Everyday Technology for Independence and Care
G.J. Gelderblom et al. (Eds.)
IOS Press, 2011

doi:10.3233/978-1-60750-814-4-481

Usability of Brain Computer Interfaces

Emanuele PASQUALOTTO[a,b,1], Stefano FEDERICI[b,c], Alessandro SIMONETTA[b,d], Marta OLIVETTI BELARDINELLI[b,e]

[a]*Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Tübingen, Germany*
[b]*Interuniversity Center for Research on Cognitive Processing in Natural and Artificial Systems (ECoNA), Sapienza University of Rome, Rome, Italy*
[c]*Department of Human and Education Sciences, University of Perugia, Perugia, Italy*
[d]*INAIL – Consulenza per l'Innovazione Tecnologica, Rome, Italy*
[e]*Department of Psychology, Sapienza University of Rome, Rome, Italy*

Abstract

Objective The aim of this work is to assess the usability of two BCI prototypes by measuring interaction with the systems in context, considering the performance, cognitive workload and satisfaction of non-disabled users in order to better understand how the interface affect these parameters. We tested two keyboard-controlled Java BCI prototypes based on the Language Support Program (LSP) and the P300 Speller (P3S).

Main Content We performed two different evaluations. In the first one, we tested the learnability of BCIs on 6 healthy users through the Thinking Aloud technique. Then, we tested BCI efficiency on 30 participants through the Copy Spelling Task (CST) and we administered the System Usability Scale (SUS) to measure usability and the Survey of Technology Use (SOTU) scale of the Matching Person and Technology (MPT) to measure predisposition to the use of technology. In the second evaluation, we tested again using 61 participants with different computer skills, and administered usability and cognitive workload questionnaires.

Results The first test showed that all users easily learned how the system worked with the LSP, but failed with P3S. We found that P3S users were more accurate in selecting and recognising letters on the screen. Both SUS and SOTU did not show any significant effects. In the second one, the results showed differences in the number of errors, in user satisfaction and in the cognitive workload.

Conclusions We found that the Thought Translation Device was more error-resistant, less stressful and more satisfactory for the users compared to the P3S.

Keywords Brain Computer Interfaces, Neurodegenerative disease, Amyotrophic lateral sclerosis, Usability, Psychotechnology, Human Computer Interaction.

[1] Corresponding Author: Emanuele Pasqualotto, Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Gartenstraße 29, D-72074 Tübingen, Germany. E-mail:emanuele.pasqualotto@uni-tuebingen.de

Introduction

A Brain Computer Interface (BCI) is a control and/or communication system in which the user's commands and messages do not depend on muscular control; information is not conveyed directly from nerves and muscles, and muscular activity is not necessary for the production of the signal that is needed to convey the message [10].

Since BCIs could provide paralysed people a means of communicating, this technology should be considered and treated as an assistive technology that facilitates daily activities [5; 6]. Therefore, BCIs need an assessment process to match person and technology in order to avoid dissatisfaction and abandonment [15].

Approaches such as "User-centered design" [2], "User interfaces for all" [1] and the most recent "Integrated model of usability" [24; 27] already highlight the importance of a complete and full evaluation of the interaction. As already pointed out by De Kerckhove [3; 4], technology has a big impact on the properties of the mind. He defined the role of psychotechnology as an artefact capable of turning the human mind into an electronic sensory extension. Federici overtook this definition, embracing it in an intrasystemic perspective [17]. Following this approach, he defined psychotechnology as any "technology that emulates, extends, amplifies and *modifies* sensory-motor, psychological or cognitive functions of the mind" [14; 22]. The "modification" component that the author included in the original definition emphasised that any technology is able to permit the adaptation of the human being to the environment-system and, at the same time, force users toward cognitive and cultural modification and adaptation [25].

1. Aim

The tradition of Human Computer Interaction (HCI) has already given us most of the tools we need to analyse and evaluate technology. Although studies have aimed at evaluating human factors, the tools developed in HCI are still poorly used in the design of BCIs. During the development of BCIs, there is a lack of consideration of the human factors involved in user-technology interaction. Such factors could lead to better structuring of the information, with a consequent increase of bit rate and writing accuracy, and thus allowing a more usable communication. Moreover, more pleasant and effective interfaces could lead to more motivating and satisfactory conditions, minimising the risk of discontinuation.

