BNCI Horizon 2020

FP7-ICT-2013-10 609593
Nov 2013–Apr 2015

Deliverable: **D4.4**
Title: **Quality Assessment Guidelines**

Work package: WP4
Due: M15
Type: ☑ PU[1] ☐ PP[2] ☐ RE[3] ☐ CO[4]

Main author: Ruben Real (UNIWUE)
Other authors: Francesca Schettini (FSL)
Floriana Pichorri (FSL)
Donatella Mattia (FSL)
Begonya Otal (BDIGITAL)

Abstract: The aim of D4.4 is to provide a matrix of the features different BCI stakeholders require with a special focus on practical applicability, broad spectrum of end users, and commercialization.
In addition, it provides a review of methodologies to evaluate and compare BNCI solutions. D4.4 is based on previous deliverables (D2.2, D3.1, D3.2, and D4.1) and several focus group discussions and interviews conducted throughout the project (D4.2).

Keywords: Guidelines, review, user-centred design, UCD

[1] Public
[2] Restricted to other program participants
[3] Restricted to a group specified by the consortium
[4] Confidential, only for members of the consortium

1 Introduction

This deliverable discusses different requirements for future developmental issues in terms of practical applicability in a broad spectrum of end-users and commercialization of BNCI technology (section 2). In addition, a framework of procedures and metrics is described which allows for the evaluation of BNCI applications and technology in a real world setting (section 3).

During the project we identified several unique application scenarios of BNCI technology (see D4.1: improve, replace, enhance, supplement, restore, & research) and three unique classes of users (see D4.1; Table 1). For each application scenario the consortium identified three promising use cases (see D1.3). However, an extensive revision of these use cases revealed that the supplement scenario was too unrealistic at the moment. Indeed, many use cases originally assigned to this scenario are more reflective of the replace/restore scenarios. Thus we decided to focus only on five scenarios (improve, replace, enhance, restore, & research). To reach the aims of this deliverable we conducted focus group discussions and expert interviews with users of BNCI technologies. Participating users were presented with a selection of six different use cases and their opinions elicited. A summary of the main results is presented in section 2 of this deliverable (see also D4.2).

While this approach provides insights into precisely defined topics and issues raised by BNCI users, it does not present an evaluative framework able to provide guidelines on how to ensure that BNCI systems may meet these diverse requirements. Based on the consortium's experiences in the user-centred design approach, we developed a framework to help evaluating BNCI technologies and applications (see section 3).

1.1 Terms and definitions

The following terms and definitions will be used throughout the document.

Table 1 – Terms and Definitions.

Term	**Definition**
Accessibility	The quality of a technology to be available to all members of society.
Differential Indication	Concerns choosing the optimal technology for a given user, with respect to the individual's aims and goals.
Technological limitations	Aspects of current BNCIs which limit its usability.
End users, or primary users	Persons who actually use the product.
Secondary users	Persons who will occasionally use the product or those who use it through an intermediary.
Tertiary users (professional users or other stakeholders)	Persons who will be affected by the use of the product or make decisions about its purchase.

2 BNCIs - requirements for success

2.1 Topics

Based on the Focus Group / Interview summaries provided by the project partners for D4.2 a total of 63 statements relating to BNCI requirements were extracted. Across the different use cases 25% of all statements related to Societal and Ethical Implications, 19% referred to

questions of the Differential Indication of BNCI technology, 14% concerned BNCIs strong interdisciplinarity and 11% referred to current technological limitations. Other important aspects included Usability (10%), Financing (8%), and Flexibility (5%). The distribution of statements showed little differentiation between user classes, indicating that primary, secondary, and tertiary users shared a common understanding of the most important concerns. Thus, in the following, we summarize the most important topics across use cases and user classes with the understanding that these summaries apply to all user classes and use cases, with exceptions highlighted.

2.1.1 Ethical and Social Issues

Participants highlighted several aspects, which provide a challenge for the individual primary users but also for societal and political stakeholders. While the majority of topics (56 %) in this category, such as the need to avoid raising false hope, or **risk-benefit ratio** has been discussed before (see Haselager, Vlek, Hill & Nijboer, 2009; Vlek et al., 2012), other factors require separate analysis. For example, several participants were concerned about issues of **accessibility** of enhancing technologies, or issues relating to **privacy** of brain data, or whether our concept of **personhood** could change if interaction with computers becomes more intuitive using a BNCI.

2.1.2 Differential Indication

Participants highlighted the need to precisely know the **indications** of BNCIs (which patients, training duration and quantity, aims of the patient), before they would be able to prescribe such a device. Consequently, BNCIs should have high **flexibility** to allow for easy customization for each individual end user (see below "flexibility").

2.1.3 Interdisciplinarity

BNCI is an inherently interdisciplinary field requiring close cooperation of various stakeholders, and this is highlighted in the need for in-depth knowledge of participating fields. Thus, several participants called for more **basic research** into the processes targeted by BNCIs, before this technology could be widely adopted.

2.1.4 Technological limitations

Participants highlighted the fact that many potential applications of BNCIs are limited by the available **signal acquisition** and **signal analysis** technologies (see D2.2).

2.1.5 Usability

Several participants highlighted the need of BNCIs to be unobtrusive, easy-to-use, to be user-friendly, and to demonstrably help the end user to achieve his/her goals. In this regard it should also be mentioned that aesthetic design maybe of great importance to some primary users (c.f. Zickler, Halder, Kleih, Herbert, & Kübler, 2013; Holz, Höhne, Staiger-Sälzer, Tangermann, & Kübler, 2013).

