interruptions in the course of an action not due to system failures (e.g. rotation interruptions). Three circumstances in which breakdowns occurred during the task execution were observed: (1) rotation interruptions: the gesture with the arm was not executed correctly; (2) zoom-in interruptions: the participant moved outside of the field within which the Kinect can detect the body; in these situations the interface stopped working until the participant came back inside the area tracked by the Kinect; (3) cursor accuracy: the participant was not able to maintain the cursor stable while performing an action.

Table 2. N. of participants experiencing each type of breakdown per condition.

	Keyboard/Mouse	**Kinect360**	**KinectOne**
Rotation interruptions	1	5	2
Zoom-in interruption	0	4	0
Cursor accuracy	1	8	1

Table 2 shows that a higher number of participants experienced such issues when they utilized the Kinect360 as input system. Instead, when using the KinectOne and Keyboard/mouse, a lower number of breakdowns was observed. Thus, using the Kinect360 participants had difficulties chiefly under the following circumstances: (a) involuntary movements interpreted by the system as commands; (b) difficulty in recognizing the participant if outside of the tracking volume (too close or too far from the device); (c) difficulty in recognizing gestures if not performed properly.

3.3 Questionnaires

Task-Related Experience. Negative items were recoded so that higher scores corresponded to a positive evaluation. The Kronback's alpha of the questionnaire in the three conditions was high (keyboard/mouse α = .97; Kinect360 α = .98; KinectOne α = .98), showing consistency between the items. In order to see if the input system utilized had an effect on the UX, a repeated-measures ANOVA was performed with the device (three levels), task (ten levels), and question (six levels) as within-participants factors.

Table 3. Marginal Means, Standard Errors and *p*-value of User experience scores by Task and Type of Sensor (data from Task-related questionnaire). * refers to the comparison between Keyboard/Mouse and Kinect360, ** between Keyboard/Mouse and KinectOne, and *** between Kinect360 and KinectOne. In the *p*-value column only statistically significant comparisons are reported.

Task	Keyboard/Mouse (n=12)	Kinect 360 (n=12)	KinectOne (n=12)	*p*-value
	M (SE)	*M (SE)*	*M (SE)*	
Multiple pointing	4.33 (0.18)	3.33 (0.18)	3.81 (0.27)	* = .001
Bookmark button	4.40 (0.15)	3.51 (0.19)	3.82 (0.25)	* = .002
Horizontal leftward rotation	3.98 (0.24)	4.01 (0.20)	4.00 (0.20)	-
Horizontal rightward rotation	4.01 (0.22)	4.08 (0.19)	4.10 (0.19)	-
Vertical downward rotation	4.29 (0.16)	3.99 (0.21)	3.71 (0.26)	** = .030
Vertical upward rotation	4.36 (0.14)	3.93 (0.22)	4.00 (0.20)	-
Zoom-in	4.33 (0.13)	4.32 (0.17)	4.40 (0.16)	-
Zoom-out	4.24 (0.11)	4.35 (0.15)	4.22 (0.18)	-
Inject activity button	4.26 (0.19)	3.85 (0.23)	4.10 (0.22)	-
Remove button	4.36 (0.18)	3.79 (0.24)	4.22 (0.19)	* = .034 *** = .009

A main effect of the device emerged, $F(2,22) = 3.80$, $p = .04$, $\eta^2_p = .257$, but the pairwise comparisons did not show any significant differences between devices. Also a main effect of the task emerged, $F(3.663, 40.291) = 4.132$, $p = .008$, $\eta^2_p = .27$, indicating that, regardless of the device or the specific question, the ten tasks differed in their overall evaluation.

A two-way interaction between device and task emerged, $F(18,198) = 4.687$, $p < .001$, $\eta^2_p = .30$, indicating that the three devices differed in the overall evaluation of the UX depending on the task. However, only in four tasks out of ten there was a difference in the pairwise comparisons, and it was in favor of the keyboard/mouse condition at the expenses of the Kinect360 condition. The pairwise comparisons are summarized in Table 3.

The analysis also revealed a two-way interaction between the device and the questionnaire items, $F(4.069,44.758) = 6.88$, $p < .001$, $\eta^2_p = .385$, indicating that, regardless of the task, participants evaluated the UX differently when using different devices depending of the specific question. The pairwise comparisons are summarized in Table 4.

Table 4. Marginal Means and Standard Errors of User Experience scores, by question and Type of Sensor (data from Task-related questionnaire). * refers to the comparison between Keyboard/Mouse and Kinect360, ** between Keyboard/Mouse and KinectOne, and *** between Kinect360 and KinectOne. In the *p*-value column only statistically significant comparisons are reported.

Item	Keyboard/Mouse (n=12)	Kinect 360 (n=12)	KinectOne (n=12)	*p*-value
	M (SE)	*M (SE)*	*M (SE)*	
The execution of the commands is easy	4.70 (0.12)	4.13 (0.16)	4.30 (0.16)	* = .017
The execution of the commands is pleasant	3.71 (0.17)	3.83 (0.27)	4.18 (0.24)	-
The meaning of the command is intuitively associated with its function	4.12 (0.15)	4.22 (0.18)	4.31 (0.21)	-
The system responds promptly to the command	4.27 (0.19)	3.72 (0.20)	3.70 (0.20)	* = .005 ** = .017
The execution of the command is complicated	1.34 (0.14)	1.95 (0.17)	1.97 (0.19)	* = .014 ** = .014
The cursor/brain on the display moves smoothly	4.09 (0.28)	3.52 (0.21)	3.71 (0.25)	* = .025

Device-Related Experience. In answering the question “which system is easier to use” the majority of the participants (83%) evaluated the keyboard/mouse as the easiest. A small number of participants (17%) evaluated the KinectOne as the easiest system to use. None of the participant considered the Kinect360.

Noteworthy, in answering the question “which system is more pleasant to use” the majority of the participants (58,4%) evaluated the KinectOne as the more pleasant. A third of the participants (33,3%) evaluated the keyboard/mouse as the more pleasant. One (8.3%) of the participants considered the Kinect360 as the more pleasant.

With regard to the questions “which movement was the most difficult to perform with the Kinect360 (or the KinectOne or the Keyboard/mouse)” a lower number of issues have been observed in the keyboard/mouse condition. In fact for this condition 9 participants out of 12 responded “none”, whereas for the Kinects conditions they indicated a variety of problems (e.g. selecting a node, horizontal rotation). Relevantly, for the Kinect360 5 participants out of 12 responded “cursor accuracy”.

To the question “which system seemed smoother” all participants responded “Keyboard/mouse”.

4 Conclusions

The re-emergence of interest in the human-computer symbiosis calls for the adoption of new, more adequate input systems in order to better support their interaction.

The present research aimed at comparing three types of input systems in an immersive mixed-reality environment, namely, the keyboard and mouse and two input systems based on body movements and natural gestures (the Microsoft Kinect360 and the KinectOne).

As for the interaction with gestural inputs, two motion-sensing input devices were selected, the widely used Microsoft Kinect360, and its subsequent version, the KinectOne.

Therefore, in the present study both performance and UX were evaluated in a mixed-reality environment while engaging with a neuroscience tool called BrainX[3], which required them to interact with a 3D model of the human brain by using the aforementioned devices.

Regarding the performance (measured in terms of time needed to complete each task), participants were faster in completing the tasks when they utilized the keyboard and mouse compared to both the other devices. This is in line with the basic consideration that participants are more familiar with the keyboard and mouse as a command input system.

Regarding the UX, we hypothesized that participants would evaluate better the interaction via the Kinects, and especially via the KinectOne, compared to the interaction with the keyboard/mouse. As for the results, firstly, the six-item UX questionnaire did not show differences in UX scores between the three input systems in the majority of tasks. When a difference emerged it regarded the comparison between the keyboard and mouse and the Kinect360 and favored of the former. However, it is important to underline that, in general, participants did not express better evaluations when they used the keyboard/mouse compared with the KinectOne despite the fact that they were more accustomed to using the keyboard/mouse. Moreover, considering the final questionnaire, the majority of participants evaluated the KinectOne as the more pleasant input system. Thus, even if the time needed to accomplish the tasks was greater when using the KinectOne, the participants nonetheless judged this device to be the more pleasant to use during the entire experiment. This result partially supports our hypothesis that participants would express better evaluations of the interaction via the KinectOne. However, we should also consider the "novelty effect", namely the possibility that participants favored the KinectOne over the keyboard and mouse in part due to the fact that it was a novel, and thus more stimulating and engaging way to interact.

At the same time, the video analysis showed that the number of people that experienced difficulties during the tasks execution was higher in the Kinect360 condition, while the number of participants that experienced difficulties when using both the keyboard/mouse and the KinectOne was almost equal. This could in part explain the worst judgment in term of preference received by the Kinect360.

The results of the present study seem to suggest that, although participants showed a better efficiency when using the keyboard/mouse, in a high immersive environment as the XIM, an input system exploiting gestures and body movements and that does not require the use of any physical artifact (i.e. the mouse) seems to be the preferred one. Although this is what emerges from the final questionnaire, we cannot completely exclude that the preference for the KinecOne was due, at least in part, to its novelty.

In conclusion, these findings, together with those of previous studies [18] which demonstrated that the embodied interaction is advantageous for learning when interacting with large datasets, suggest that input systems exploiting body movements without requiring the use of physical artifacts (i.e. keyboard and mouse) could be the most appropriate in symbiotic systems with large immersive spaces.

However, future research might consider the opportunity to make the participants accustomed to a system that involves body movements and gestures. An appropriate and longer training could be implemented in order to increase the level of expertise in utilizing these (rather new) input systems. In fact, the common interaction paradigm with technology risks to be the one that results in better performances not because it is the best way to interact *per se*, but only because people are more used to it. In the present study we did not put in place any training before measuring the performance and the UX in the experimental tasks. An adequate initial training can increase the familiarity of the participants with the new system and can thus reduce the likelihood that the traditional system produces the best performance only due to the foreseeable positive impact of a higher expertise level on the system evaluation.

Acknowledgements. This research was supported by the European Project CEEDS, The Collective Experience of Emphatic Data Systems, European Integrated Project (Number: 258749; Call: ICT 2009.8.4 Human-Computer Confluence).

References

1. Licklider, J.C.R.: Man-computer symbiosis. Hum. Factors Electron. IRE Trans. **1**, 4–11 (1960)

2. Jacucci, G., Spagnolli, A., Freeman, J., Gamberini, L.: Symbiotic interaction: a critical definition and comparison to other human-computer paradigms. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 3–20. Springer, Heidelberg (2014)
3. Lessiter, J., Freeman, J., Miotto, A., Ferrari, E.: Ghosts in the machines: towards a taxonomy of human computer interaction. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 21–31. Springer, Heidelberg (2014)
4. Moreno, J.C., Asin, G., Pons, J.L., Cuypers, H., Vanderborght, B., Lefeber, D., Ceseracciu, E., Reggiani, M., Thorsteinsson, F., del-Ama, A., Gil-Agudo, A., Shimoda, S., Iáñez, E., Azorin, J.M., Roa, J.: Symbiotic wearable robotic exoskeletons: the concept of the biomot project. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 71–82. Springer, Heidelberg (2014)
5. Andolina, S., Forlizzi, J.: A multi-touch interface for multi-robot path planning and control. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 125–130. Springer, Heidelberg (2014)
6. Åman, P., Liikkanen, L.A., Jacucci, G., Hinkka, A.: OUTMedia – symbiotic service for music discovery in urban augmented reality. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 60–70. Springer, Heidelberg (2014)
7. Pluchino, P., Gamberini, L., Barral, O., Minelle, F.: How semantic processing of words evokes changes in pupil. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 97–110. Springer, Heidelberg (2014)
8. Hore, S., Tyrvainen, L., Pyykko, J., Glowacka, D.: A reinforcement learning approach to query-less image retrieval. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 119–124. Springer, Heidelberg (2014)
9. Hsieh, Y.-T., Jylhä, A., Jacucci, G.: Pointing and selecting with tactile glove in 3D environment. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 131–135. Springer, Heidelberg (2014)
10. Serim, B.: Querying and display of information: symbiosis in exploratory search interaction scenarios. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 113–118. Springer, Heidelberg (2014)
11. Bandyopadhyay, P., Ruotsalo, T., Ukkonen, A., Jacucci, G.: Navigating complex information spaces: a portfolio theory approach. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 136–142. Springer, Heidelberg (2014)
12. Barral, O., Jacucci, G.: Applying physiological computing methods to study psychological, affective and motivational relevance. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 34–45. Springer, Heidelberg (2014)
13. Negri, P., Gamberini, L., Cutini, S.: A review of the research on subliminal techniques for implicit interaction in symbiotic systems. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 46–57. Springer, Heidelberg (2014)
14. Aranyi, G., Kouider, S., Lindsay, A., Prins, H., Ahmed, I., Jacucci, G., Negri, P., Gamberini, L., Pizzi, D., Cavazza, M.: Subliminal cueing of selection behavior in a virtual environment. Presence: Teleoperators Virtual Environ. **23**(1), 33–50 (2014)
15. Barral, O., Aranyi, G., Kouider, S., Lindsay, A., Prins, H., Ahmed, I., Jacucci, G., Negri, P., Gamberini, L., Pizzi, D., Cavazza, M.: Covert persuasive technologies: bringing subliminal cues to human-computer interaction. In: Spagnolli, A., Chittaro, L., Gamberini, L. (eds.) PERSUASIVE 2014. LNCS, vol. 8462, pp. 1–12. Springer, Heidelberg (2014)

16. Cetnarski, R., Betella, A., Prins, H., Kouider, S., Verschure, P.F.: Subliminal response priming in mixed reality: the ecological validity of a classic paradigm of perception. Presence Teleoperators Virtual Environ. **23**(1), 1–17 (2014)
17. Spagnolli, A., Guardigl, E., Orso, V., Varotto, A., Gamberini, L.: Measuring user acceptance of wearable symbiotic devices: validation study across application scenarios. In: Jacucci, G., Gamberini, L., Freeman, J., Spagnolli, A. (eds.) Symbiotic 2014. LNCS, vol. 8820, pp. 85–96. Springer, Heidelberg (2014)
18. Betella, A., Bueno, E.M., Kongsantad, W., Zucca, R., Arsiwalla, X.D., Omedas, P., Verschure, P. F.: Understanding large network datasets through embodied interaction in virtual reality. In: Proceedings of the 2014 Virtual Reality International Conference, p. 23. ACM (2014)
19. Bernardet, U., i Badia, S.B., Duff, A., Inderbitzin, M., Le Groux, S., Manzolli, J., Verschure, P.F.: The experience induction machine: a new paradigm for mixed-reality interaction design and psychological experimentation. In: The Engineering of Mixed Reality Systems, pp. 357-379. Springer, London (2010)
20. Eng, K., Babler, A., Bernardet, U., Blanchard, M., Briska, A., Conradt, J., Verschure, P.F.: Ada: constructing a synthetic organism. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 2, pp. 1808-1813. IEEE (2002)
21. Omedas, P., Betella, A., Zucca, R., Arsiwalla, X.D., Pacheco, D., Wagner, J., Verschure, P.F.: XIM-Engine: a software framework to support the development of interactive applications that uses conscious and unconscious reactions in immersive mixed reality. In: Proceedings of the 2014 Virtual Reality International Conference, p. 26. ACM (2014)
22. Arsiwalla, X.D., Zucca, R., Betella, A., Martinez, E., Dalmazzo, D., Omedas, P., Verschure, P.F.: Network dynamics with BrainX3: a large-scale simulation of the human brain network with real-time interaction. Front Neuroinformatics, 9 (2015)
23. Betella, A., Cetnarski, R., Zucca, R., Arsiwalla, X.D., Martínez, E., Omedas, P., Verschure, P.F.: BrainX 3: embodied exploration of neural data. In: Proceedings of the 2014 Virtual Reality International Conference, p. 37. ACM (2014)

Tapping Neural Correlates of the Depth of Cognitive Processing for Improving Human Computer Interaction

Irina-Emilia Nicolae[1(✉)], Laura Acqualagna[2], and Benjamin Blankertz[2]

[1] University Politehnica of Bucharest, Bucharest, Romania
irina.nicolae@aut.pub.ro
[2] Fachgebiet Neurotechnologie, Technische Universität Berlin, Berlin, Germany
{laura.acqualagna,benjamin.blankertz}@tu-berlin.de

Abstract. The classical interaction between human and a computer or a machine relies solely on explicit behaviour (input with keyboard, mouse, gestures etc.). In many situations and tasks, the access to implicit information about the user could enhance human-computer interaction (HCI). Recent research has shown a number of examples of how such hidden user states could be extracted from signals of peripheral physiology and of the brain. While these approaches are still premature and not readily available for real application, further exploration seems worthwhile. Here, we present an approach towards monitoring the level of cognitive processing. A special experimental paradigm has been designed to detect event-related potentials (ERPs) of brain activity related to cognitive processes using tasks in different cognitive domains. Neural correlates indicating different levels of cognitive processing have been singled out and the classifiability was quantified using multivariate decoding methods. The results indicate the feasibility of monitoring the depth of cognitive processing for neurotechnological applications in BCI and industrial scenarios.

Keywords: Cognitive processing · Event-related potentials (ERPs) · Classification · Brain-Computer Interface (BCI) · Electroencephalography (EEG)

1 Introduction

The interaction of a human and a machine may be enhanced if the software adapts to the momentary state of the user. New perspectives for such an adaptation in a seamless manner are opened by the presumed possibility of estimating hidden user states (i.e., those which cannot be observed from outside) from physiological signals. The applicability ranges from human-computer interaction, like information seeking, to industrial workplaces (e.g. operator monitoring) [1]. An interface that is aware of which information is more significant for the user, can adapt efficiently according to the user's needs. For example, an information seeking application could display meaningful keywords in appropriate positions, sort the information or assign different weights to the results of a query. In this context, it would be helpful to develop methods that estimate the user's level of the cognitive processing of presented information. Our genuine interest is in the natural fluctuations of cognitive processing [4],

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 126–131, 2015.
DOI: 10.1007/978-3-319-24917-9_13

e.g., caused by distraction, mind wandering, and variations in vigilance. However, in this study, we decided to take the approach of inducing different levels of processing by task instructions in order to have a better control and also because we experience in earlier studies that the natural fluctuations are not so easy to observe in laboratory settings. The aim of this work (which extends previous work [2]) is to find specific feature markers of the depth of cognitive processing, exploitable in future user state adaptation. In order to achieve this, we developed a specific experimental design, inspired by the odd-ball paradigm involving shallow and deep levels of processing in three cognitive domains: memory, language and visual imagery that can appear in a HCI system. The depth of cognitive processing refers to the degree to which information can be processed. A shallow process represents a short-term retention of information, e.g. color appearance and a deep process requires a more elaborated process, e.g. semantic correlations [3]. Discriminative neurophysiological markers are extracted using measures of separability applied to ERPs waveforms and further quantified by classification methods based on multivariate data analysis.