In this study, we aimed to assess the usability of two BCI prototypes by measuring interaction with the systems in context, considering the performance, cognitive workload and satisfaction of non-disabled users in order to better understand how the interface affect these parameters. We planned two studies. The first one was an exploratory assessment study (see paragraph 3 below) of two BCI prototypes divided into two parts: the first part to evaluate the learnability of the prototypes and the second part to evaluate the ease of use. In both cases, our aim was also to test our material and methods for the subsequent study. The second study (see paragraph 4 below) was an extensive assessment performed using the same prototypes.

2. Java mock-up

BCIs are usually EEG-, MEG- or fMRI-based. The BCI systems are normally developed by using two separate functional blocks: the transducer and the control interface [28]. The transducer translates the user's brain activity into usable control signals. The BCI component uses these control signals to perform a desired activity or function. Due to the difficulties in the processing of the transducer, several issues

Figure 1 The LSP works by halving the letters in two blocks at each new step. The user can accept the block if it contains the target letter or reject it if it does not.

associated with BCI are related to this block. In order to evaluate only visual layout features and to remove all the EEG and transducer-related issues, we developed two keyboard-controlled Java prototypes. We decided to investigate the Language Support Program (LSP) used in the Thought Translation Device (TTD) [18; 19; 29] as well as the P300 Speller (P3S) [7; 12]. While the TTD (see Figure 1) is based on Slow Cortical Potentials (SCP) and allows users to select between two alternatives (e.g., yes and no), the P3S is based on the P300 event-related potential and allows users to use a flashing 6x6 matrix of characters. We developed the prototypes in such a way that they can be controlled using a PC keyboard, by considering the original nature of the input.

The original LSP works by halving the block of letters at each new step [19]. These letters are constantly displayed in two groups in order to be selected using dichotomous logic. At each step, the user can accept the block if it contains the target letter or reject it if it does not.

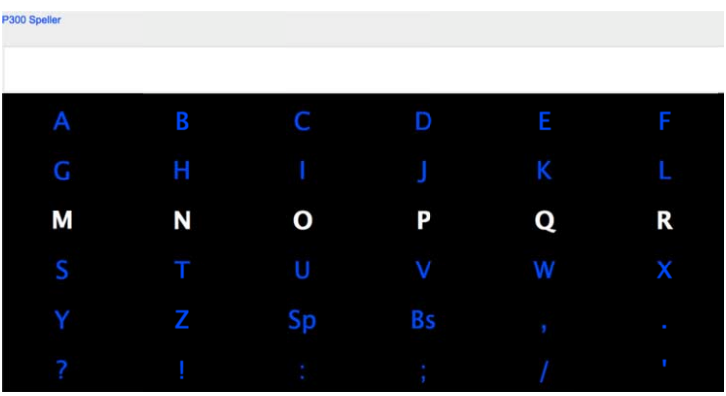

Figure 2 The P3S is a matrix of letters, where each row and each column are intensified and the intensifications are presented in a random sequence.

Our version of the LSP works in the same way, in addition to the fact that it is controlled by the arrow keys on the keyboard. By selecting the left arrow key, the user can accept the displayed half while by selecting the right arrow key, the user can reject the displayed half and pass on to the next one.

The P300 Speller described by Farwell and Donchin [12] presents a 6x6 matrix of symbols (see Figure 2). Each row and each column are intensified and the intensifications are presented in a random sequence. The user focuses attention on the desired cell in the matrix. The intensification of rows and columns constitutes the Oddball Paradigm necessary to elicit a P300 event-related potential [13].

By using one key to select when the target appears, it is possible to control the P3S prototype developed by our team. As with the P300 Speller, it is necessary to press the key twice to hit a target, once for the row and once for the column where the target is. Due to the EEG components used, the P300 Speller and the LSP represent two different ways of using input in a speller device: single hit and double hit modality, respectively.

3. Assessment of learnability and ease of use

3.1. Learnability

3.1.1. Procedure

The aim of this study was to assess which interface, between the LSP and P3S, is more supportive for the user in the learning or re-learning phase. The more the interface is self-explanatory, the shorter the learning procedure is.

An assessment of the interface using the Thinking Aloud (TA) technique [11; 21; 26] was performed on six subjects, who were healthy gender-balanced university students (mean age: 24.8, range: 22 to 28). All subjects were presented with both interfaces, balancing the presentation order, and none had ever seen the interfaces before. Subjects were informed of the final aim of the interfaces but not about how they worked. They were instructed first to understand the functioning of the interfaces and then to perform a Copy Spelling Task (CST).

3.1.2. Results

All six users of the LSP understood the system's functioning in less than 9 minutes. The results of this first exploratory study indicated that the LSP had a more intuitive, conceptual and functioning model compared to that of the P3S.