2 Methods

2.1 Experimental Setup

The experimental design consists in a sequence of stimuli presented on a screen, with 2500 ms stimulus onset asynchrony, using the following structure: fixation cross 500 ms, stimulus duration 1250 ms and 750 ms of blank screen. Subjects were requested to stay seated, relaxed, and try to focus in the center of the screen. Each stimulus consisted of a pair of images, having same color (either red, green, blue, or magenta) and same category (either animals, fruits, or mobility). The task varied according to whether the color, or color and category of the current stimulus matched the specified target color and category. The stimuli induced tasks involves one of the following three types of cognitive processing: non-targets (NT) requiring no further processing (neither color nor category match), shallow targets (ST) associated with a 'shallow level' of processing (only color matches: performing counting) and deep targets (DT) triggering a deep level of processing (color and category match: performing a cognitive processing task as describe below and additional counting). The experiment consists in 5 runs per condition, with 120 stimuli presented for each run, in a ratio of 75:12.5:12.5 (NT:ST:DT).

The memory condition requires the fulfillment of a task that resembles a complex form of a 1-back task (similar work in [5]): decide whether one of the objects of the current stimulus was also present in the last pair of target images (last pair with color and category matching with target). In the language condition, the subject had to compare the number of syllables of the objects presented, namely: is the number of syllables in the left object image, greater than the number of syllables of the right object? The visual imagery condition requires imagining the objects in reality and performing size comparisons, i.e.: is the object in the left, two times bigger than the one in the right? In all conditions, in case the specific question was fulfilled, 10 had to be added to the counting. The resulting number of the counting task was entered on a response keyboard at the end of the run. The brain activity was recorded using 64 channels (Brain Products, Germany).

2.2 Data Analysis

The EEG signals were off-line processed using MATLAB and EEGLAB software [6]. First, a low-pass filter at 50Hz and a high-pass filtered at 1 Hz frequency was applied. In order to clean the data of eye movements, muscle artifacts and loose electrodes, artifact removal was performed using Independent Component Analysis with MARA features [7]. The data was segmented in epochs with respect to the stimulus duration and baseline correction was performed using 100 ms of the pre-stimulus signal. The brain activity is investigated by event-related potentials (ERPs) analysis. In order to assess discriminative information between classes, we use as measure of separability, the: signed square of the point-biserial correlation coefficient, sgn r^2, computed between classes of each condition. Pairwise classification was performed using a regularized linear discriminant analysis with shrinkage [8] in 10 fold cross-validation. Spatio-temporal features [8] were considered for classification, and performed on the first five time intervals selected with maximum discrimination between classes determined by signed r2. The search interval was considered same as the stimuli interval: 0 - 1250 ms. The performance of the classification is assessed by the area under curve scores.

3 Results

3.1 Behavioral Data

Participants' answers are assessed considering the absolute differences between their responses and the correct number divided by the correct number. Fig. 1 shows box plots for the data pooled over all runs and participants. The black asterisk indicates the corresponding mean values.

We observe a slightly wider range for memory and visual imagery condition, meaning that answers were more far away from the correct number in comparison to the language case which is in line with the higher difficulty reported by the users in performing memory and visual imagery condition.

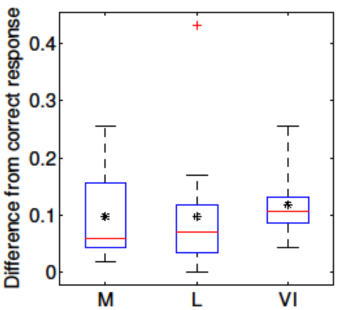


Fig. 1. The box plots representing the relative deviation of the responses from the correct number for each condition (M – memory, L – language, VI – visual imagery)

3.2 Neurophysiological Data

The data consists of brain signals recorded from 17 participants (14 right-handed) of which one set (right-handed) has been removed due to high noise probably caused by improper recording. The grand average difference between each pair of conditions is analyzed by the signed biserial correlation coefficient, sgn r^2, and presented in Fig. 2. The top plots show the spatial distribution on scalp of the sgn r^2 values at the intervals represented by the shaded areas in the underneath time plots. The highest discrimination between classes is observed in the centro-parietal area, corresponding to the ERP component, P300. The amplitude and duration of the P300 varies between conditions and reflects the different cognitive processing of the stimulus [9]. The most prominent difference can be observed as a positive deviation which starts around 300 ms and gradually decreases until about 1000 ms, depending on the condition.

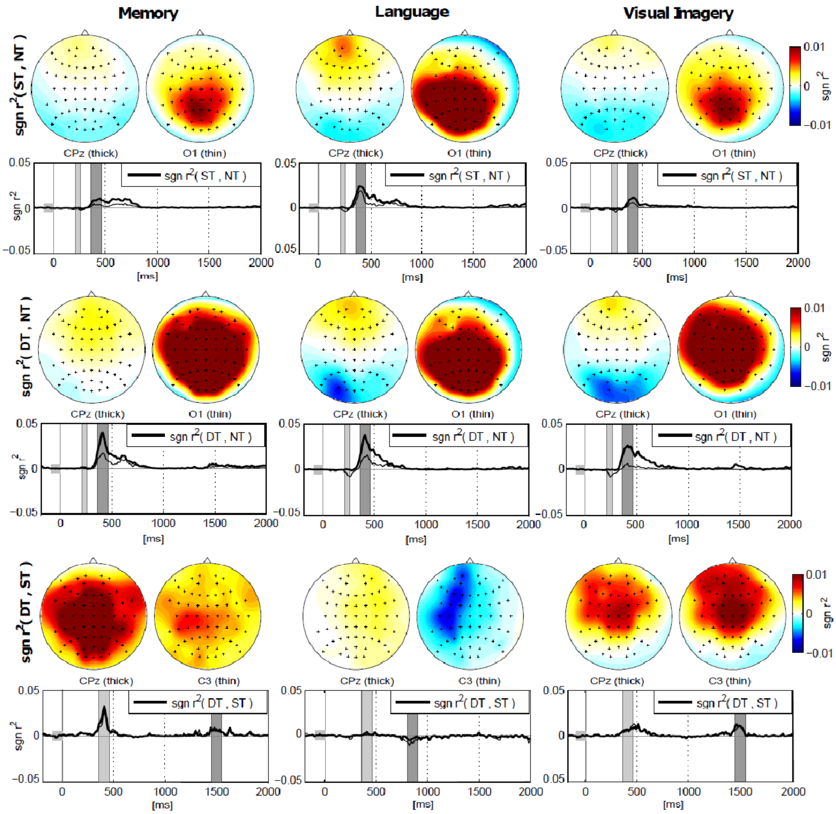


Fig. 2. Grand-average of the event-related potentials for all conditions:discrimination between ST – NT and DT – NT at channel CPz (thick line) and O1 (thin line), followed by DT – ST at channel CPz (thick) and C3 (thin), quantified by the signed biserial correlation coefficient

Exclusively for the language condition, we observe left lateralized scalp topography, which is in line with the literature that reports that language areas are mostly on the left hemisphere [10]. Comparing the amplitude evolution of the deep level of processing to the shallow one, we can still observe activation in the late interval: a negative difference in language condition around 850ms and a positive difference, observable for example at 1500ms for memory and visual imagery processes.

4 Classification

The classification performance based on spatio-temporal features is presented in Fig. 3 by the bar plots based on the area under the roc curve values. The grey antennas on top indicate the standard error of the mean, SEM. The classification performances were averaged over trials and subjects, giving a high overall performance above chance level. A performance over 0.75 was obtained between non-targets and shallow targets, with the highest values for the language condition. The same thread is visible between non-targets and deep targets, but with higher mean performance up to 0.85. The mean classification performance is under 0.7 between shallow and deep processing, which corresponds to the smaller difference of the sgn r^2 in Fig. 2.

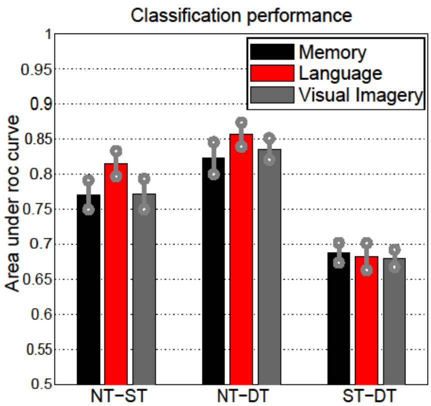


Fig. 3. Pairwise classification performance between classes (NT – non-targets, ST – shallow targets, DT – deep targets)

5 Discussion and Conclusion

ERP analysis points out neural markers that differ with the levels of processing. In the sgn r^2 scalp plots we observe higher ST-NT difference in the language case, which might be explained by a general lower level of workload in this task (memory and visual imagery were reported to be more complex by participants). The classification results demonstrate that the levels of cognitive processing can be distinguished by multivariate data analysis applied to ERPs on a single-trial basis, up to 0.85 performance. In this first approach we used a binary classification on the three cognitive levels. Future developments could exploit continuous measures like regression analysis. We will also consider extending the classification interval to 2000 m, since the sgn r^2 plots show important discriminative information for the deep process also in the late timing which is due the fact that a deeper cognitive processing requires more time to be fulfilled, longer than the stimulus presentation.

In conclusion we show that the levels of cognitive processing can be well differentiated by analyzing the neurophysiological features. If the BCI-system is able to infer the cognitive user state, it could use this information to adapt the computer interface accordingly. In this regard, the concept of monitoring the level of cognitive processing can complement human capabilities with a non-invasive BCI system like for human-computer interaction applications.

Acknowledgement. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 611570, supported by the BMBF under contract 01GQ0850 and financed by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreement POSDRU/159/1.5/S/134398.

References

1. Venthur, B., Blankertz, B., Gugler, M.F., Curio, G.: Novel applications of BCI technology: psychophysiological optimization of working conditions in industry. In: Proceedings of the 2010 IEEE Conference on Systems, Man and Cybernetics (2010)
2. Nicolae, I.E., Acqualagna, L., Blankertz, B.: Neural indicators of the depth of cognitive processing for user-adaptive neurotechnological applications. In: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Milan, Italy, August 25–29 (2015)
3. Craik, F.I.M., Tulving, E.: Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General **104**, 268–294 (1975)
4. Polich, J., Kokb, A.: Cognitive and biological determinants of P300: an integrative review. Biological Psychology **41**, 103–146 (1995)
5. Chen, Y.N., Mitra, S., Schlaghecken, F.: Sub-processes of working memory in the N-back task: an investigation using ERPs. Clin Neurophysiol **119**, 1546–1559 (2008)
6. Delorme, A., Makeig, S.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods **134**, 9–21 (2004)
7. Winkler, I., Haufe, S., Tangermann, M.: Automatic classification of artifactual ICA-components for artifact removal in EEG signals. Behavioral and Brain Functions, **7**(30) (2011)
8. Blankertz, B., Lemm, S., Treder, M., Haufe, S., Muller, K.R.: Single-trial analysis and classification of ERP components a tutorial. NeuroImage **56**(2), 814–825 (2011)
9. Polich, J.: Task difficulty, probability, and inter-stimulus interval as determinants of P300 from auditory stimuli. Electroencephalography and Clinical Neurophysiology **68**(4), 311–320 (1987). Elsevier
10. Griggs, R.A.: Psychology: A Concise Introduction. p. 69 (2012)

Brain–Robot Interfaces Using Spatial Tactile BCI Paradigms

Symbiotic Brain–Robot Applications

Tomasz M. Rutkowski[1,2(✉)], Kensuke Shimizu[1], Takumi Kodama[1], Peter Jurica[2], and Andrzej Cichocki[2]

[1] Life Science Center of TARA, University of Tsukuba, Tsukuba, Japan
tomek@bci-lab.info
http://bci-lab.info/
[2] RIKEN Brain Science Institute, Wako-shi, Japan

Abstract. Two novel approaches to a direct brain–robot interface using tactile brain–computer interface (BCI) technologies are presented in the paper. We propose to utilize two previously developed by our team stimulus driven BCI paradigms, which are based on tactile pin pressure and full body vibrotactile modalities. The user intentions are decoded from the brainwaves in real time and translated to a symbiotic humanoid robot NAO navigation. A communication protocol between the BCI output and the robot is realized in a symbiotic brain–robot communication scenario using an user datagram protocol (UDP). Results obtained from healthy users reproducing simple brain–robot control tasks support the research hypothesis of the possibility to interact with robotic devices using symbiotic BCI technologies.

Keywords: Brain–computer interfaces · Brain–robot interfaces · Symbiotic brain–robot interaction

1 Introduction

A brain-computer interface (BCI) is a technology that decodes neurophysiological signals of a user to allow a direct thought–based communication with others or a control of external devices (e.g. a direct brain–robot interface) without any body muscle activities [13]. The majority of BCI applications are based on a visual [1] or auditory [2,10] modalities. However, the tactile BCI [8] seems to offer the better communication options in comparison with visual and auditory modalities in case of locked?-in?-syndrome (LIS) patients [5,8]. The two BCI–based direct brain–robot interfaces reviewed in this paper are employing brain event related responses (ERPs) to tactile modality. Namely those are the developed by our team tactile pin pressure (tpBCI) [12] and full body vibrotactile (fbvBCI) [3] BCIs.

K. Shimizu and T. Kodama contributed equally.

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 132–137, 2015.
DOI: 10.1007/978-3-319-24917-9_14

We present the recently developed by our team tactile BCIs paradigms applied for online control of a humanoid robot NAO in the brain–robot symbiotic configurations utilizing the UDP network protocol. Brainwave responses in the all presented online BCIs, tested with two healthy users, are captured with eight active EEG electrodes g.LADYbird connected to g.USBamp portable amplifier from g.tec medical instruments GmbH, Austria. The experiments are conducted in oddball style paradigm eliciting the P300 responses, classified next with stepwise linear discriminant analysis (SWLDA) method [4], and not exceeding four steps averaging procedures in online experiments. The users first have to learn command mappings associated hand six pin positions, in tpBCI case, or larger body areas for fbvBCI. The tpBCI and fbvBCI experiments are conducted in six command robot control set up (go straight, back, left, right, sit down, and say goodbye).

The paper from now on is organized as follows. In the following section we describe materials and methods used in the study. Next the obtained results are presented. Results discussion and conclusions summarize the paper.

2 Materials and Methods

The experiments reported in this paper were performed in the Life Science Center of TARA, University of Tsukuba, Japan. All the details of the two reviewed experimental procedures and the research targets of the BCI–based brain–robot control paradigms were explained in detail to the participating users, who agreed voluntarily to participated in the study. The EEG experiments were conducted in accordance with *The World Medical Association Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects*. The experimental procedures were approved and designed in agreement with the ethical committee guidelines of the Faculty of Engineering, Information and Systems at University of Tsukuba, Tsukuba, Japan (experimental permission no. 2013R7).

2.1 Brain–robot Control with Tactile Pin–pressure BCI

In this first reviewed project the tactile stimuli were generated via the tactile pin–pressure generator composed of nine solenoids arranged in the 3×3 matrix [12]. There were six linear patterns of tactile pin–pressure stimuli delivered in random order to the user fingers. Three of them were horizontal lines ordered from the top to bottom of user?s fingers respectively. The remaining patterns were the vertical lines in left to right positions order. The solenoids generated pin–pressures 100 ms long each time. The users performed first short psychophysical experiments with visual command feedback in order to familiarize themselves with robot command mappings. The commands were sent to the robot using the UDP network protocol with a wireless connection. The successfully classified P300 responses to stimuli delivered to user's hand (intended robot commands) were delivered to the robot for an execution of pre–programmed movements.

In the training BCI experiments, first conducted with digits one to six spellings (users had to reproduce command sequences), EEG signals were captured with from eight active wet EEG electrodes. Those were attached to the head locations *Cz, Cpz, P3, P4, C3, C4, CP5,* and *CP6.* A reference electrode was attached to a left earlobe and a ground electrode on the forehead at *FPz* position respectively. The set up reproduced our previously published study [11]. The users put on polyethylene gloves to limit any electric interference possibly leaking from the tactile stimulators. The users were also requested to limit their eye–blinks and body movements to avoid electromagnetic and electromyography (EMG) interferences. The EEG signals were recorded and preprocessed by an in–house enhanced BCI2000–based application [9], using a stepwise linear discriminant analysis (SWLDA) classifier [4] with features drawn from ERP intervals of $0 \sim 800$ ms. The trained SWLDA classifier parameters were next applied for online BCI experiments using the BCI2000 environment. The EEG capture sampling rate was set to 256 Hz, the high pass filter at 0.1 Hz, and the low pass filter at 40 Hz. The ISI was 400 ms and each stimulus? duration was of 100 ms.

The NAO robot was next controlled using six pre–programmed commands decoded and classified by the tpBCI. The commands were transmitted in form of numbers one to six (representing the movement commands to be executed) via a wireless network using UDP protocol. The tpBCI experimental setup is depicted in Figure 1 and a demo video from the online direct brain–robot interfacing experiment is available at [7].