Regarding user satisfaction, all users explicitly declared that the LSP was more intuitive. However, two of the six subjects preferred the P3S, showing how the routine quality of the system allowed greater control of the situation.

3.2. Ease of use

3.2.1. Procedure

In order to evaluate the ease of use of the two interfaces, we administered the CST in the same way as the previous test to 30 healthy gender-balanced subjects (mean age: 27.6, range: 19 to 42). The subjects were divided into two groups, one for each interface. At the end of the tasks, we administered the Survey of Technology Use Version C (SOTU-C) [15; 16; 23] scale and the System Usability Scale (SUS) [9; 20].

3.2.2. Results

We used the Number of Errors made during the CST as a measure; moreover, we differentiated between two types of error: Choice-based Error and a Recognition-based Error. The former represented either a wrong choice or a false positive (a choice that led to the wrong result) while the latter represented the lack of recognition of the letter to be selected or a false negative, leading to continued search for the desired target. We analysed SUS as a global score and as a two-factor scale: SUS Usability and SUS Learnability [20].

We analysed the data by means of an analysis of variance (ANOVA) using the variable Interface (LSP and P3S) as an independent variable and the scores of SUS, SOTU-C, Number of Errors, Choice-based Error, and Recognition-based Error as dependent variables. Results showed a difference in the use of the two interfaces in favour of the P300 Speller. The differences in the global Number of Errors ($F_{(1,28)}$=7.84, p=.009), Choice-based Error ($F_{(1,28)}$=5.18, p=.031) and Recognition-based Error ($F_{(1,28)}$=7.67, p=.01) were significant. The analyses of the SOTU-C scores and of the SUS did not show any significant differences between the two groups of participants. Subsequently, in order to assess the presence of possible unforeseen effects, a median split based on the number of errors of the groups was performed but no significant results emerged.

3.3. Discussion

The difference in the result could lead to the conclusion that although P3S is less intuitive during the learning phase, it allows a higher control of the status of the system, as seen during the TA. The user satisfaction result obtained initially, however, was not confirmed. It has been hypothesised that the questionnaire used may not have been sensitive enough to assess the satisfaction using these kinds of interfaces. Moreover, a methodological issue needs to be pointed out: the investigators recorded the errors manually; although this method is widely used, it clearly suffers from human imprecision. The performance speed of both interfaces could have affected the accuracy of the measure.

4. Evaluation of BCI

4.1. Aim

The aim of the second study was to understand whether the subjects interacted with the two different interfaces in different ways. As a consequence of the previous findings, we decided to modify the Java Interfaces. Using an automatic collection of the errors, it was possible to avoid the drawbacks of the observational method. Due to this change, we expected to have a more accurate measure. In this study, we also introduced the cognitive workload as a measure. It was our aim to additionally assess whether gender differences and computer skills could affect performance and the measure on usability and cognitive workload.

4.2. Material and methods

We tested our BCI prototypes on 61 healthy participants (mean age: 22.3, range: 18 to 38) with different computer skills assessed by a questionnaire. Subjects were randomly assigned to one of two groups, one for each prototype. Using the CST, where users were asked to write on the screen through the interface, our participants wrote in one session the same list of ten randomised words used previously. None of the participants was involved in the previous studies. We then administered the SUS to measure usability and the NASA Task Load Index (TLX) [34] to measure cognitive workload.

4.3. Results

We analysed the data using a General Linear Model procedure and a 2x2x3 design, where the independent variables were Interface (PS3 and LSP), Gender, and Computer Skills (low, medium and high). We used the Number of Errors, Choice-based Errors, Recognition-based Errors, and the SUS and NASA-TLX scores as dependent variables.

Gender and Computer Skills did not interact with the measures. On the contrary, the Interface played a major role. Using the P3S, users were less accurate, as shown from the Number of Errors ($F_{(1,60)}$=24.8, p<.00) and Choice-based Errors ($F_{(1,60)}$=30.1, p<.00) made. This is consistent with the finding that users considered the LSP more satisfactory and effortless, as shown by the SUS ($F_{(1,60)}$= 8.7, p=.005) and the NASA-TLX score ($F_{(1,60)}$=4.08, p=.049). The analysis of the NASA-TLX factors did not show any difference, except for the Frustration Level ($F_{(1,60)}$=5.9, p=.019). The analysis of the two factors of SUS confirmed a difference in the SUS Usability score ($F_{(1,60)}$=10.76, p=.002) but not in the SUS Learnability score.