2.2 Brain–robot Control with Full Body Vibrotactile BCI

The final reviewed research project reported in this paper was based on application of vibrotactile generators applied to user shoulders, arms, backs and the both legs in order to create a somatosensory response–based oddball BCI paradigm [8] as described in detail in [3]. Active EEG electrodes were attached to the sixteen locations *Cz, Pz, P3, P4, C3, C4, CP5, CP6, P1, P2, POz, C1, C2, FC1, FC2 and FCz*, as in 10/10 international system. A reference electrode was attached to the left mastoid, and a ground electrode to the forehead at the *FPz* position. The EEG signals were captured and classified by BCI2000 software [9] using a stepwise linear discriminant analysis (SWLDA) classifier [4] applied to the $0 \sim 800$ ms ERP time range latencies. The EEG recording sampling rate was set at 512 Hz, and the high and low pass filters were set at 0.1 Hz and 60 Hz, respectively. The notch filter to remove power line interference was set for a rejection band of $48 \sim 52$ Hz.

In each trial, the stimulus duration was set to 250 ms and the ISI to random values in a range of $350 \sim 370$ ms in order to break rhythmic patterns of presentation. Each online experiment comprised of 10 trials used for epochs averaging in the classifier training sessions. In the online brain–robot control scenarios four responses averaging set up was used for a more smooth interaction.

The vibrotactile spatial pattern stimuli (stimulated body locations used as cues) were generated using the same *MAX 6* program, and the trigger onsets

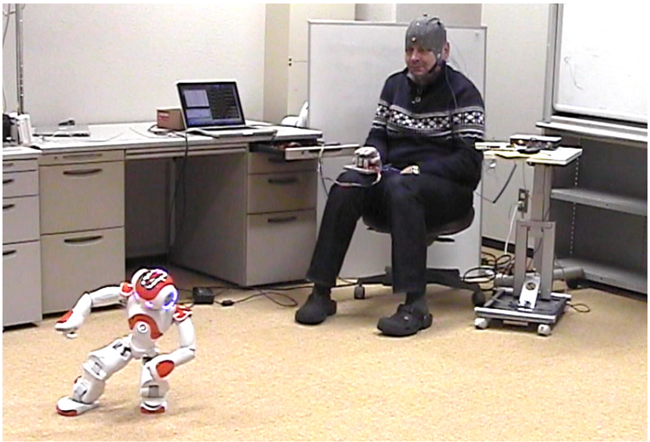


Fig. 1. The brain–robot control example using the tactile pin–pressure BCI–based NAO humanoid robot navigation in the brain–robot configuration. A demo video is available online at [7].

were generated by BCI2000 EEG acquisition and ERP classification software [9]. During the training run the cues were given in form of vibrations of the target location before each sequence. In testing phase the six robot control commands of walking straight, back, left, right; siting down; and saying goodbye were transmitted from fbvBCI to the NAO using also via wireless network using UDP protocol.

The fbvBCI experimental setting is depicted in Figure 2. A video from the online brain–robot interfacing experiment is available at [6].

3 Results

The two, described in the above sections, direct brain–robot interfacing techniques utilizing tactile modality experiments resulted with successful online control outcomes as depicted in Figures 1 and 2. The videos documenting the presented results are available online at [6,7]. The online tpBCI accuracies in offline training scenarios of the five users scored well above the chance level of 16% and around 70% in six digits spelling test on average, which was an encouraging outcome of the tested prototype. Based on the obtained accuracies we trained the SWLDA classifier in cross–validation setting for the final brain–robot control evaluation. The users were able to control perfectly (100% accuracy scores) the humanoid robot, as documented online [7], using six commands interfacing set up. The six digit spelling fvbBCI–based pilot experiment classification accuracies resulted also above chance level of 16% of the three participating users (67% on the average of the three participants). The above user results allowed to train in cross–validation scenario the SWLDA classifiers for target responses

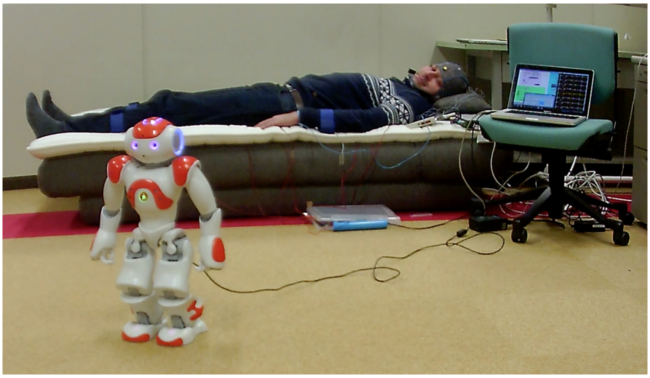


Fig. 2. The brain–robot control using the body–tactile BCI–based NAO humanoid navigation in the symbiotic brain–robot configuration. A demo video is available online at [6].

classification in oddball settings for each user separately. The online direct brain–robot interfacing using the fvbBCI experiments were performed as depicted in Figure 2 (the documentary online video available at [6]). Also in this final BCI case, the participating users were able to score with perfect accuracies of 100% while controlling the humanoid NAO robot using six commands.

4 Conclusions

Two direct brain–robot interfacing scenarios have been discussed in this paper. In order to realize the purpose, we aimed to test practical applications of two previously developed by our team tactile BCIs. This paper reported a successful implementation of the six commands–based direct and symbiotic brain–robot BCIs in online control scenarios. We conducted experiments to verify the feasibility and user experience of the proposed two tactile BCI–based methods. According to the results obtained from the participating users, even if offline training classification accuracies were not perfect (though clearly above the chance levels), the users were able to improve to achieve the perfect (100% scores) during the online robot control trials. Those final results shall contribute to the BCI–training related brain plasticity, which will be a future target of our research. For future research, we aim to validate the prototypes and further explore the brain and BCI co–adaptivity with more users to additionally proof the results. We will also aim at reduction of the brainwave acquisition and analysis time windows for even better usability of the proposed direct brain–robot interfacing paradigms.

References

1. Chang, M., Nishikawa, N., Struzik, Z.R., Mori, K., Makino, S., Mandic, D., Rutkowski, T.M.: Comparison of P300 responses in auditory, visual and audiovisual spatial speller BCI paradigms. In: Proceedings of the Fifth International Brain-Computer Interface Meeting 2013, p. Article ID: 156. Graz University of Technology Publishing House, Austria, Asilomar Conference Center, Pacific Grove, CA USA, June 3–7, 2013. http://castor.tugraz.at/doku/BCIMeeting2013/156.pdf
2. Chang, M., Mori, K., Makino, S., Rutkowski, T.M.: Spatial auditory two-step input Japanese syllabary brain-computer interface speller. Procedia Technology **18**, 25–31 (2014). http://www.sciencedirect.com/science/article/pii/ S2212017314005283
3. Kodama, T., Makino, S., Rutkowski, T.M.: Spatial tactile brain-computer interface paradigm applying vibration stimuli to large areas of user's back. In: Mueller-Putz, G., Bauernfeind, G., Brunner, C., Steyrl, D., Wriessnegger, S., Scherer, R. (eds.) Proceedings of the 6th International Brain-Computer Interface Conference 2014, pp. Article ID 032-1-4. Graz University of Technology Publishing House (2014)
4. Krusienski, D.J., Sellers, E.W., Cabestaing, F., Bayoudh, S., McFarland, D.J., Vaughan, T.M., Wolpaw, J.R.: A comparison of classification techniques for the P300 speller. Journal of Neural Engineering **3**(4), 299 (2006). http://stacks.iop.org/1741-2552/3/i=4/a=007
5. Patterson, J.R., Grabois, M.: Locked-in syndrome: a review of 139 cases. Stroke **17**(4), 758–764 (1986). http://stroke.ahajournals.org/content/17/4/758.abstract
6. Rutkowski, T.M.: Tactile-body BCI-based NAO robot control. http://youtu.be/sn6OEBBKsPQ
7. Rutkowski, T.M.: Tactile-pressure BCI-based NAO robot control. http://youtu.be/dSrZ5O59vhI
8. Rutkowski, T.M., Mori, H.: Tactile and bone-conduction auditory brain computer interface for vision and hearing impaired users. Journal of Neuroscience Methods **244**, 45–51 (2015). http://www.sciencedirect.com/science/article/pii/S0165027014001265, brain Computer Interfaces; Tribute to Greg A. Gerhardt
9. Schalk, G., Mellinger, J.: A Practical Guide to Brain-Computer Interfacing with BCI2000. Springer-Verlag, London (2010)
10. Schreuder, M., Blankertz, B., Tangermann, M.: A new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue. PLoS ONE **5**(4), e9813 (2010)
11. Shimizu, K., Makino, S., Rutkowski, T.M.: Inter-stimulus interval study for the tactile point-pressure brain-computer interface. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. (accepted, in press). IEEE Engineering in Medicine and Biology Society. IEEE Press, August 25–29, 2015. http://arxiv.org/abs/1506.04458
12. Shimizu, K., Mori, H., Makino, S., Rutkowski, T.M.: Tactile pressure brain-computer interface using point matrix pattern paradigm. In: 15th International Symposium on Soft Computing and Intelligent Systems (SCIS), 2014 Joint 7th International Conference on and Advanced Intelligent Systems (ISIS), pp. 473–477, December 2014. doi:10.1109/SCIS-ISIS.2014.7044756
13. Wolpaw, J., Wolpaw, E.W. (eds.): Brain-Computer Interfaces: Principles and Practice. Oxford University Press (2012)

Joint Stiffness Tuning of Exoskeleton Robot H2 by Tacit Learning

Shingo Shimoda[1(✉)], Álvaro Costa[2], Guillermo Asin-Prieto[3], Shotaro Okajima[1,4], Eduardo Ináẽz[2], Yasuhisa Hasegawa[4], Jose M. Azorín[2], Jose L. Pons[3], and Juan C. Moreno[3]

[1] RIKEN, Brain Science Institute, Wako, Japan
shimoda@brain.riken.jp
[2] Brain-Machine Interface Systems Lab, Miguel Hernández University of Elche, Elche, Spain
[3] Neural Rehabilitation Group, the Spanish National Research Council, Madrid, Spain
[4] Nagoya University, Nagoya, Japan

Abstract. Joint stiffness of the exoskeleton robot is one of the most important factors to support bipedal walking. In this paper, we discuss the robot joint stiffness tuning algorithm using the bio-mimetic learning method called tacit learning. We experimentally showed that the proposed controller can tune the joint stiffness of the exoskeleton robot by tuning the integral gain in the controller. The walking experiment wearing the exoskeleton robot suggest that the stiffness tuning is applicable to control the walking speed.

Keywords: Exoskeleton robot · Walking support · Tacit learning

1 Introduction

In order to support humans behaviors by using exoskeleton robots, how to decode the wearers' motion intention is one of the most important factors. Our approach is to synchronize the exoskeleton robot controller with the human control systems based on the biological control principle.

We mimic the two features of biological controllers for the synchronization between these two controllers. One of the features is the data analysis structure. It is said that biological control systems can be modeled as *bow-tie structure*[1][2] described in Fig. 1a. Bow-tie structure represents the essence of biological control systems where there are great diversity of inputs and outputs while a smaller diversity of protocol is used to connect these diverse outputs and inputs. The inputs here imply the sensor signals and the outputs imply the signals to the body effectors like muscles. In the case of the patients who need prosthetic devices or exoskeleton robot supports, bow-tie structure is described in Fig. 1b where some of the parts both in inputs and outputs are missing. Our idea to synchronize the patient control systems and the exoskeleton robot controller

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 138–144, 2015.
DOI: 10.1007/978-3-319-24917-9_15

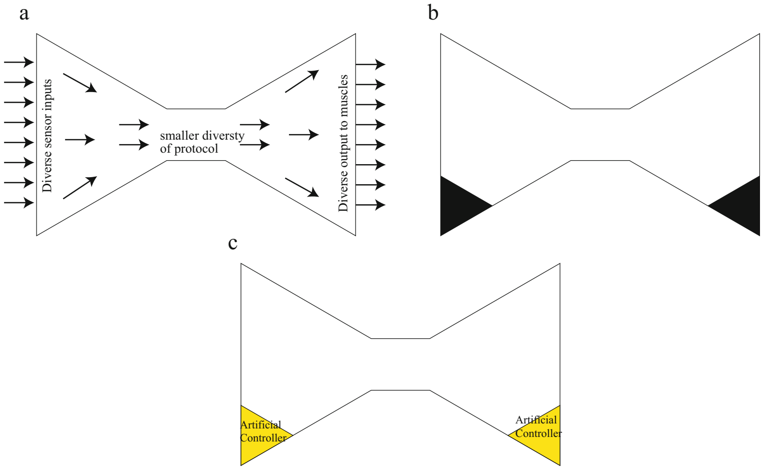


Fig. 1. a: Bow Tie Structure representing the biological control structure, b : State when we lose parts of our bodies, c : Proposed idea of controller to adapt exoskeleton robot motions to humans' motions

is the placement of the robot controller at the missing points. The detail how we place the robot controller is discussed in the next section.

Another important feature of biological control systems is the capability to adapt their behaviors to the environment through body/environment interactions. To synchronize the robot controllers with the human control systems, the exoskeleton controller should have the same capability. The authors have proposed the bio-mimetic learning architecture called *tacit learning* [3][4]. Tacit learning can tune the roughly-defined robot behaviors to the sophisticated ones that are adapted to the environment.

We apply these two biological features to control the lower limb exoskeleton controllers for walking support. In this paper, we focus on the joint stiffness tuning by tacit learning through human body-exoskeleton robot interactions. We first discuss in Section 2 the features of bow-tie structure and how we can design the walking support controllers based on the two biological features. The control algorithm that tune the joint stiffness interacting with the wearer's motions is also discussed in Section 2. In Section 3, we show the preliminary experimental results with healthy subjects. In Section 4, we conclude this paper.

2 Exoskeleton Controller with Bow-tie Structure

2.1 Features of Bow-tie Structure

Bow-tie structure is used to represent the feature of biological controllers where there are great diversity of inputs and outputs while a smaller diversity of protocol is used to connect them. In bow-tie structure, the input signals are grounded

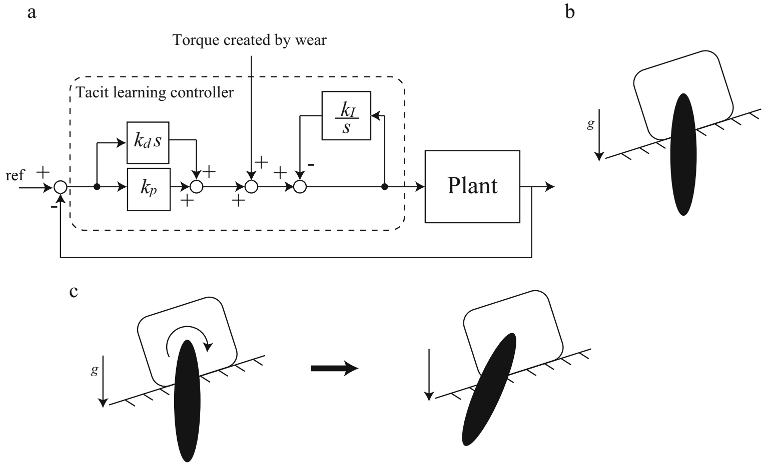


Fig. 2. a:Block diagram of proposed controller b: Posture where the joint angle is followed to the direction of gravity

on some symbols to reduce the diversity, and the output signals are created from the symbols. When we lose a part of our bodies, we lose not only the sensors and the muscles but also the processes of reducing and increasing the signal diversities in bow-tie structure as described in Fig. 1b. To develop the robot controller that can place at these missing parts of humans controller as described in Fig. 1c, the controller must have the following two features:

1. The controller should communicate with human controller at the higher level than the output from human controller such as EMG.
2. The controller should adapt to the both the human controller and the environment.

The first feature is required to know the motion intention at the appropriate level. In the previous study for the development of the forearm prosthesis[5], for instance, we measured the shoulder and the elbow joint angles to know how much the forearm prosthesis user want to rotate the wrist joint.

The second feature is required according to the biological feature where the adaptation to the environment can be observed in the various level of the brains. To synchronize to humans controller, the robot controller should be adapted to the other parts of the brains and the environments.

2.2 Controller for Tuning Joint Stiffness

To support the walking by lower limb exoskeleton robot, we tune the joint stiffness using the controller placed on the part of the bow-tie structure. In this paper as the preliminary discussion for the walking support, we propose the tacit learning controller that can tune the joint stiffness as shown in Fig. 2a. This controller can be described as follows:

$$\tau = -k_p\theta - k_d\dot{\theta} - f \tag{1}$$

$$f = k_I \int \tau dt. \tag{2}$$

Here, τ and θ imply the joint torque and the joint angle, respectively. k_p, k_d and k_I are the control gains.

In the previous study[6], we showed that this controller can create the posture where the robot joint follow the direction of gravity as illustrated in Fig. 2b. When we apply this controller for the exoskeleton joint, not only the gravity but also the torque created by the wearer is worked to the joint. In this case, the joint angle is controlled to follow the wear's torque as shown in Fig. 2c even the joint back-drivability is very low.

The important feature of this controller is that the rate to follow the external torque is tuned by the gain of the integrator. When the gain becomes larger, the joint more quickly follow the external torque, and *vice versa*. When the robot joint follows the external torque quickly, we can think that the stiffness of the joint get lower. In this paper, we experimentally show that the joint stiffness is changed by tuning the gains of the proposed controller.

3 Experimental Results with Exoskeleton Robot

3.1 Exoskeleton Robot for Experiments

We use the exoskeleton robot *H2* for the experiments as shown in Fig. 3a. H2 has 3 joints in each leg that can support the ankle, knee and hip joints in sagittal plane. Please see Reference [7] for the detail mechanism and the control systems. One important feature here is that the joint back-drivability is very low such that almost 40 Nm external torque is needed to move the joint.

3.2 Results in Stiffness Tuning Experiments

First, we tested the joint stiffness turning by the proposed controller through body/exoskeleton interactions. In this experiments, we focused on the right knee joint motions. The movies of the experimental performances can be seen in [8]. $k_p = 20.0$ and $k_d = 20.0$ were used for the right knee control. Three different k_Is, 0.1, 0.05 and 0.02, were used to change the stiffness. The experimental results are shown in Fig. 3c-d and in the movies. In the case of the largest gain $k_I = 0.1$ that implies the lowest joint stiffness, the subject moved his knee joint smoothly with small EMG of the rectus femoris muscle. The maximum exoskeleton joint torque was 1.0 Nm. The EMG results and the exoskeleton joint torque showed that the joint stiffness got larger as the integral gain k_I got smaller. From these results, we can say that the controller proposed in Fig. 2b can tune the joint stiffness through body/robot interactions.

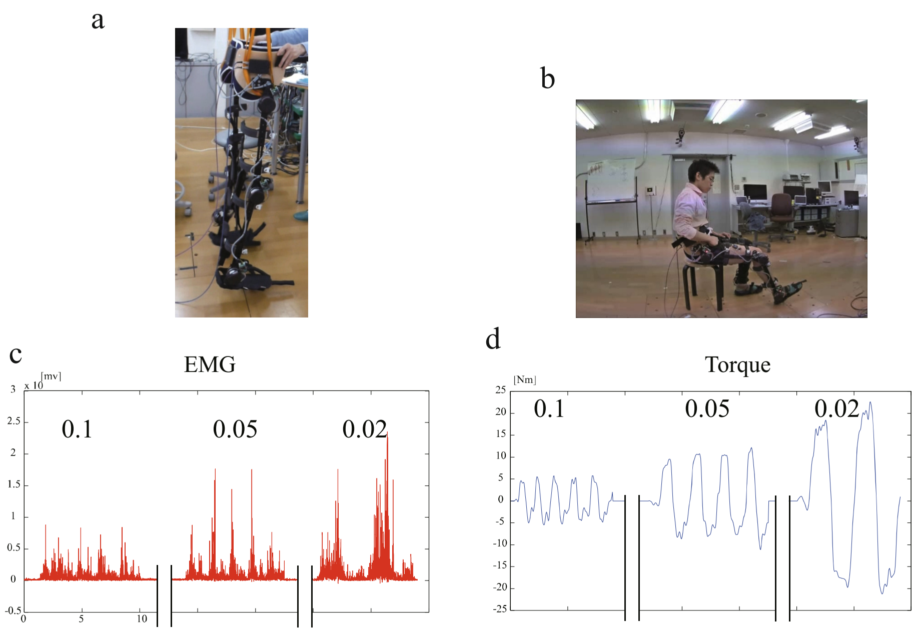


Fig. 3. a: Overview of Exoskeleton Robot H2, b: Overview of One joint stiffness tuning experiment, c: EMG of rectus femoris muscle during experiments. When we used the larger integral gain, EMG got smaller suggesting the joint stiffness got smaller. d: Knee joint torque of H2 during the experiments. The results also showed the joint stiffness was tuned by changing the integral gain.

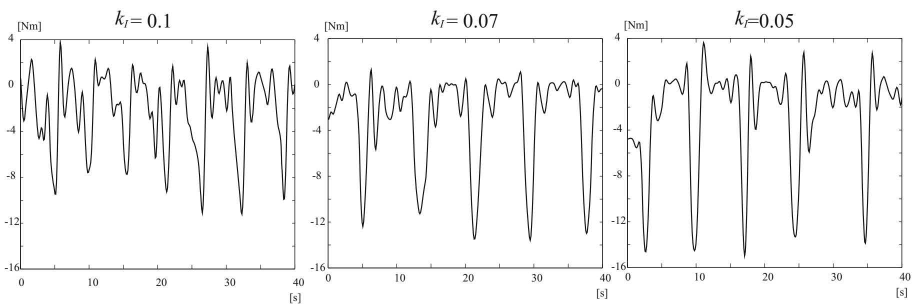


Fig. 4. Change of right knee joint torque depending of integral gain during walking with H2

3.3 Walking with Low Stiffness

We set the same controller to the all joints of H2 and asked the subject to walking wearing H2. Three different k_Is, 0.1, 0.07 and 0.05, were used to change the stiffness. In the case of k_I=0.02 that was used in the previous one joint experiment, the stiffness was too high that the subject could not walk. The right knee joint torques during walking were illustrated in Fig. 4. These results showed that the proposed controller successfully tuned the joint stiffness by tuning the integral gains as was the case with the one joint experiments. Another important result from this experiment was the change of walking speed depending on the joint stiffness. The subject walked 2.0 s for one step in the case k_I=0.1 while the subject walked 4.2 s for one step in the case of $k_I = 0.05$. There results suggest that the stiffness tuning would be used to tune the walking speed.

4 Conclusion

To support the motions by using exoskeleton robots, brains and the robot controller should be well synchronized for the good support. In this paper, we proposed the exoskeleton robot controller based on the two biological system features. One feature is the data analysis mechanisms called bow-tie structure, and the other is the behavior adaptation architecture called tacit learning. The proposed controller can change the joint stiffness by tuning the integral gain. We experimentally showed that the joint stiffness was changed when the wearer created the external torque to the joint.

Obviously, the joint stiffness tuning is not enough to support the walking of the patients. Our next step is how to use this controller to support patient walking. As shown in the walking experiments, the walking speed was changed by tuning the joint stiffness. If the controller can modify the stiffness as the wearer want, the walking speed is adjusted as the wearer want even if the wearer cannot move his leg by himself. Further discussions are needed to develop the controller for the walking support as users want.

Acknowledgments. This study has been funded by grant from the European Commission, within the Seventh Framework Programme (IFP7- ICT-2013-10-611695: BioMot - Smart Wearable Robots with Bioinspired Sensory-Motor Skills).

References

1. Scete, M., Doyle, J.: Bow ties, metabolism and disease. Trends Biotechnol. **22**(9), 446–450 (2004)
2. Alnajjar, F.S.K., Itkonen, M., Berenz, V., Tournier, M., Shimoda, S.: Sensor synergy as environmental input integration. Frontier in Neuroscience **15**, article 436 (2015)
3. Shimoda, S., Kimura, H.: Bio-mimetic Approach to Tacit Learning based on Compound Control. IEEE Transactions on Systems, Man, and Cybernetics- Part B **40**(1), 77–90 (2010)
4. Shimoda, S., Yoshihara, Y., Kimura, H.: Adaptability of tacit learning in bipedal locomotion. IEEE Transactions on Autonomous Mental Development **5**(2), 152–161 (2013)
5. Iwatsuki, K., Oyama, S., Alnajjar, F.S., Shimoda, S., Yamamoto, M., Hirata, H.: Hand prosthesis equipped with a tacit learning system to auto-regulate forearm rotation. In: The 69th Annual Meeting of the American Society for Surgery of the Hand (2014)

6. Shimoda, S., Yoshihara, Y., Fujimoto, K., Yamamoto, T., Maeda, I., Kimura, H.: Stability analysis of tacit learning based on environmental signal accumulation. In: Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (2012)
7. Bortole, M., Venkatakrishnan, A., Zhu, F., Moreno, J.C., Francisco, G.E., Pons, J.L., Contreras-Vidal, J.L.: The H2 robotic exoskeleton for gait rehabilitation after stroke: early findings from a clinical study. Journal of NeuroEngineering and Rehabilitation (2015)
8. http://btcc.nagoya.riken.jp/biologic/biomot.html

Human Computer Interaction Meets Psychophysiology: A Critical Perspective

Michiel M. Spapé[1(✉)], Marco Filetti[1], Manuel J.A. Eugster[1,2], Giulio Jacucci[3], and Niklas Ravaja[1,4]

[1] Helsinki Institute for Information Technology HIIT, Aalto University, Espoo, Finland
{michiel.sovijarvi-spape,marco.filetti,manuel.eugster}@hiit.fi
[2] Department of Computer Science, Aalto University, Espoo, Finland
[3] Department of Computer Science, Helsinki Institute for Information Technology HIIT, University of Helsinki, Helsinki, Finland
{giulio.jacucci,niklas.ravaja}@helsinki.fi
[4] Department of Social Research, University of Helsinki, Helsinki, Finland

Abstract. Human computer interaction (HCI) groups are more and more often exploring the utility of new, lower cost electroencephalography (EEG) interfaces for assessing user engagement and experience as well as for directly controlling computers. While the potential benefits of using EEG are considerable, we argue that research is easily driven by what we term naïve neurorealism. That is, data obtained with psychophysiological devices have poor reliability and uncertain validity, making inferences on mental states difficult. This means that unless sufficient care is taken to address the inherent shortcomings, the contributions of psychophysiological human computer interaction are limited to their novelty value rather than bringing scientific advance. Here, we outline the nature and severity of the reliability and validity problems and give practical suggestions for HCI researchers and reviewers on the way forward, and which obstacles to avoid. We hope that this critical perspective helps to promote good practice in the emerging field of psychophysiology in HCI.

Keywords: HCI · EEG · Psychophysiology · Reliability · Validity · Naïve Neurorealism

1 Outline

In the following subsections, we will first briefly summarise the history of EEG and describe the rise in use of psychophysiology in HCI. Following, we will discuss that the reduced costs of equipment as well as the increased popular appeal of neuroscience are likely reasons behind the explosive growth of interest. However, we argue that what we term naïve *neurorealism* can lead to unsubstantiated optimism. In short, this concerns the idea that use of psychophysiological measurements necessarily enables objective knowledge of the mind, and thereby must

This work was partly supported by TEKES (Re:Know 2, decision 5159/31/2014).

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 145–158, 2015.
DOI: 10.1007/978-3-319-24917-9_16

lead to a high degree of insight and user-control. We explain this and illustrate the point and improve discussion by first outlining a hypothetical example of an application scenarios: the BrainGuitar. The device is, to our knowledge, purely fictional and merely serves here to illustrate some of the more serious caveats that occur in the field. We explain how the known methodological aspects of reliability and validity as pertaining to psychophysiological measurements undermine the credibility of inference of mental states. In particular, the weak signal to noise ratio of EEG is discussed, and how strongly this is affected by artifacts. Finally, we provide guidelines for scientists who consider the utility of psychophysiological measurements as well as reviewers who assess the contributions of others.

2 EEG in HCI

In 1929, Hans Berger [1] dramatically showed how EEG can enable us to non-invasively measure human brain activity at a high temporal accuracy. Berger was also the first to discover alpha waves, one of the most prominent features of the EEG. Alpha waves are easily observed as a stereotypical oscillation in the range of 8 Hz–12 Hz that can be observed over much of the scalp (for a history of Berger's work, see [2]). As they appear in the absence of prominent stimuli, they are often used as an index of relaxation, or brain inactivity (however, observation of alpha waves alone is not a sufficient condition to conclude that a reduction of brain activity took place [3]. We will come back to the issue of performing reverse inferences in Sec. 4.6).

Another important discovery in the field was the P300, a brain signal initially observed in concomitance with the presentation of an unexpected stimulus. Discovered in 1965 [4], it manifests itself as a large positive potential starting at approximately 300 ms post stimulus. The P300 is currently believed to indicate saliency (due to an interaction between attention and memory, under one hypothesis [5]). Due to its characteristics (relatively high amplitude and reliability), the P3 allowed researchers to develop the P3 speller, the first working instance of a Brain-Computer Interface (BCI) [6]. This is particularly useful for patients with serious disabilities, such as the locked-in syndrome, for whom a BCI may be the most efficient way to communicate.

As computer technology increased in quality and availability, EEG became more and more available across domains. The relative ease in which raw brain-related signals can be obtained and analysed led to rapid developments. Researchers and practitioners are now able to develop tools based on EEG and other physiological measurements with common electronic devices. It is now, for example, possible to implement a Brain-Controlled address book based on the P3 speller concept even on mobile phones [7].

2.1 The Rise of Brain Informed Human Computer Interaction

In HCI, EEG is used for various purposes. Affective computing [8] and physiological computing [9] are two intertwined branches of this field: in both, cognitive

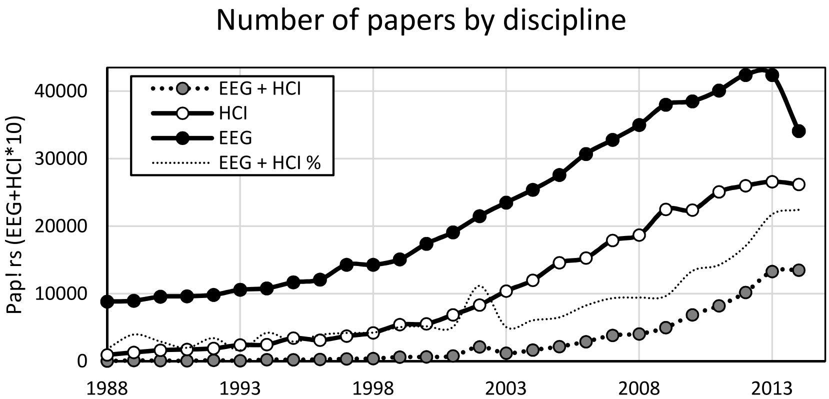


Fig. 1. Growing use of EEG in HCI. The graph displays the growth of published papers in EEG, HCI and the combination thereof, showing greater growth of EEG in HCI (EEG + HCI) than either of its constituent parts, as shown by the growing proportion of EEG in HCI as a function of EEG papers (EEG + HCI %).

and emotional states are predicted or classified based on their physiology (in affective computing, the emotional state is of particular interest). Both investigate how systems should adapt to detected changes in users' own states. For example, it has been attempted to measure task engagement by using EEG alpha asymmetry, i.e. the difference in alpha power between the two hemispheres [10]. In this kind of research, two simultaneous assumptions are made: firstly, alpha power correlates to reduced neural activity and, secondly, greater left than right activity corresponds to positive emotions, and / or high motivation. Another area worth of note is neurofeedback: this area investigates the possibility of building systems that take advantage of "tight feedback loops", so that users' (or patients', for medical applications) cognitive states in a predefined direction. For example, to remain in the domain of alpha oscillations, elderly patients have been trained to increase their peak alpha (10 Hz–11 Hz) power, which was found to be associated with increases in their processing speed and executive function [11]. However, it remains a debated issue that claims regarding the success of neurofeedback may be overblown for marketing purposes, and as many of these systems are commercial, a conflict of interest could be present [12].

The interest of human computer interaction research in psychophysiological data grew exponentially over the last few years. This can be easily seen in Fig. 1, which maps the number of papers (as indexed by Google Scholar) per year from 1988 (the first EEG brain computer interface [6]) to 2014. Of course, scientific production in general continues to grow [13], but it is fair to say that up until 1993, EEG was a fairly uncommon interest for HCI, with ca. 0.6% of publications. However, it seems that the combination of HCI and EEG took off in the subsequent years, roughly doubling in successively 1.5 (1994-1995: 24), 4 (1998-1999: 56), 3.5 (2001-2003: 138), 3 (2005-2007: 297), 3 (2009: 499) and 3 (2012: 1020) years. The same years to double from 1993 numbers for the related disciplines separately would be for 6, 4 and 6 years for HCI and 9 and 10 years for EEG. In other words, the number of published studies using EEG in HCI grows about twice as fast as HCI and three times faster than EEG in general. Whereas use of EEG was extremely rare for HCI in 1993, (at 0.6%), it is now merely uncommon (at 5.2%).

3 Naïve Neurorealism

The degree that neuroscience has captured the imagination of the popular press and interest of companies and academic institutions alike is thus understandable, but should be treated with a healthy dose of scientific scepticism. We coin the term *naïve neurorealism* to describe the idea that "the brain cannot lie" and that somehow, a subjective measure immediately becomes objective, accessible and trustworthy because the brain is "directly" involved. The navity might stem from the Cartesian assumption that because the brain "causes thought", the measurement of the brain necessarily brings one closer to "the truth".

In HCI, this can for example lead to the idea that one simply plugs in a brain signal and thereby improve an existing user-interface. To illustrate, let us imagine the following, hypothethical scenario in which EEG in HCI could, but should not, be applied:

> *The BrainGuitar.* The guitar remains an extremely popular musical interface. However, its bi-manual multi-touch design is characterised by a steep learning slope that can be an obstacle to many a beginner. We imagine a future in which, rather than relying on our hands, we can directly control the guitar through the use of our brain. The BrainGuitar relies on spectral analysis of EEG signals to determine whether ongoing music is enjoyable or not. If the music produced by the BrainGuitar is not liked, we realign the style to suit the musician's objective taste. A user study was carried out in which 20 practitioners, none of whom had prior experience playing guitar, played 5-10 minutes either with BrainGuitar or normal, classical guitar. We prove with surveys that satisfaction and usability are significantly better with Brain- than normal guitars. Self-reports indicated that BrainGuitar users were surprised to discover their subjective and objective musical taste did not always correspond, showcasing the exploration value of the interface. Finally, we discuss the neurofuture in which everyone can enjoy playing guitar.

In our hypothetical scenario, naïve neurorealism is demonstrated by the invalid assumption that because we use a neural source (possibly), the BrainGuitar is better able to determine whether the user likes a song or not. Combined with the intrinsic appeal of the human mind, this can easily lead to big claims of "mind reading", "mind controling" or "thought identification". However, already in the first BCI-related publication, the authors clarify [6]: "*[...], there is no more 'mind reading' in the procedures we describe than there is when a person is handed a pencil and asked to record impressions.*" Mind reading is made vastly more difficult due to concerns regarding reliability and validity.

4 Reliability and Validity Concerns

Reliability concerns the degree to which measurements are consistent. For example, a measuring tape, properly handled, can reliably indicate a user's height in feet, meters, inches and/or centimetres. It is unlikely to give a very different metric if the measuring is taken on a different day of the week, or by a different person. The validity of a measurement tool concerns the degree to which it measures what it is supposed to measure. A measuring tape can be used to measure one's height, but not weight.

4.1 Source Localization

EEG, by contrast, is a very indirect measure: since the electrodes are placed on the skin, rather than in the brain directly, it can only measure the electric potential of the scalp. Electrical potentials with neuronal sources, therefore, are picked up only after having passed other cortical areas, as well as the cerebrospinal fluid, skull and skin, resulting in the well known problem of spatial blurring, and poor spatial resolution of EEG [14]. Indeed, given the already weak signals of individual sources, the only reason we can measure EEG in the first place is because large groups of neurons fire synchronously in the same direction. Accordingly, the topography of an EEG potential gives a poor approximation of its neuronal source, which has given rise to various proposed solutions for predicting scalp activity as a function of a known source (the forward problem, c.f. [15]) and localising sources as a function of measured scalp activity (the inverse problem, [16]).