4.4. Discussion

Gender and Computer Skills did not interact with the measures, showing that performance and satisfaction were not influenced by these features. The effects of the interfaces, on the other hand, were significantly evident and they demonstrated how a difference in the interface could change the perspective of the user. Users were more accurate with the LSP interface than with the P3S interface. The LSP supports the users more while writing and requires less mental effort. This result emerges clearly from the difference in the usability scores, which highlight user preference for the LSP. Surprisingly, the findings of the second pilot study revealed that a more accurate measure obtained using an automatic procedure reversed the findings of the first study.

Moreover, our results are particularly interesting considering the findings of Njiober and colleagues [8]. In a study that involved Amyotrophic lateral sclerosis (ALS) patients using a Sensorimotor and a P300 BCI, they recommend the use of the P300, if possible, as first choice because of better immediate results. Although their findings appear to contradict our results, it is important to underline that their differences could be more related to the difficulties controlling the EEG than to the interface itself.

5. Conclusions

Although BCIs are innovative and still progressing technology, the lack of evaluation is still underestimated. The findings of our study highlight the need for adequate tools to assess usability and demonstrate that a change in perspective can affect results on a large scale.

In our first pilot study, we assessed the learnability and the ease of use of two interfaces. The initial data we had with the Thinking Aloud technique showed that the LSP had greater effectiveness in the ease of learning and was therefore preferred by the subjects. The functioning of the LSP did not need much effort to be explained and understood.

The experimental data obtained were partially conflicting. In the pilot study, the P3S appeared to be preferable to the LSP while in the subsequent evaluation study, the opposite was true. Due to the difference in the error-collecting method, we consider the results of the second evaluation to be more reliable. We found that in the use of our interfaces, features such as gender and computer skills did not affect the measures. Using the Language Support Program, users were more accurate and satisfied, and employed less mental effort compared to when they used the P3S.

As highlighted by Federici, technology has a big impact on the properties of the mind. The way an interface works can affect not only variables such as performance speed but also how the information conveyed by the interface is used and how much effort is needed to process such information. Moreover, this process affects satisfaction related to the use of the technology. In fact, following the definition of psychotechnology, a BCI shares and modifies the functions of the mind and participates in the process of configuring and constructing relationships within the user's experience. For this reason, we suggest that future studies about BCIs should involve disabled user and consider more carefully the processes above mentioned.

When designing new technology for disabled people, we should always try to understand how and if this will be used. Disabled users are too often ignored and it is assumed that they will use the technology only because they have no other choice. The fact that disabled people abandon technology clearly shows that the idea that users will use technology because it is needed is wrong. Users want something they can appreciate and be satisfied with, no matter their health conditions.

References

[1] C. Stephanidis, User interfaces for all: New perspectives into human-computer interaction, in: *User Interfaces for All - Concepts, Methods and Tools,* C. Stephanidis, editor. Lawrence Erlbaum Associates, Mahwah, NJ, 2001, 3-17.

[2] D.A. Norman, S.W. Draper, *User Centered System Design; New Perspectives on Human-Computer Interaction,* L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1986.

[3] D. De Kerckhove, *Brainframes: Technology, mind and business,* Bosch & Keuning, Utrecht, NL, 1991.

[4] D. De Kerckhove, *The skin of culture: Investigating the new electronic reality,* Somerville, Toronto, CA, 1995.

[5] E. Pasqualotto, A. Simonetta, S. Federici, M. Olivetti Belardinelli, Usability Evaluation of BCIs, in: *Assistive Technology From Adapted Equipment to Inclusive Environments - AAATE 2009,* P.L. Emiliani, L. Burzagli, A. Como, F. Gabbanini, A.L. Salminen, eds., IOS Press, Amsterdam, 2009, 882.

[6] E. Pasqualotto, A. Simonetta, V. Gnisci, S. Federici, M. Olivetti Belardinelli, *International Journal of Bioelectromagnetism,* in press.