4.2 Signal to Noise Ratio

However, as debate regarding localisation problems continues, it is still true that human brain-related EEG has been reliably measured for more than 85 years now [1], and related to specific cognitive functions for over half a century [17]. One of the reasons of this gap again has to do with reliability: although Berger's [1] alpha activity can be easily discerned by the eyes (see Fig. 2, C), being in the range of ca. 30 μV, EEG related to specific sensory or response related events are relatively much smaller. For instance, in the top-right panel, Fig. 2 shows what raw EEG looks like and how much it is affected by visual events (here: focally presented pictures of people).

It is impossible to discern any EEG related to the stimulus due to the amount of background noise relative to the signal (i.e. the low signal to noise ratio). However, as the background noise is presumably unrelated to the event, it can be steadily reduced by repeating the exact same conditions over and over again, and averaging across measurements. Sutton et al. [4] thus used between 30 and 360 repetitions to create the average event related potentials used to discover the P300 (previously mentioned in Sec. 2). Clearly, the number of data points involved in computation over subjects, conditions, channels, timepoints and repetitions was considerable and advancements in EEG benefited accordingly from the availability of computers in university campuses. Much has changed since 1965, but the problem of low SNR in EEG remains. Indeed, both the number of suggested repetitions and emphasis on strong experimental control remains similar in modern EEG [18]. Of particular concern, in this regard, are artifacts in EEG.

4.3 Artifacts

One of the reasons that we require so many repetitions is because EEG is commonly contaminated with artifacts. In particular, EEG is extremely sensitive to eye-blinks and movements, as is portrayed in Fig. 2. In the top left panel, EEG is shown during episodes of eye movements (A) and blinks (B), resulting in activity levels of $>100\,\mu V$. The traditional way to deal with such artifacts is commonly referred to as "artifact rejection by visual inspection", which means that an expert looks at visualisations of the entirety of the data during or after collection, and selects all data that is suspected of being contaminated with artifacts related to eye-movements and blinks, as well as head movements, muscle activity, and so on. The top left panel, in this regard, contains problematic data while the top right seems more normal. These days, EEG studies tend to rely on automatic or semi-automatic classification to distinguish contaminated from clean data, but as the latter tend to account for more variance than the former, artifact rejection is still commonly employed. As this leads to the removal of significant amounts of data, more repetitions are required to sustain the reliability.

Another family of methods commonly employed to enhance SNR are artifact-correction methods. Rather than removing time-points from the data if artifacts are suspected, these methods aim to subtract their contribution from the signal. The classic method for doing this is via linear regression removing the correlation with the electro-oculogram (EOG, [19]), which is normally collected with electrodes placed at sites near the eyes. However, there are drawbacks to this method – it will, for example, also remove the EEG that is collected with EOG electrodes. For this reason, methods that decompose the EEG into components that can be related to artifacts or uncontaminated EEG are becoming more popular [20]. An example is provided in the central top panel of Fig. 2, showing the top left panel as it appeared after removing EOG related components using independent component analysis [21]. Again, the degree of activity after artifact correction is much lower (here ca. 2 times), although it is still clearly higher than during the "clean" interval. For this reason, we have added specific guidelines here regarding artifacts.

4.4 Measuring Mental States or Systematic Artifacts

A common misconception is that artifacts only reduce reliability. However, artifacts may also be disastrous to the validity of a study. For example, the eye movements shown in Fig. 2 strongly affect activity, but not necessarily consistently across spectral frequencies. That is to say, eye movements cause extreme power in the lowest ($<8\,Hz$) frequencies but do not affect alpha ($8\,Hz$–$12\,Hz$)

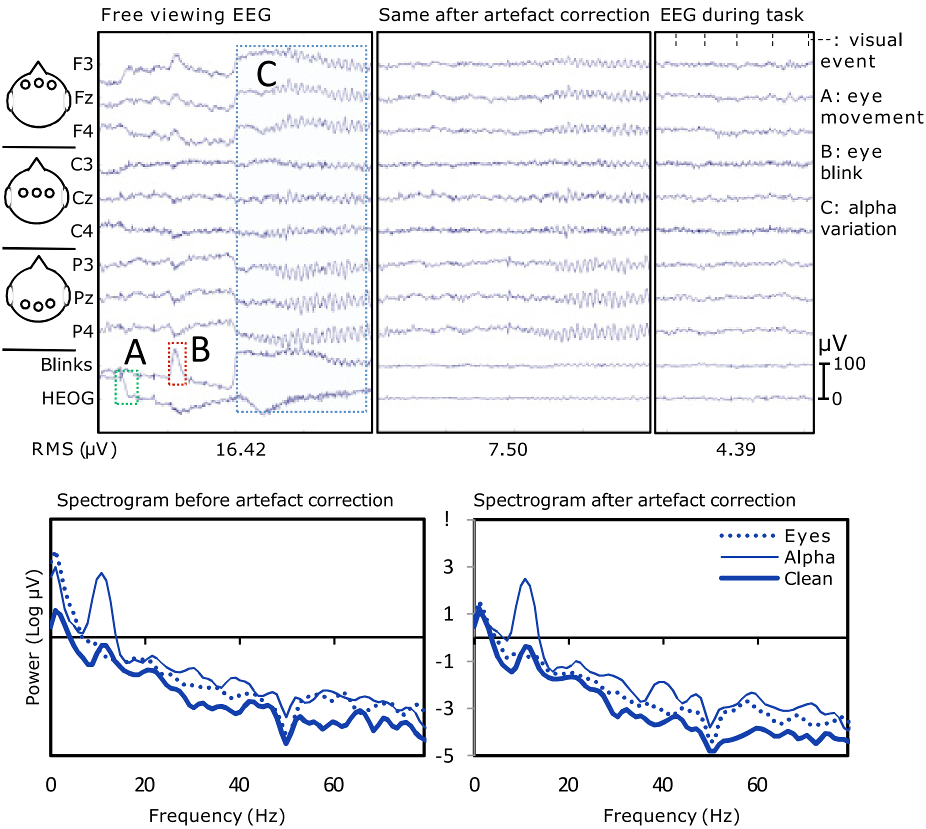


Fig. 2. Effects of artifacts on EEG measurements in time (upper panels) and frequency (lower panels) domains. A, B, and C indicate three common types of contaminating sources related to eye movements, eye blinks and alpha activity. The central top panel shows how artifact correction affects the strongly contaminated top left panel. The top right panel displays typical EEG activity during a task. The lower left panel show the spectral power of the first ("eyes": artifacts A and B), and second ("alpha": C) half of the top left panel, as well as the top-right panel ("clean" data). Finally, The lower right panel shows the spectral power after artifact correction. Data available from www.hiit.fi/manuel.eugster/aomm2015/.

and beta (13 Hz–29 Hz) frequencies as much (a common finding, cf. [22]). The higher frequencies (30–200 Hz), meanwhile, are of considerable interest for people interested in consciousness [23], meditation [24] and neurofeedback [25] but as illustrated in Fig. 2, and investigated closer elsewhere [26], [27], it is also possible that eye and muscle movements cause activity within these frequency bands. Consequently, when neuroscientists study how a cognitive function causes differential spectral activity, they aim to control extraneous effects as much as possible to avoid confounds that is unrelated effects that explain the observed findings. As a result, there is little consensus regarding the relationship between specific frequency bands and cognitive functions.

4.5 Data Analysis

Typically, the use of EEG in HCI starts with a simple application idea (here, the BrainGuitar). The first obstacle, as discussed so far, is to map such an idea to a sound neurophysiological paradigm. The second obstacle then, after recording the data, is to rigorously perform the data analysis. Unfortunately, decoding brain states is a difficult data analytic endeavour: major issues are often very specific experimental designs, the unfavourable signal to noise ratio, the vast dimensionality of the data, and the high trial-to-trial variability [28]. Fortunately, recent literature provides comprehensive frameworks for rigorous statistical analysis and predictive modelling: see [29] for a tutorial on single-trial analysis, and [30] for a general approach on evaluating prediction algorithms.

From an application point of view, the final goal of the data analysis is to develop a predictive model which discriminates the different brain states with the highest accuracy. In our hypothetical scenario, this is for example a classifier predicting one of the two classes "I like the song" versus "I don't like the song". Estimating the predictive power and the generalisation error of such a classifier on the measured data is a easy source of mistakes (see [28] for a list of typical pitfalls). Here, we want to underscore one. Neurophysiological paradigms often rely on an imbalance of the brain states under investigation. Consequently, the prediction problem is imbalanced and the used cross-validation scheme as well as the performance measures have to take this into account (see [28], Section 5.6). In the BrainGuitar example, no user had prior experience playing guitar. Therefore, most of the observations are probably "dislikes" and only a few "likes" will be available. A default "I don't like this song" classifier will have a high accuracy but no validity.

4.6 The Seductive Allure of Neurorealism

Neuroscience has brought many advances to our understanding of the brain and mind, to the point that the expectations of its capabilities are clearly exaggerated. Thus, people are known to find even bad explanations more convincing if they come wrapped in neuroscience talk [31]. One of the problems inherent in neuroscience is that the same cognitive (or artefactual) function can map onto various spectral frequencies (see Sec. 4.4) or brain areas, and conversely, that different functions may affect the same frequencies or brain areas. As a result, it is often possible to predict activity in various brain areas (IF mental function X, THEN brain activity Y), but the reverse inference (IF activity Y then mental function X) is fallacious [32], and not necessarily true. Similarly, localising cognitive functions to specific areas can be challenging: for example, hemispheric activity has been linked to both emotional valence (positive / negative emotions) and motivation. Simply observing greater left than right activity would not be informative enough to conclude whether someone is highly motivated, or in a positive mood [3]. Moreover, mappings between brain areas and functions within the same person might not be stable over time, since specialisations in brain areas can adapt to changes in its environment [33]. While it is debatable (c.f. [34]) that the reverse inference can sometimes lead to valuable information, this is not necessarily true.

In the BrainGuitar, one can argue that a correlate for "liking" a song might be found in frontal asymmetry (but see [35]). The fallacy of reverse inference and the naïve of the researchers is demonstrated by the surprise of the user: While this should suggest the measurement of liking was invalid, instead they feel the BrainGuitar revealed something beyond the knowledge of the users. Are they sure, we should ask, if they do not pick up correlates that in themselves may be caused by the brain, but are not equivalent to brain signals? For example, eye movements and muscle activity are, of course, caused by brain activity, and contaminate EEG activity, but are not brain activity themselves.

In other words, reverse inferences should be treated with great caution, and overly positive statements such as "enjoyment was determined using EEG" should be avoided. We urge the field (see also [36]) to tread cautiously, particularly when making strong claims in academic work and when talking to the popular media.

5 Other Psychophysiological Measurements

In this article, we tend to use EEG and psychophysiology interchangeably. This is, on the one hand, because concerns over reliability and validity are not as striking for other physiological measurements such as electrodermal activity (EDA, or galvanic skin response), heart rate (electrocardiogram), respiration rates, and so on. In some cases, combining signals from these sources into a single predictor can provide a better assessment of the users' state than EEG alone (see [37] for a review).

The source of these measures is rather well localised (the hand, the heart), and the number of associated psychological constructs is limited (usually arousal). However, even for these measures, the relationship is not as simple as it first appears: emotionally exciting stimuli tend to show increased EDA but slowing of heart rate, arousal (as an emotional state) tends to have increased EDA and increased heart rate [38].

Functional near-infrared spectroscopy is another brain imaging technique which measures the blood oxygenation level dependent (BOLD) response, similar to fMRI[39]. This is done by making use of the different light absorption properties of oxygenated and de-oxygenated haemoglobin. It has been successfully employed to image the human brain (see for a review [40]). Although the field remains relatively young, it is generally held that the spatial localisation of superficial frontal sources in particular is good. Given that it is also relatively cheap, at least compared to MRI, and easier to prepare than EEG, it is possible that the technique will be very popular with interdisciplinary disciplines such as HCI. However, it should be stressed that knowing that, for example, the right Brodmann Area 10 is active, does not relieve one from naïve neurorealism: the area has been mapped onto functions of recollection of episodes [41] and odours [42], non-speech sounds [43], risk and reward [44]. Is the BrainGuitar familiar, sound-making, challenging or does it simply smell familiar?

6 Consumer Devices

Consumer grade EEG devices are relatively very cheap, usually wireless and are often easier to set up. These two qualities have created a popular sense that the future is mind controlled in mainstream media [45–47]. This future, however, for now largely remains science fiction. In particular, consumer grade devices (e.g. Epoc Emotiv, Neurosky) focus on low cost materials and ease of setting up, which will adversely affect SNR and validity. Science or clinical grade electrodes use highly conducive materials such as silver-chloride (AgCl) and gold in order to capture as much signal from the scalp as possible materials of which the cost can be prohibitive for single consumers. Furthermore, anything (e.g. air, skin flakes) between the electrodes and scalp will adversely affect SNR, for which reason researchers often prepare the skin (by scraping, use of cleaning materials) and use materials like conducive gel to fill in the space between electrode and scalp. Of course, such procedures are not particularly comfortable and generally require assistance from an extra person.

Finally, while psychophysiology experts prefer the use of many electrodes placed at standardised, equidistant locations on the scalp both in order to increase SNR and to enhance external validity with other research groups, this naturally increases cost and effort. Accordingly, many consumer grade devices use few (<16) electrodes. The Emotiv EPOC is a common, noteworthy exception at 14 channels (and two references), although the extra electrodes have a focus on eye and facial muscle activity rather than EEG (but see [48] for a way around). Consumer grade devices provide ready made quantified emotional and cognitive state analysis, but the validity of these classifications cannot easily be assessed as they tend to rest on trade-secrets and subjective reports. This, in fact, leads to exactly the circular problem the present paper is aiming to address: the SNR is poor and it is unclear what is measured. In sum, consumer grade devices may well provide good fun for consumers, but for science these benefits are likely offset by the extra costs and efforts involved if a submission is rejected.

7 Conclusion

The use of psychophysiology in HCI has been remarkable. We have seen many instances in which the fusion of neuroscience and HCI can create new insights and applications. However, the popularity and naïve neurorealism can lead to an overly optimistic idea of making psychophysiology a simple plugin of the human-computer interaction. More importantly, we discussed issues of reliability and validity that make claims regarding direct mind-control tenuous.

To help this exciting new field, we would like to conclude with a few questions. From our experience, it is useful to keep these questions in mind while developing and presenting EEG-in-HCI applications. We hope they may prove beneficial to other researchers and reviewers.

1. How much does the quality of the apparatus involved reflect the study's aims?
2. How much does the conclusion reflect the known limitations of the measurements?
3. Which method was used to correct and/or reject data? How many repetitions were used in the analysis (BCI: training)?
4. Which electrode sites were used as channels and which reference(s) were employed?
5. Do the psychophysiological markers correspond to what the authors aim/claim to measure or could differences have been caused by correlated variables? Is the paradigm sound?
6. Does the control condition provide a valid comparison?
7. Has the work been communicated to the press with an unbiased, factual report, and were all communications with the press reviewed by the involved researchers?

References

1. Berger, H.: Über das elektroencephalogramm des menschen. Psychiatry **87**, 527–570 (1929)
2. Karbowski, K.: Hans berger (1873–1941). J. Neurol. **249**(8), 1130–1131 (2002)
3. Miller, G.A., Crocker, L.D., Spielberg, J.M., Infantolino, Z.P., Heller, W.: Issues in localization of brain function: the case of lateralized frontal cortex in cognition, emotion, and psychopathology. Frontiers in Integrative Neuroscience **7** (2013)
4. Sutton, S., Braren, M., Zubin, J., John, E.: Evoked-potential correlates of stimulus uncertainty. Science **150**(3700), 1187–1188 (1965)
5. Linden, D.E.J.: The p300: Where in the brain is it produced and what does it tell us? The Neuroscientist **11**, 563–576 (2005)

6. Farwell, L., Donchin, E.: 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)
7. Campbell, A., Choudhury, T., Hu, S., Lu, H., Mukerjee, M.K., Rabbi, M., Raizada, R.D.: Neurophone: brain-mobile phone interface using a wireless EEG headset. In: MobiHeld 2010 Proceedings of the Second ACM SIGCOMM Workshop on Networking, Systems, and Applications on Mobile Handhelds, pp. 3–8 (2010)
8. Picard, R.W., Vyzas, E., Healey, J.: Toward machine emotional intelligence: Analysis of affective physiological state. IEEE Transactions on Pattern Analysis and Machine Intelligence **23**(10), 1175–1191 (2001)
9. Fairclough, S.H.: Fundamentals of physiological computing. Interacting with Computers **21**(1), 133–145 (2009)
10. Fairclough, S.H., Ewing, K.C., Roberts, J.: Measuring task engagement as an input to physiological computing. In: 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009, pp. 1–9 (2009)
11. Angelakis, E., Stathopoulou, S., Frymiare, J.L., Green, D.L., Lubar, J.F., Kounios, J.: Eeg neurofeedback: a brief overview and an example of peak alpha frequency training for cognitive enhancement in the elderly. The Clinical Neuropsychologist **21**(1), 110–129 (2007)
12. Rabipour, S., Raz, A.: Training the brain: fact and fad in cognitive and behavioral remediation. Brain Cogn. **79**(2), 159–179 (2012)
13. Larsen, P.O., Von Ins, M.: The rate of growth in scientific publication and the decline in coverage provided by science citation index. Scientometrics **84**(3), 575–603 (2010)
14. Junghöfer, M., Elbert, T., Leiderer, P., Berg, P., Rockstroh, B.: Mapping eeg-potentials on the surface of the brain: a strategy for uncovering cortical sources. Brain Topogr. **9**(3), 203–217 (1997)
15. Acar, Z.A., Makeig, S.: Neuroelectromagnetic forward head modeling toolbox. J. Neurosci. Methods **190**(2), 258–270 (2010)
16. Pascual-Marqui, R.D.: Review of methods for solving the eeg inverse problem. International Journal of Bioelectromagnetism **1**(1), 75–86 (1999)
17. Walter, W.G., Cooper, R., Aldridge, V., McCallum, W., Winter, A.: Contingent negative variation: an electric sign of sensorimotor association and expectancy in the human brain. Nature **8**(203), 380–384 (1964)
18. Luck, S.: An Introduction to the Event-related Potential Technique. MIT Press, Cambridge (2005)
19. Gratton, G., Coles, M., Donchin, E.: A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology **55**(4), 468–484 (1983)
20. Jung, T.P., Humphries, C., Lee, T.W., Makeig, S., McKeown, M.J., Iragui, V., Sejnowski, T.J.: Removing electroencephalographic artifacts: comparison between ICA and PCA. In: Proceedings of the 1998 IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing VIII, pp. 63–72 (1998)
21. Delorme, A., Makeig, S.: Eeglab: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods **134**(1), 9–21 (2004)
22. Whitton, J.L., Lue, F., Moldofsky, H.: A spectral method for removing eye movement artifacts from the EEG. Electroencephalography and Clinical Neurophysiology **44**(6), 735–741 (1978)
23. Singer, W., Gray, C.M.: Visual feature integration and the temporal correlation hypothesis. Annu. Rev. Neurosci. **18**(1), 555–586 (1995)

24. Lutz, A., Greischar, L.L., Rawlings, N.B., Ricard, M., Davidson, R.J.: Long-term meditators self-induce high-amplitude gamma synchrony during mental practice. Proc. Natl. Acad. Sci. USA **101**(46), 16369–16373 (2004)
25. Keizer, A.W., Verment, R.S., Hommel, B.: Enhancing cognitive control through neurofeedback: A role of gamma-band activity in managing episodic retrieval. Neuroimage **49**(4), 3404–3413 (2010)
26. Muthukumaraswamy, S.D.: High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations. Frontiers in Human Neuroscience **7** (2013)
27. Whitham, E.M., Pope, K.J., Fitzgibbon, S.P., Lewis, T., Clark, C.R., Loveless, S., Broberg, M., Wallace, A., DeLosAngeles, D., Lillie, P.: Scalp electrical recording during paralysis: quantitative evidence that EEG frequencies above 20hz are contaminated by emg. Clin. Neurophysiol. **118**(8), 1877–1888 (2007)
28. Lemm, S., Blankertz, B., Dickhaus, T., Müller, K.R.: Introduction to machine learning for brain imaging. Neuroimage **56**(2), 387–399 (2011). http://www.sciencedirect.com/science/article/pii/S1053811910014163
29. Blankertz, B., Lemm, S., Treder, M.S., Haufe, S., Müller, K.R.: Single-rial analysis and classification of ERP components - A tutorial. NeuroImage **56**(2), 814–825 (2011). http://www.sciencedirect.com/science/article/pii/S1053811910009067
30. Eugster, M.J.A., Hothorn, T., Leisch, F.: Domain-based benchmark experiments: Exploratory and inferential analysis. Austrian Journal of Statistics **41**(1), 5–26 (2012). http://www.stat.tugraz.at/AJS/ausg121/121Leisch.pdf
31. Weisberg, D.S., Keil, F.C., Goodstein, J., Rawson, E., Gray, J.R.: The seductive allure of neuroscience explanations. J. Cogn. Neurosci. **20**(3), 470–477 (2008)
32. Poldrack, R.A.: Can cognitive processes be inferred from neuroimaging data? Trends in Cognitive Sciences **10**(2), 59–63 (2006)
33. Duncan, J.: An adaptive coding model of neural function in prefrontal cortex. Nat. Rev. Neurosci. **2**(11), 820–829 (2001)
34. Poldrack, R.A.: The role of fMRI in cognitive neuroscience: where do we stand? Curr. Opin. Neurobiol. **18**(2), 223–227 (2008)
35. Chai, Jing, Ge, Yan, Liu, Yanfang, Li, Wen, Zhou, Lei, Yao, Lin, Sun, Xianghong: Application of frontal EEG asymmetry to user experience research. In: Harris, Don (ed.) EPCE 2014. LNCS, vol. 8532, pp. 234–243. Springer, Heidelberg (2014)
36. Brouwer, A.M., Zander, T.O., Van Erp, J.B.F., Korteling, H., Bronkhorst, A.W.: Using neurophysiological signals that reflect cognitive or affective state: Six recommendations to avoid common pitfalls. Frontiers in Neuroscience 9(136) (2015). http://www.frontiersin.org/neuroprosthetics/10.3389/fnins.2015.00136/abstract
37. Novak, D., Mihelj, M., Munih, M.: A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. Interact. Comput. **24**(3), 154–172 (2012)
38. Cacioppo, J.T., Tassinary, L.G., Berntson, G.: Handbook of Psychophysiology. Cambridge University Press (2007)
39. Cui, X., Bray, S., Bryant, D.M., Glover, G.H., Reiss, A.L.: A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. Neuroimage **54**(4), 2808–2821 (2011)
40. Ferrari, M., Quaresima, V.: A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. Neuroimage **63**(2), 921–935 (2012)
41. Paulesu, E., Sambugaro, E., Torti, T., Danelli, L., Ferri, F., Scialfa, G., Sberna, M., Ruggiero, G., Bottini, G., Sassaroli, S.: Neural correlates of worry in generalized

anxiety disorder and in normal controls: a functional MRI study. Psychological Medicine **40**(01), 117–124 (2010)
42. Royet, J.P., Koenig, O., Gregoire, M.C., Cinotti, L., Lavenne, F., Le Bars, D., Costes, N., Vigouroux, M., Farget, V., Sicard, G.: Functional anatomy of perceptual and semantic processing for odors. J. Cogn. Neurosci. **11**(1), 94–109 (1999)
43. Wong, D., Pisoni, D.B., Learn, J., Gandour, J.T., Miyamoto, R.T., Hutchins, G.D.: Pet imaging of differential cortical activation by monaural speech and nonspeech stimuli. Hearing Research **166**(1), 9–23 (2002)
44. Rogers, R.D., Owen, A.M., Middleton, H.C., Williams, E.J., Pickard, J.D., Sahakian, B.J., Robbins, T.W.: Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. The Journal of Neuroscience **19**(20), 9029–9038 (1999)
45. Isaacson, B.: Mind control: How EEG devices will read your brain waves and change your world (2012). http://www.huffingtonpost.com/2012/11/20/mind-control-how-eeg-devices-read-brainwaves_n_2001431.html (retrieved on May 22, 2015)
46. Le, T.: A headset that reads your brainwaves (2010). http://www.ted.com/talks/tan_le_a_headset_that_reads_your_brainwaves?language=en (retrieved on May 22, 2015)
47. Li, S.: Mind reading is on the market (2010). http://articles.latimes.com/2010/aug/08/business/la-fi-mind-reader-20100808 (retrieved on May 25, 2015)
48. Debener, S., Minow, F., Emkes, R., Gandras, K., de Vos, M.: How about taking a low-cost, small, and wireless EEG for a walk? Psychophysiology **49**(11), 1617–1621 (2012)

Evaluation of Suitable Frequency Differences in SSVEP-Based BCIs

Piotr Stawicki, Felix Gembler, and Ivan Volosyak(✉)

Faculty of Technology and Bionics, Rhine-Waal University of Applied Sciences, Kleve, Germany
ivan.volosyak@hochschule-rhein-waal.de
http://www.hochschule-rhein-waal.de

Abstract. Frame-based frequency approximation methods are a popular approach to realize visual stimuli that can be used to elicit steady-state visual evoked potentials (SSVEPs) at various frequencies on computer screens and allows the development of multi-target Brain-Computer Interface (BCI)-Systems. In this paper we investigate appropriate selection of visual stimuli for multi-target BCIs using a frequency approximation method. Twelve sets of frequencies from different bands and with different resolutions have been compared among each other during an on-line BCI-task with six healthy subjects. Our results confirm that equidistant frequency sets are not optimal, as the results from the sets with lower frequency ranges (<12 Hz) surpass those of the mid-range sets, even if a higher resolution is used. Interestingly, the study shows that SSVEPs elicited by stimuli from lower bands with a very high frequency resolution of 0.05 Hz could still be classified with adequate accuracy (around 90%). The results confirm that careful stimuli choice has high impact on SSVEP based BCI performance.

Keywords: BCI (Brain-Computer Interface) · SSVEP (Steady-State Visual Evoked Potential) · LCD (liquid crystal display) · Frequency

1 Introduction

Brain-Computer Interfaces (BCIs) translate brain signals, usually acquired non-invasively using electroencephalogram (EEG), in computer commands without using the brain's normal output pathways of peripheral nerves and muscles [15]. Such communication technologies have the potential to help people with physical impairments, if they provided special interfaces that worked independently of the person's limitations. One of the BCI paradigms used for realization of multi-target interfaces is the Steady-state visual evoked potential (SSVEP)-based BCI which measures the brain responses to a visual stimulation at specific constant frequencies [15]. A popular source for the visual stimuli are computer screens. Since implementation mainly relies on software development, use of computer screens offers flexibility for combining BCI stimulation with the controlled application and makes it possible for the stimulation interface to easily be

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 159–165, 2015.
DOI: 10.1007/978-3-319-24917-9_17

fine-tuned during BCI development [16]. For SSVEP-based spelling applications, the arrangement and the number of classes (simultaneously displayed stimulation frequencies) influence speed and usability. The number of stable frequencies that can be rendered on a monitor are always limited by the refresh rate since the number of frames in a stimulation cycle needs to be a constant [3,9,13]. That is why SSVEP based spelling systems using those stable frequencies use usually four to seven different targets [1,5,8]. Two or three successive commands are needed to spell a single target character in those applications, due to the fact that the number of target characters outnumbers the number of visual stimuli. However, the so-called frequency approximation method, as proposed by Wang et al. [14], allows the realization of visual flickers with a high frequency resolution (e.g., 0.25Hz) and the implementation of high-speed multiple target BCI on a computer screen [2,3,9,14]. The choice of the stimulation frequencies has impact on the BCI-perfomance. In many research articles using multi-target SSVEP-based BCIs, equidistant stimulation frequencies are used [3,7,14]. Effects of stimuli choice has been discussed extensively throughout the BCI-literature [12]. The best SSVEP responses are obtained using stimulation frequencies between 5 and 20 Hz and 15 Hz is the stimulation frequency at which the SSVEP response is maximum [10]. Also one has to bear in mind, that mutual influences between stimulating frequencies should be avoided. Each pair or triple of simultaneously flickering stimuli should not harm the restriction rules $f_i \neq [f_j + f_k]/2, f_i \neq 2f_j - f_k, f_i \neq 2f_k - f_j$. According to Gao et al. two flickering targets with a frequency difference as low as 0.2 Hz can be successfully distinguished in the SSVEP response [4]. Lately the maximal SSVEP frequency resolution for reliable detection has been updated to 0.1 Hz [3,7]. In this paper we investigated the question of optimal frequency selection regarding multi-target stimulation, as an evenly divided frequency band might not be optimal for the implementation of such multi-target BCIs. We compared the BCI performance among frequency sets from different domains and with different resolutions.

2 Methods and Materials

2.1 Subjects

This study was carried out in accordance with the guidelines of the Rhine-Waal University of Applied Sciences. All subjects gave written informed consent in accordance with the Declaration of Helsinki. Six healthy volunteer subjects with a mean (SD) age of 25.33 (4.6) years participated in the study. One subject was female. All subjects were students or employees of the Rhine-Waal University of Applied Sciences and had little or no previous experience with BCI-systems.

The EEG recording took place in a laboratory room (approx. 36 square meters) with low background noise and luminance. Spectacles were worn when appropriate.

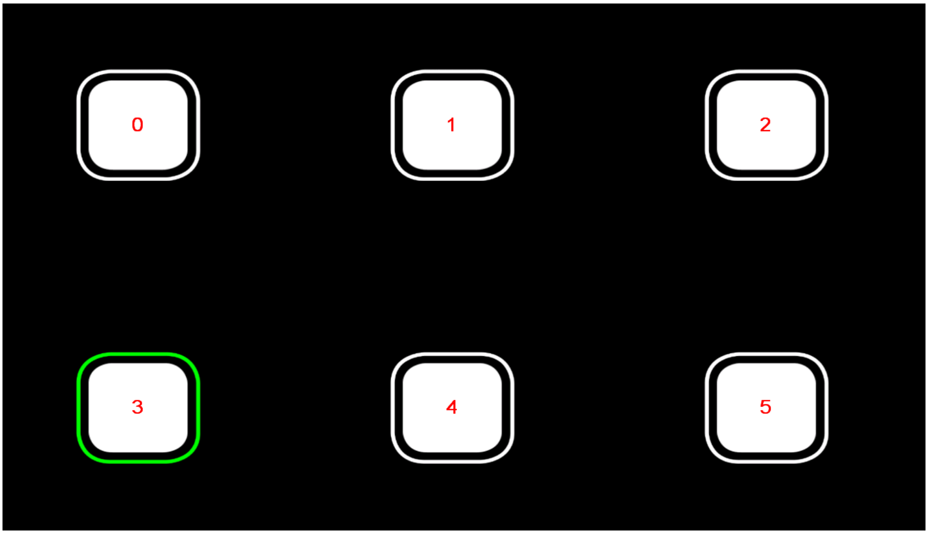


Fig. 1. GUI of the BCI-system during the online experiment. The subject had to concentrate on the box containing the number three.

2.2 Signal Acquisition

Subjects were seated in front of a LCD screen (BenQ XL2420T, resolution: 1920×1080 pixels, vertical refresh rate: 120 Hz) at a distance of about 60 cm. The used computer system operated on Microsoft Windows 7 Enterprise running on an Intel processor (Intel Core i7, 3.40 GHz). Standard Ag/AgCl electrodes were used to acquire the signals from the surface of the scalp.

The ground electrode was placed over AF_Z, the reference electrode over C_Z, and the eight signal electrodes were placed at predefined locations on the EEG-cap marked with $P_Z, PO_3, PO_4, O_1, O_2, O_Z, O_9$, and O_{10} in accordance with the international system of EEG electrode placement. Standard abrasive electrolytic electrode gel was applied between the electrodes and the scalp to bring impedances below 5 $k\Omega$. An EEG amplifier, g.USBamp (Guger Technologies, Graz, Austria), was utilized. The sampling frequency was set to 128 Hz. During the EEG signal acquisition, an analogue band pass filter (between 2 and 30 Hz) and a notch filter (around 50 Hz) were applied directly in the amplifier.

2.3 Graphical User Interface

For this study we designed a 6-target BCI-system (see Fig. 1). The graphical user interface (GUI) is a 2×3 stimulus matrix containing the numbers 0 to 5. Each stimulus was presented within a 144×128 pixels box and the distance between two adjacent boxes was roughly 300 pixels. One of the boxes was surrounded by a green frame. Each box flickered with a specific frequency. The user was told to concentrate on the highlighted box. After correct classification, another number would be highlighted and the user had to shift his gaze and focus on this box. The order in which the numbers where highlighted was randomized. When every number was selected once, the task was completed and the flickering stopped. If no error was made, the task was completed after six selections. After a short break the task was repeated with the next set of frequencies. In total twelve different frequency sets were tested. In order to reduce the information load of the visual channel, every command classification was followed by an audio feedback. For SSVEP signal classification a minimum energy combination method was used. The SSVEP classification was performed online every 13 samples (ca. 100ms) on the basis of the adaptive time segment length of the acquired EEG data. More details about the used SSVEP detection method can be found in [11]. In order to make the system more robust we included larger time windows (8 and 16 seconds) and increased the number of time segment lengths, as discussed in [6]. In this study the minimal classification time window was set to 2 seconds. In order to implement multiple targets on the LCD-screen, a frame-based stimulus approximation method was used. If a varying number of frames in each cycle is used any frequency up to half of the refresh rate can be approximated and a much higher frequency resolution can be achieved. The stimulus signal at frequency f can be generated by the following equation:

$$\mathrm{stim}(f, i) = \mathrm{square}[2\pi f(i/\mathrm{RefreshRate})],$$

where $\mathrm{square}(2\pi ft)$ generates a square wave with frequency f and i is the frame index. For example, the one-second stimulus sequence for the frequency 17 Hz on a 120 Hz refresh rate monitor is $\underline{4\ 4}$ 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 $\underline{4\ 4}$ 3 4 3 4 3 4 3 4 3 4 3 4 3 4 3 (see Fig. 2).

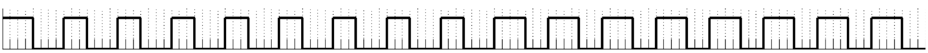


Fig. 2. One second stimulus sequence for the approximated frequency 17 Hz. The black/white reversing interval for the approximated frequency 17 Hz includes 17 cycles of varying length (three or four frames).

2.4 Experimental Setup

After signing the consent form, each subject was prepared for the EEG recording. Each subject had to complete the task twelve times with different frequency sets (one recording for each set). Four groups of frequency sets were used, each containing three sets of six equidistant frequencies (see Table 1). Each selection phase ended automatically when all numbers were selected correctly. After a brief pause (approximately 20 seconds) the program was restarted manually and the procedure was repeated with the next set. The entire procedure took about 40 min on average per subject.

Table 1. Overview of the different frequency sets.

Set	Frequencies [Hz]						Set	Frequencies [Hz]						ΔHz
1	6.10	6.20	6.30	6.40	6.50	6.60	7	15.10	15.20	15.30	15.40	15.50	15.60	0.100
2	6.10	6.18	6.25	6.33	6.40	6.48	8	15.10	15.18	15.25	15.33	15.40	15.48	0.075
3	6.10	6.15	6.20	6.25	6.30	6.35	9	15.10	15.15	15.20	15.25	15.30	15.35	0.050
4	10.10	10.20	10.30	10.40	10.50	10.60	10	20.10	20.20	20.30	20.40	20.50	20.60	0.100
5	10.10	10.18	10.25	10.33	10.40	10.48	11	20.10	20.18	20.25	20.33	20.40	20.48	0.075
6	10.10	10.15	10.20	10.25	10.30	10.35	12	20.10	20.15	20.20	20.25	20.30	20.35	0.050

3 Results

BCI performance for each subject was evaluated by calculating the commonly used ITR in bits/min, employing the formula as discussed e.g. in [15]. In the GUI presented here, the overall number of possible choices was six. The accuracy was calculated based on the number of correct command classifications divided by the total number of classified commands. The overall BCI performance is given in Table 2.

Table 2. Results for the BCI-task. ITR values in bits/min for each frequency set and subjects are displayed. Mean values are given at the bottom of the table.