[7] E.W. Sellers, A. Kübler, E. Donchin, Brain-computer interface research at the University of South Florida Cognitive Psychophysiology Laboratory: the P300 Speller, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **14** (2006), 221-224.
[8] F. Nijboer, N. Birbaumer, A. Kübler, The influence of psychological state and motivation on brain-computer interface performance in patients with amyotrophic lateral sclerosis - a longitudinal study, *Frontiers in Neuroscience* **4** (2010),
[9] J. Brooke, SUS: a 'quick and dirty' usability scale, in: *Usability Evaluation in Industry,* P.W. Jordan, B. Thomas, B.A. Weerdmeester, I.L. McClelland, eds., Taylor & Francis, 1996, 189-194.
[10] J.R. Wolpaw, N. Birbaumer, W.J. Heetderks, D.J. McFarland, P.H. Peckham, G. Schalk, E. Donchin, L.A. Quatrano, C.J. Robinson, T.M. Vaughan, Brain-computer interface technology: a review of the first international meeting, *IEEE Transactions on Rehabilitation Engineering* **8** (2000), 164-173.
[11] K. Hannu, P. Pallab, A comparison of concurrent and retrospective verbal protocol analysis, *The American Journal of Psychology* **113** (2000), 387-404.
[12] L.A. Farwell, E. Donchin, Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials, *Electroencephalography and Clinical Neurophysiology* **70** (1988), 510-523.
[13] M. Fabiani, G. Gratton, D. Karis, E. Donchin, Definition, identification, and reliability of measurement of the P300 component of the event-related brain potential, *Advances in Psychophysiology* **2** (1987), 78.
[14] M.J. Scherer, S. Federici, *Assistive Technology Assessment: A Handbook for Professionals in Disability, Rehabilitation and Health Professions,* CRC Press, London, UK, in press.
[15] M.J. Scherer, *Living in the state of stuck: How assistive technology impacts the lives of people with disabilities,* 3rdBrookline Books, Cambridge, MA, 2000.
[16] M.J. Scherer, G. Craddock, T. Mackeogh, The relationship of personal factors and subjective well-being to the use of assistive technology devices, *Disability and Rehabilitation* (2010),
[17] M. Olivetti Belardinelli, *La costruzione della realtà,* Boringhieri, Torino, 1973.
[18] N. Birbaumer, A. Kübler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann, H. Flor, The thought translation device (TTD) for completely paralyzed patients, *IEEE Transactions on Rehabilitation Engineering* **8** (2000), 190-193.
[19] N. Neumann, A. Kübler, J. Kaiser, T. Hinterberger, N. Birbaumer, Conscious perception of brain states: Mental strategies for brain-computer communication, *Neuropsychologia* **41** (2003), 1028-1036.
[20] S. Borsci, S. Federici, M. Lauriola, On the dimensionality of the System Usability Scale: a test of alternative measurement models, *Cognitive Processing* **10** (2009), 193-197.
[21] S. Borsci, S. Federici, The Partial Concurrent Thinking Aloud: A New Usability Evaluation Technique for Blind Users., in: P.L. Emiliani, L. Burzagli, A. Como, F. Gabbanini , A.-L. Salminen , eds., Assistive Technology from Adapted Equipment to Inclusive Environments – AAATE 2009IOS Press, Florence, Italy, 2009, 421-425.
[22] S. Federici, Linee-guida di Psicotecnologie, in: *Graduate Course on Psychotechnology for Learning,* Sapienza University of Rome, Rome, Italy, 2002.
[23] S. Federici, M. Scherer, A. Micangeli, C. Lombardo, M. Olivetti Belardinelli, A cross-cultural analysis of relationships between disability self-evaluation and individual predisposition to use assistive technology, in: *Assistive Technology: Shaping the Future,* G.M. Craddock, L.P. McCormack, R.B. Reilly , H.T.P. Knops , eds., IOS Press, Amsterdam, 2003, 941-946.
[24] S. Federici, A. Micangeli, I. Ruspantini, S. Borgianni, F. Corradi, E. Pasqualotto, M. Olivetti Belardinelli, Checking an integrated model of web accessibility and usability evaluation for disabled people, *Disability and Rehabilitation* **27** (2005), 781-790.
[25] S. Federici, S. Borsci, M.L. Mele, G. Stamerra, Web popularity: an illusory perception of a qualitative order in information, *Universal Access in the Information Society* **9** (2010), 375-386.
[26] S. Federici, F. Stefano, S. Borsci, G. Stamerra, Web usability evaluation with screen reader users: implementation of the partial concurrent thinking aloud technique, *Cognitive Processing* **11** (2010), 263-272.
[27] S. Federici, S. Borsci, Usability evaluation: models, methods, and applications, in: *International Encyclopedia of Rehabilitation,* J.H. Stone, M. Blouin, eds., Available online: http://cirrie.buffalo.edu/encyclopedia/en/article/277/, 2011.
[28] S.G. Mason, M.M. Jackson, G.E. Birch, A general framework for characterizing studies of brain interface technology, *Annals of Biomedical Engineering* **33** (2005), 1653-1670.
[29] T. Hinterberger, A. Kübler, J. Kaiser, N. Neumann, N. Birbaumer, A brain-computer interface (BCI) for the locked-in: comparison of different EEG classifications for the thought translation device, *Clinical Neurophysiology* **114** (2003), 416-425.