	Set Nr.	1	2	3	4	5	6	7	8	9	10	11	12
ITR [bits/min]	Subject 1	35.4	10.3	10.8	20.9	18.8	14.2	6.7	9.0	23.3	0.4	0.0	0.5
	Subject 2	25.9	11.4	24.6	41.6	7.8	27.8	11.7	7.3	25.4	20.7	3.2	9.4
	Subject 3	32.4	30.2	11.8	36.8	31.0	24.6	33.8	29.8	21.8	13.2	12.0	14.5
	Subject 4	30.1	20.8	33.3	11.6	15.1	39.2	32.0	11.6	8.2	0.0	0.3	1.8
	Subject 5	16.7	18.7	26.6	27.9	12.8	21.1	16.1	15.2	24.6	27.5	4.3	16.7
	Subject 6	14.6	5.3	4.6	27.4	14.9	12.4	23.2	9.6	11.1	0.5	0.2	2.5
	Mean	**25.8**	**16.1**	**18.6**	**27.7**	**16.7**	**23.2**	**20.6**	**13.8**	**19.1**	**10.4**	**3.3**	**7.5**
	SD	**8.5**	**8.9**	**11.1**	**10.8**	**7.9**	**9.8**	**11.0**	**8.3**	**7.4**	**11.9**	**4.6**	**7.1**
Accuracy [%]	Subject 1	100	75	75	86	86	86	60	67	100	25	0	29
	Subject 2	86	67	100	100	60	100	67	60	100	86	46	75
	Subject 3	100	100	86	100	100	100	100	100	100	86	86	100
	Subject 4	86	75	100	60	67	100	100	67	67	17	26	43
	Subject 5	75	86	100	100	75	100	75	86	100	100	55	100
	Subject 6	75	55	50	86	75	75	86	67	75	29	23	46
	Mean	**87**	**76**	**85**	**89**	**77**	**93**	**81**	**74**	**90**	**57**	**53**	**65**
	SD	**11**	**16**	**20**	**16**	**14**	**11**	**17**	**15**	**15**	**37**	**27**	**31**

4 Discussion and Conclusion

It can be seen in Fig. 3 that there is a substantial difference between the BCI performance with the different frequency sets. Subjects reached generally higher ITR with the lower stimulation frequency sets. 50% of the subjects were not

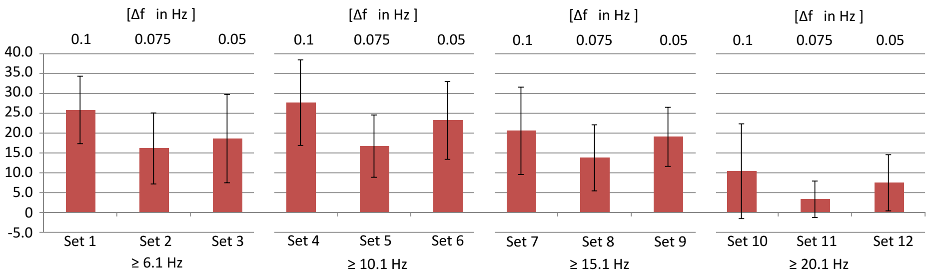


Fig. 3. Average ITRs for the different frequency sets.

able to control the system with the frequency sets 10, 11 and 12 (all frequencies above 20 Hz) at all. The performance in each frequency domain is best with the 0.1 Hz resolution set. Surprisingly, the sets with 0.05 Hz resolution (sets 3, 6, 9, 12) yielded better results than the 0.075 Hz resolution. A reason for this could be a training effect, as the low resolution tasks were recorded last. However, the results indicate, that the previously reported maximal frequency resolution ($\geq$ 0.1 Hz) for SSVEP-based BCIs [3,4,7] can be updated to 0.05 Hz. Interestingly, the results with the sets with lower frequencies with the highest resolution (set 3 and set 6) are still better than those of the highest frequency set with the lowest resolution. These results confirm the assumption, that an equidistant frequency selection for multi target BCI-systems is not optimal. The results suggest, that the lower frequencies can be spread denser than the higher once. So, when frequencies are spanned over a large range (e.g. from 8 to 20Hz), selection of frequencies from a low frequency-range (below 12 Hz) with denser resolution, while selecting lesser frequencies from a middle-frequency range (12-30 Hz) with a sparser resolution, could yield overall a better BCI performance.

Acknowledgments. This research was supported by the German Federal Ministry of Education and Research (BMBF) under Grants 16SV6364 and 01DR14014. We also thank all the participants of this research study.

References

1. Cecotti, H.: A self-paced and calibration-less SSVEP based Brain-Computer Interface speller. IEEE Trans. Neural Syst. Rehabil. Eng. **18**(2), 127–133 (2010)
2. Chen, X., Chen, Z., Gao, S., Gao, X.: A high-ITR SSVEP-based BCI speller. Brain-Computer Interfaces **1**(3–4), 181–191 (2014)
3. Chen, X., Wang, Y., Nakanishi, M., Jung, T.P., Gao, X.: Hybrid frequency and phase coding for a high-speed SSVEP-based BCI speller. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3993–3996 (2014)
4. Gao, X., Xu, X., Cheng, M., Gao, S.: A BCI-based environmental controller for the motion-disabled. IEEE Trans. Neural Syst. Rehabil. Eng. **11**(2), 137–140 (2003)

5. Gembler, F., Stawicki, P., Volosyak, I.: Towards a user-friendly BCI for elderly people. In: Proceedings of the 6th International Brain-Computer Interface Conference Graz (2014)
6. Gembler, F., Stawicki, P., Volosyak, I.: A comparison of SSVEP-based BCI-performance between different age groups. In: Rojas, I., Joya, G., Catala, A. (eds.) IWANN 2015. LNCS, vol. 9094, pp. 71–77. Springer, Heidelberg (2015)
7. Hwang, H.J., Lim, J.H., Jung, Y.J., Choi, H., Lee, S.W., Im, C.H.: Development of an SSVEP-based BCI spelling system adopting a QWERTY-style LED keyboard. Journal of Neuroscience Methods **208**(1), 59–65 (2012)
8. Kick, C., Volosyak, I.: Evaluation of different spelling layouts for SSVEP based BCIs. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1634–1637. IEEE (2014)
9. Nakanishi, M., Wang, Y., Wang, Y.T., Mitsukura, Y., Jung, T.P.: Generating visual flickers for eliciting robust steady-state visual evoked potentials at flexible frequencies using monitor refresh rate. PloS one **9**(6), e99235 (2014)
10. Pastor, M.A., Artieda, J., Arbizu, J., Valencia, M., Masdeu, J.C.: Human cerebral activation during steady-state visual-evoked responses. J. Neurosci. **23**(37), 11621–11627 (2003)
11. Volosyak, I.: SSVEP-based Bremen-BCI interface - boosting information transfer rates. J. Neural Eng. **8**(3), 036020 (2011)
12. Volosyak, I., Cecotti, H., Gräser, A.: Impact of frequency selection on LCD screens for SSVEP based brain-computer interfaces. In: Cabestany, J., Sandoval, F., Prieto, A., Corchado, J.M. (eds.) IWANN 2009, Part I. LNCS, vol. 5517, pp. 706–713. Springer, Heidelberg (2009)
13. Volosyak, I., Cecotti, H., Gräser, A.: Optimal visual stimuli on LCD screens for SSVEP based Brain-Computer Interfaces. In: Proc. 4th Int. IEEE/EMBS Conference on Neural Engineering NER 2009, pp. 447–450, May 2009
14. Wang, Y., Wang, Y.T., Jung, T.P.: Visual stimulus design for high-rate SSVEP BCI. Electronics Letters **46**(15), 1057–1058 (2010)
15. Wolpaw, J., Wolpaw, E.W. (eds.): Brain-computer interfaces: principles and practice. Oxford University Press (2011)
16. Zhu, D., Bieger, J., Molina, G.G., Aarts, R.M.: A survey of stimulation methods used in SSVEP-based BCIs. Computational Intelligence and Neuroscience **1**, 702357 (2010)

EEG Correlates of Visual Recognition While Overtly Tracking a Moving Object

Marija Uščumlić(✉), Miriam Hägele, and Benjamin Blankertz

Neurotechnology Group, Technische Universität Berlin, Berlin, Germany
{marija.uscumlic,haegele.miriam}@gmail.com,
benjamin.blankertz@tu-berlin.de

Abstract. Although our natural visual environment is dynamic, to date EEG studies on visual cognition are mainly based on the fixed-gaze visual paradigms or static images as stimuli. On the other side, scenes' dynamic significantly influence our visual behavior, i.e., the occurrence of saccadic movements, smooth pursuit and fixations. Since smooth-pursuit eye-movements do not occur in a static scene, in this study we address the EEG-based intention decoding in presence of smooth-pursuit eye-movements at slow speed ($\sim 2.8°$/s) using the state-of-the-art EEG decoding methods. Our results suggest that the decoding performance remain high (with reference to the fixed-gaze paradigm) even when subjects are additionally engaged in tracking a moving object. In contrast to the pursuit movements, the uncertainty of the change perception remains one of the major challenges for the EEG decoding as we additionally demonstrated in this study.

Keywords: Electroencephalography (EEG) · Event-related potentials · Smooth-pursuit · Visual recognition

1 Introduction

Visual cognition allows us to perform decision-based scene analyses in our daily activities. We successfully sample task-relevant information from the scene by employing different voluntary eye-movements such as saccades and smooth-pursuit eye movements. Saccades represent a sudden shift of gaze ending in the fixation, while smooth-pursuit eye-movements represent foveal tracking of a moving object. Interestingly, in contrast to saccades the smooth-pursuit eye movements can not be performed in absence of the moving stimulus, i.e., when presented with a static scene.

Brain imaging experiments on visual cognition are mainly designed to isolate certain cognitive processes. Consequently, the experiments are based on simplified stimuli and temporally-controlled sensory input. Additional advantage of the simplified and strictly controlled paradigms is minimization of non-neural sources of EEG artifacts, in particular ocular artifacts. While the fundamental cognitive neuroscience knowledge advances, the simplification of the experiments

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 166–171, 2015.
DOI: 10.1007/978-3-319-24917-9_18

limits the knowledge transfer to real-life scenarios for brain-computer interface (BCI) applications. The main reason is the discrepancy between the experimental setting and real-world scenarios with respect to visual behavior and parallel cognitive processes.

The conventional visual oddball paradigm that elicit cognitive P300 visual evoked potential in response to the target stimulus recognition is a typical example of the fixed-gaze paradigm [5], [6]. Stimuli are rapidly presented one after another at the same position in the screen at a constant presentation rate, while the percentage of target stimuli in the sequence is kept low. The P300 potential is elicited over the centro-parietal scalp region not earlier than 300 ms of target stimuli onset. Advances in eye-tracking technology established a new trend in studying visual cognition – i.e., joint recording of neural signals and eye-movements, that allows studying the cognitive processes in active visual tasks. Reported results evidence that the fixation-related EEG potentials when attending the target content closely resemble the ERPs in the classical visual oddball paradigm [3], [1], [2], [4].

In the context of real-life scenarios, we are typically faced with complex spatio-temporal scenes as the following examples demonstrate. In a busy public place such as a train station, while waiting for someone we inspect people entering our visual field. This requires tracking of a person until it is close enough that we can recognize if it is the person we are waiting for. When playing a video game we often control behavior of an avatar that navigates through the space and interacts with the static and moving objects in the scene, a situation which also requires overt tracking of the content in motion. Finally, similar ocular behavior occurs while watching long tracking shots in a movie, e.g., following a character we move through the street.

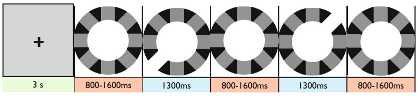
Motivated by our visual behavior in complex spatio-temporal scenes, the present study evaluates the EEG responses to the visual recognition event in an oddball task while a subject overtly tracks a moving object at low speed $(2.8°/s)$. We considered a fixed-gaze oddball task as a reference condition. In addition, we compared the decoding performance between two cases of the stimulus appearance style – an instantaneous stimulus appearance, and its transient appearance, that introduces temporal uncertainty in the perception process.

2 Methods

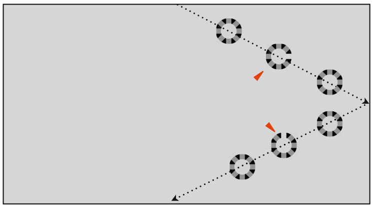
2.1 Subjects and Data Acquisition

Subjects were seated in front of the screen (1680 X 1050 pixels, 60Hz) with their head resting in a chin rest positioned at $\sim$ 61 cm from the screen. Six subjects participated in the study (2 male and 4 female subjects, between 23 and 28 age old). The EEG was recorded using a Brain Products actiCAP active electrode system with 64 electrodes (International $10-20$ system), 1000 Hz. The EEG was referenced to the linked-mastoid and filtered between 1 and 30 Hz. Single-trial data epochs were extracted with respect to the stimuli appearance as specified in Section 2.3.

(a) Modified Landolt rings



(b) Timeline of stimuli presentation.



(c) The *Smooth-pursuit* condition

Fig. 1. Illustration of the stimuli and experimental protocols.

2.2 Experimental Protocol

Subjects were asked to perform a visual search task, silently counting the target stimuli in a sequence. At the end of each sequence they verbally reported the number of recognized targets. Before each sequence, subjects were presented with the target stimulus. We used modified Landolt rings as stimuli (Figure 1a). A target was randomly selected from eight broken Landolt rings (each ring had the unique gap position). The remaining seven rings were used as non-target stimuli.

The experiment consisted of three conditions: *i* - **Static condition (St)**, in which subjects were instructed to look at the center of the screen where stimuli sequence was presented; *ii* - **Smooth-pursuit (Sp) condition**, in which subjects were asked to overtly track a moving stimulus (a Landolt ring). Starting from the center of the screen a Landolt ring moves with the constant speed ($2.8°/s$) in the same direction until it bounds from a field's edge and change its direction (Figure 1c). The field is defined as a rectangular gray area (1260 by 735 pixels, the gray-level value: 0.7) centered in the screen; *iii* - **Fading smooth-pursuit (FSp) condition**, that differs from the (Sp) condition only with respect to the way the target/distactor stimuli appear in the scene. Namely, while in the (Sp) condition the stimuli appear instantaneously in the scene, in the (FSp) condition their appearance is rather a transient process of 1000ms. Note that in all the conditions, a Landolt ring is constantly present in the screen. Therefore, appearance of a new stimulus results in an appearance of the gap in one of eight possible positions.

In all three conditions, the inter-stimulus interval was uniformly distributed between 800 ms and 1600 ms, while the target and non-target stimuli were presented for 1300ms. The timeline of stimuli presentation is shown in Figure 1b. Experiment was organized in 10 blocks. A block contained three runs, i.e., a search task per each condition, in a random order. Stimuli sequence consisted of forty stimuli, where the occurance of targets in a sequence was within the range $20 - 30\%$.

2.3 Discriminative EEG Analysis

The
EEG discriminant analysis is performed independently for each experimental condition (St, Sp and FSp). The analysis consisted of three steps. First, we estimated the discrimination between target and non-target trials for each channel and time point using signed squared biserial correlation coefficient (sgn r^2) [7]. Then, we created the temporal profile of the discriminant information by applying a shrinkage LDA classifier [7] at each time point using all the EEG channels, in a ten-fold cross-validation setting. Finally, we estimated single-trial decoding performance using a hierarchical LDA classifier integrating spatio-temporal discriminative activity over a trial [8]. The considered time intervals (from 100 ms to 800 ms of stimulus onset) was divided in seven non-overlapping windows.

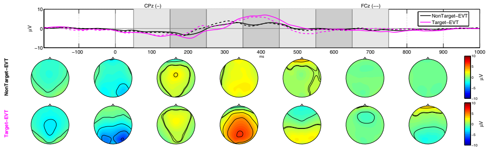
3 Results and Discussion

The grand average ERP potentials are presented separately for each experimental condition in Figure 2. One can notice a positive activity over the centro-parietal region evoked by target stimuli peaking at 400 ms of stimulus onset in the *Static* and *Smooth-pursuit* condition (Figure 2a-b). Similar component is found in the *Fading Smooth-pursuit* condition, but temporally more spread between 450 ms and 650 ms of stimulus onset (Figure 2c). This temporal shift and larger span might indicate the uncertainty of the perception of stimulus over trials due to the fading effect.

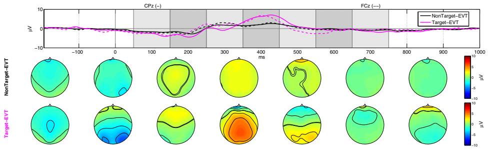
Single-trial decoding performance are comparable between the *Static* (St) and the *Smooth-pursuit* (Sp) condition, while a decrease in performance is found in the *Fading Smooth-pursuit* (FSp) condition (Figure 4). These results suggest that smooth-pursuit eye-movements at low speed do not challenge the EEG decoding of visual recognition if the timing of the event is known. Interestingly, median performance of the *Smooth-pursuit* condition was better than in the *Static* condition. This might indicate that tracking an object as a background task increase the subjects' engagement in the primary search task. Behavioral responses support this idea, since more erroneous answers were reported in the (St) condition than the (Sp) condition (see Table 1). Future work will investigate this research question further and in particular acquire data from a large number of participants in order to be able to test for statistical significance.

Table 1. Responses: Number of erroneous runs (mean absolute error).

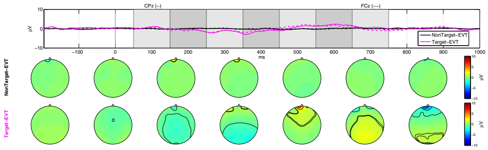
Protocol/Subjects	S1	S2	S3	S4	S5	S6	Sum
St	0 (0)	1 (1)	3 (1.3)	0 (0)	1 (1)	4(1)	9
Sp	0 (0)	0 (0)	0 (0)	0 (0)	1 (3)	2 (1)	3
FSp	0 (0)	2 (1)	3 (1)	0 (0)	2 (1.5)	0 (0)	7



(a) Static condition

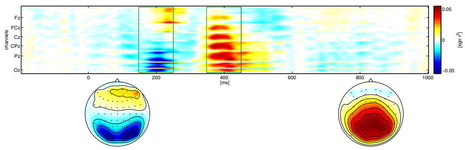


(b) Smooth-pursuit condition

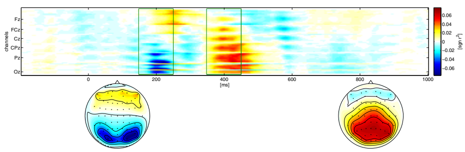


(c) Fading Smooth-pursuit condition

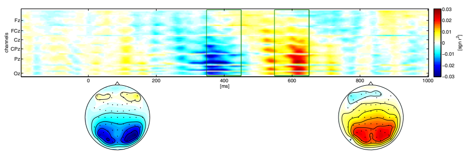
Fig. 2. Grand average event-related potentials.



(a) Static condition



(b) Smooth-pursuit condition



(c) Fading Smooth-pursuit condition

Fig. 3. Visualization of signed r^2-values.

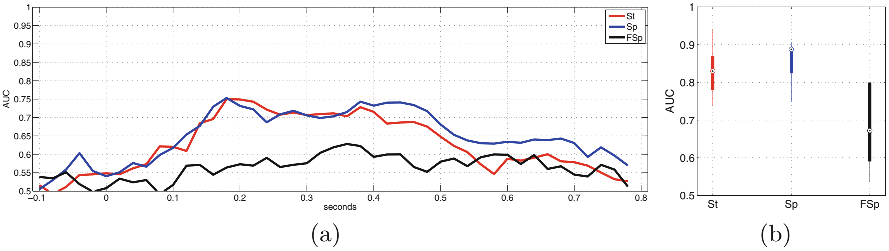


Fig. 4. (a) Temporal profile of discriminative activity. (b) Single-trial classification performance.

4 Conclusion

The aim of the present study was to investigate the EEG-based decoding performance of visual recognition during the slow-speed smooth pursuit eye-movements. Our preliminary results indicated that the overt tracking of a moving object do not challenge the state-of-the-art EEG decoding if the precise time of the event is known. Moreover, the results motivate further studies on potential favorable effects which overt tracking might have on the EEG-based recognition decoding.

Acknowledgments. This study is funded by the European Union Seventh Framework Programme (FP7/2007-2013) (grant agreement #611570), and in part by the BMBF (contract 01GQ0850 BFNT).

References

1. Kamienkowski, J.E., Ison, M.J., Quiroga, R.Q., Sigman, M.: Fixation-related potentials in visual search: A combined EEG and eye tracking study. Journal of Vision **12**(7) (2012)
2. Brouwer, A.M., Reuderink, B., Vincent, J., van Gerven, M.A., van Erp, J.B.: Distinguishing between target and nontarget fixations in a visual search task using fixation-related potentials. Journal of Vision **13**(3) (2013)
3. Rämä, P., Baccino, T.: Eye-fixation related potentials (efrps) during object identification. Visual Neuroscience **27**, 187–192 (2010). 11
4. Jangraw, D.C., Wang, J., Lance, B.J., Chang, S.F., Sajda, P.: Neurally and ocularly informed graph-based models for searching 3d environments. Journal of Neural Engineering **11**(4), 046003 (2014)
5. Farwell, L.A., Donchin, E.: 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)
6. Polich, J.: Updating P300: An integrative theory of P3a and P3b. Clinical Neurophysiology **118**, 2128–2148 (2007)
7. Blankertz, B., Lemm, S., Treder, M.S., Haufe, S., Müller, K.R.: Single-trial analysis and classification of ERP components – a tutorial. NeuroImage **56**, 814–825 (2011)
8. Gerson, A.D., Parra, L.C., Sajda, P.: Cortically coupled computer vision for rapid image search. IEEE Transactions on Neural and Rehabilitation Systems Engineering **14**(2), 174–179 (2006)

Neural Responses to Abstract and Linguistic Stimuli with Variable Recognition Latency

Markus A. Wenzel(✉), Carlos Moreira, Iulia-Alexandra Lungu, Mihail Bogojeski, and Benjamin Blankertz

Neurotechnology Group, Technische Universität Berlin, Berlin, Germany
markus.wenzel@tu-berlin.de

Abstract. Electroencephalography (EEG) can provide information about which words or items are relevant for a computer user. This implicit information is potentially useful for applications that adapt to the current interest of the individual user. EEG data were used to estimate whether a linguistic or abstract stimulus belonged to a target category that a person was looking for. The complex stimuli went beyond basic symbols commonly used in brain-computer interfacing and required a variable assessment duration or gaze shifts. Accordingly, neural processes related to recognition occurred with a variable latency after stimulus-onset. Decisions involving not only shapes but also semantic linguistic information could be well detected from the EEG data. Discriminative information could be extracted better if the EEG data were aligned to the response than to the stimulus-onset.

Keywords: EEG · Single trial classification · Physiological computing · User relevance estimation

1 Introduction

Physiological sensors can capture neural processes during human-computer interaction and can potentially provide access to information about the user that is otherwise not accessible. The performance of a wide range of applications could benefit from collecting user-related information directly with neurophysiological sensors instead of observing the user's behaviour or asking questions. The computer could, for instance, estimate which words or items are relevant for the user. Exploiting this implicit information promises to allow completely new scenarios where devices adapt to the current interest of the individual user.

In this study, information present in the electroencephalogram (EEG) was used to estimate whether a stimulus was relevant for a person because it belonged to a target category that required a particular response. The stimuli were either words or abstract items and required a variable assessment duration or gaze shifts, which can be expected in application scenarios out-of-the-lab. Accordingly, neural processes related to recognition occurred with a variable latency after stimulus-onset, which poses a challenge for EEG-based detection algorithms. For this reason, neural processing of complex linguistic and abstract stimuli was characterised in the present study.

B. Blankertz et al. (Eds.): Symbiotic 2015, LNCS 9359, pp. 172–178, 2015.
DOI: 10.1007/978-3-319-24917-9_19

2 Materials and Methods

2.1 Experimental Design

The experiment was a variation of the classic oddball paradigm where rare stimuli, that are being paid attention to, elicit a different neural response than other, more frequently presented stimuli, that are not being focused on [1]. Abstract and linguistic stimuli were presented on a 22-inch screen while EEG was recorded. The subjects were asked to press either the right or the left arrow key if each stimulus was part of an announced target category or not. The participants used the same hand for both keys to avoid that response-related activity in the motor cortex was discriminable between targets and non-targets. The three experimental conditions covered the range from relatively complex linguistic stimuli (*Semantic, Pseudo-words*) to rather simple stimuli (*Abstract – Shapes*):

Linguistic – Semantic (LS). Decision if a word belongs to a semantic category. Words were selected that were part of certain semantic categories, i.e. groups of related words. Target words belonged to a designated semantic category, while non-target words belonged to other semantic categories. For example, the participant had to press the target key when the current target category was 'furniture' and the word presented was 'chair', whereas the participant had to press the non-target key if the word 'water' was shown. Lists of fifteen words each from ten different categories were assembled – covering both easier categories (e.g. 'animals', 'food') and more broad or difficult ones (e.g. 'science', 'time'). Depending on the mother tongue of the participant, English or German words were used.

Linguistic – Pseudo-words (LP). Decision if a word is a pseudo-word or a real word. Pseudo-words are words that could exist in a given language but do not have any meaning, e.g. "diagrant", "persided", "occation" and "yeanings". In this experimental condition, targets were either pseudo-words or true words. A list of pseudo-words was generated by picking random letters with a Markov process using the most common words from either English or German as input [2]. The resulting pseudo-words were hand-picked to a total of 120 to avoid close resemblances to real words in each respective language and to maintain a similar letter per word distribution to the chosen true words.

Abstract – Shapes (AS). Decision if a polygon has a specific number of corners. Several different tetragons, pentagons, hexagons and octagons were prepared respectively (c.f. figure 1). Polygons with a specific number of corners were the respective targets, all other polygons were non-targets. The angles within each polygon type were varied to enforce counting and to preclude the immediate recognition of the shape itself from memory. Heptagons were not used because it was difficult to discriminate them from octagons.

Each experimental session contained thirty runs, with ten runs per condition (LS, LP and AS). In each run, 60 stimuli were presented with a 1:3 target to non-target ratio. Before each run, the target category or shape was announced.

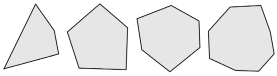


Fig. 1. Examples of polygons used as stimuli.

After a preparation countdown of three seconds, the stimuli were presented one by one in the centre of the screen for 1000 ms, preceded by a fixation cross during 500 ms. After stimulus presentation and the next fixation cross, the screen turned blank for 500 ms. The order of the runs was randomised. Each stimulus occupied a larger area and required the evaluation of different parts (and possibly gaze shifts) in order to solve the task.

2.2 Data Acquisition

Three female and three male subjects participated in the experiments. The EEG signals were recorded with 64 active electrodes (*BrainAmp, ActiCap*, BrainProducts, Munich, Germany; sampling frequency of 1000 Hz). The electrodes were positioned according to the international 10–20 system, using the linked mastoid as reference and an electrode on the forehead as ground. One of the electrodes was positioned below the left eye for electrooculography (EOG). The study was approved by the ethics committee of the Department of Psychology and Ergonomics of the Technische Universität Berlin.

2.3 Data Analysis

EEG Data Processing. The multi-channel EEG data were low-pass filtered for anti-aliasing (2nd order Chebyshev, 42 Hz pass-band, 49 Hz stop-band), down-sampled to 100 Hz, re-referenced to the linked-mastoids and high-pass filtered to reduce drifts (FIR least square filter, 0.5 Hz). Artefacts were rejected based on a variance criterion and the time-series were segmented into short epochs aligned either to the onset of the *stimulus* or to the *response* of the participant. Only correct responses were considered for the further analysis. Artefact epochs were removed based on a maximum-minimum criterion (peak difference over 150 μV within the interval [0 ms, 1200 ms] for stimulus- and [-950 ms, -50 ms] for response-aligned epochs). The epochs were baseline corrected by subtracting the average signal measured during an interval of 100 ms before the stimulus-onset (and respectively within the interval [-950 ms, -850 ms] before the response) and averaged for target and non-target stimuli. To investigate which channels and time points contained discriminative information between targets and non-targets, the area under the curve (AUC) of the receiver operating characteristic [3] was computed for each channel and time point.

Classification of the EEG Epochs. Target versus non-target EEG epochs (either stimulus-aligned or response-aligned) were classified with regularized linear discriminant analysis – the shrinkage parameter was determined analytically [4] [5].

Classification performance was evaluated in 10x10-fold cross-validations using the AUC as metric [3]. Spatio-temporal features were used for the classifications [6]: The means were computed of 50 ms intervals from stimulus-onset to 1250 ms after stimulus-onset, as well as the means of 50 ms intervals from 350 ms before to 400 ms after the response. For the classifications, the EOG channel was removed.

3 Results

Participants' Behaviour. The six participants gave correct responses in over 90 % of the cases, with a slightly lower performance in the condition 'Linguistic – Semantic' in comparison to the two other conditions. The latencies between stimulus presentations and corresponding responses of the participants via key press are presented as histograms in figure 2. In all three experimental conditions, the latency of recognition was variable. Average latencies for targets were larger than for non-targets (LS 759 ms/721 ms, LP 775 ms/738 ms, AS 750 ms/700 ms).

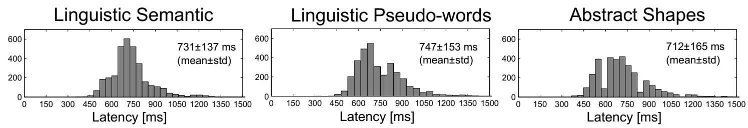


Fig. 2. Response latency histograms for the three conditions.

EEG Epochs Aligned to Stimulus-Onset and Response. The event-related potentials for the experimental condition 'Linguistic – Semantic' are presented in figure 3. The statistical differences between target and non-target EEG epochs are depicted in figure 4 for each experimental condition. Stimuli elicited late positive components at central electrodes. The components were larger for targets than for non-targets, as it can be expected for the oddball paradigm [1]. Augmented positive components at central sites also preceded the participants' responses to targets.

Classification of Target and Non-Target EEG Epochs. Classification performance was above chance level (0.5) in all conditions (cf. figure 5). Alignment to the responses resulted in a better performance than when the EEG epochs were aligned to the stimulus-onsets.

4 Discussion

Linguistic and Abstract Stimuli. Decisions involving not only basic information coded in simple abstract shapes but also semantic linguistic information are well detectable from EEG data (cf. figure 5). This finding is promising for application cases where the computer would assign a user-interest-level to each word displayed on the screen.

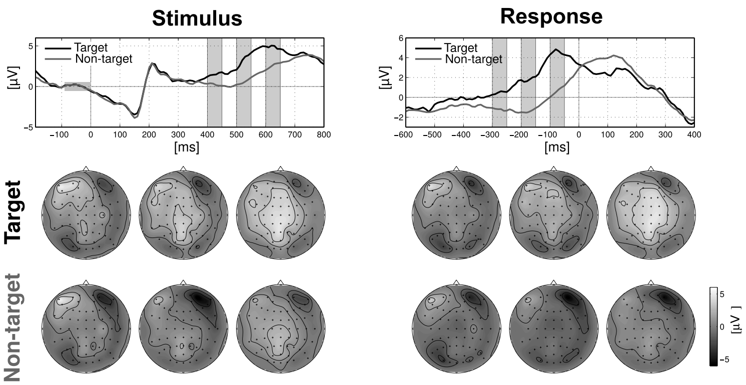


Fig. 3. Event-related potentials for the experimental condition 'Linguistic – Semantic' averaged across-subjects. Epoch alignment to stimulus (*left*) and response (*right*). The time courses at electrode Cz are displayed at the *top* of the figure. The scalp topographies (*bottom*) indicate the potentials at all electrodes within the three intervals shaded in grey in the time courses.

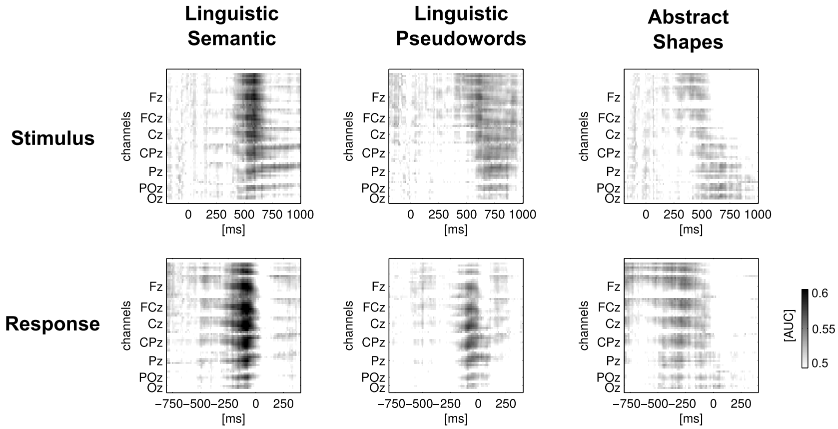


Fig. 4. AUC-score matrices for the three experimental conditions (*columns*). Epoch alignment to stimulus (*top*) and response (*bottom*).

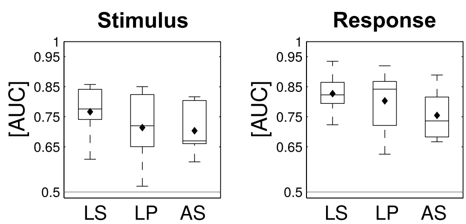


Fig. 5. ERP classification results of each condition for all participants. Epoch alignment to stimulus (*left*) and response (*right*).

EEG Epochs Aligned to Stimulus-Onset and Response. The electroencephalogram reflects multiple simultaneous brain processes that are difficult to differentiate. The process of interest can be extracted by averaging epochs of the continuous signal that are aligned to a repeated event of one type (event-related potential). Brain activity unrelated to this time point of reference is averaged out. The characteristics of sensory processing can be uncovered by analysing the change in the brain activity after stimulus-onset (cf. figure 3, *left*, and figure 4, *top*). In this study, the participants responded as soon as they had recognised whether a stimulus was a target or not. Thus, the response indicated the moment of target recognition. Due to the complex stimuli used, the latency of recognition was variable with respect to the stimulus-onset (cf. figure 2). The alignment of the EEG data to the subjects' responses improved the classification performance (cf. figure 5). Probably, response-aligned epochs were more informative than stimulus-aligned epochs because the temporal variability of discriminative neural activity within the epochs was reduced (cf. figure 4). However, it has to be considered that the response latencies were larger for targets than for non-targets, which constitutes a confounding factor for the response-locked analyses. The response-aligned EEG epochs captured not only the EEG potential evoked by recognition but also the EEG potential evoked by the stimulus. The time points captured from the latter differed between classes, because the response latencies were different on average. Accordingly, classification performance could be improved by this factor, too.

Conclusion. For more complex decision tasks, discriminative information is rather response than stimulus locked. However, a user response to each stimulus can not be expected in 'real-world' scenarios. For applications without responses, appropriate classification techniques need to be developed. Future research will strive to improve classification performance in the absence of an overt user response.

Acknowledgements. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 611570, and in part by the BMBF (contract 01GQ0850, BFNT). The authors want to thank the anonymous reviewers for very helpful comments.

References

1. Picton, T.W.: The p300 wave of the human event-related potential. J. Clinical Neurophysiology **9**(4), 456–479 (1992)
2. ibbly - Pseudo-words. (acc: September 18, 2014) http://ibbly.com/Pseudo-words.html

3. Fawcett, T.: An introduction to ROC analysis. Pattern Recogn. Lett. **27**(8), 861–874 (2006)
4. Friedman, J.H.: Regularized discriminant analysis. J. Am. Stat. Assoc. **84**(405), 165 (1989)
5. Schäfer, J., Strimmer, K.: A Shrinkage Approach to Large-Scale Covariance Matrix Estimation and Implications for Functional Genomics. Statistical Applications in Genetics and Molecular Biology **4**(1) (2005)
6. Blankertz, B., Lemm, S., Treder, M., Haufe, S., Müller, K.-R.: Single-trial analysis and classification of ERP components – a tutorial. NeuroImage **56**(2), 814–825 (2011)

Author Index