2.1.6 Financing

Participants agreed that the development of practical BNCIs requires great investments, and highlighted that **simple yet flexible** systems might be preferable to expensive but feature-complete off-the-shelf systems. Depending on the intended field of application of a BNCI, strong competition from other technologies might even increase the need for cheap BNCI systems.

2.1.7 Flexibility

Flexibility was an important topic throughout discussions. With BNCI technology leaving the lab and being more widely adopted in other fields, tensions arise between demands for highly flexible system on the one hand, and the industry's interest in developing feature-complete but proprietary off-the-shelf systems. In this situation, focusing on **interoperability** of BNCI components through the use of open **standards** might be a key factor to help the BNCI sector reach faster and persistent growth (see D3.1). Maintaining interoperability would meet

primary, secondary and tertiary users' interest in Differential Indication, Interdisciplinarity, Financing, and Usability (see above).

2.2 Requirements

Table 2 – General requirements for BNCI success.

Topic	**Requirements**
Technology	• Enhanced signal acquisition and analysis techniques • Efficacy • Reliability • Robustness
Interdisciplinarity	• Paradigm development • Interoperability • Flexibility
Finance	• Cost reduction • Market studies • Demonstrate relative benefits over competing technologies • Safety
Usability	• Enhance autonomy • User-friendliness • Aesthetic design

Table 3 – Specific requirements for BNCI success per scenario.

Scenario	**Requirements**
Replace (invasive)	• Signal acquisition and analysis • Robustness • Aesthetic design
Replace (non-invasive)	• Research into control signals • Reliability • Usability
Restore	• User-friendliness • Reliability • Robustness • Proof-of-principle • Relative benefits
Improve	• Studies on differential indication • Flexible systems • Interoperability of diverse soft- and hardware components
Enhance	• Basic research into neuropsychological processes • Market studies
Research	• Reliability • Robustness • Signal acquisition and analysis

While the above analysis provided insights into concrete topics and issues raised by BNCI users, it did not present an evaluative framework able to provide guidelines on how to ensure that BNCI systems may meet these diverse requirements. In the following, we develop such a framework.

3 An evaluation framework for current and future BNCIs

Although BNCI systems are ultimately designed for what might be called "independent home use", much research still takes place within the comforts of a well-equipped psychophysiological laboratory. Thus, there exists a translation gap, which manifests itself in a lack of knowledge about end-users of BNCI technology and the biological, psychological and social aspects of human-computer interaction (Kübler et al., 2014).

In the following, we describe the application of the user-centred design (UCD) process as a viable approach to bridge this gap and bring BNCI technology to the market. This framework is derived from the consortium's long standing experience and mirrored in several recent publications (Kübler et al., 2014; Schettini et al., 2015; Riccio et al., 2011; Zickler et al., 2011). It is based on the user-centred design approach, which posits "early and continuous involvement of potential users, understanding of user requirements and the whole user experience, and iterative processes between developers and users" (Kübler et al., 2014, p. 2). These principles can be implemented using a four-stage development mode (see Table 4), which focuses on understanding and specifying the user's needs, defining the context of use, evaluating prototypes against these specification, and developing ever-more refined prototypes to meet these requirements (see Figure 1).

These principles and stages derive from the concept of **usability**, which ISO standard 9241–210 defines as the ''extent to which a […] product […] can be used by specified users to achieve specified goals with *effectiveness*, *efficiency* and *satisfaction* in a specified *context of use*'' (p. 3). Whereas *effectiveness* refers to the accuracy with which a user can accomplish a given task, *efficiency* relates effectiveness to invested costs (time and personal efforts). Earlier conceptualizations of *user satisfaction* defined this to mean the perceived comfort and acceptability while using a BNCI product, e.g. a BCI controlled application (Kübler et al. 2014). Here, its focus is broadened to also include satisfaction with using BNCI technology components. Finally, *context of use* refers to users, tasks, equipment (e.g. hard- and software, materials), and the physical and social environments in which a product/technology is used.

In the UCD process, participants should be selected from the **intended user population**, even if this may mean spending substantial efforts in recruiting these participants (e.g. involvement of motor impaired individuals). In addition, prototype evaluation always refers to evaluating a product based on **actual experience**. Asking participants to imagine how interacting with a (fictional) product would be like does not suffice and may even be impossible for end-users. Finally, tasks selected for evaluation need to be representative of actual product use, as restricting evaluation to subsets of tasks may severely limit generalizability beyond the sampled tasks.

Table 4 – Principles of user-centred design and their application for BNCI technology (from Kübler et al., 2014).

Principle	**Application**
Understand the user, the task and environmental requirements	Chose appropriate metrics - apply interviews/questionnaires for first definitions
Encourage early and active involvement of users	Interaction between users and developers to define the first version of a prototype
Be driven and refined by user-centred evaluation	Valid evaluation metrics
Include iteration of design solutions	Continuous interaction between developers and end-users in their home environment leading to several prototypes
Address the whole user experience	Evaluation metrics that covers all aspects of ‘‘usability’’, i.e. effectiveness, efficiency, satisfaction
Encourage multi-disciplinary design	Continuous involvement of experts of relevant fields

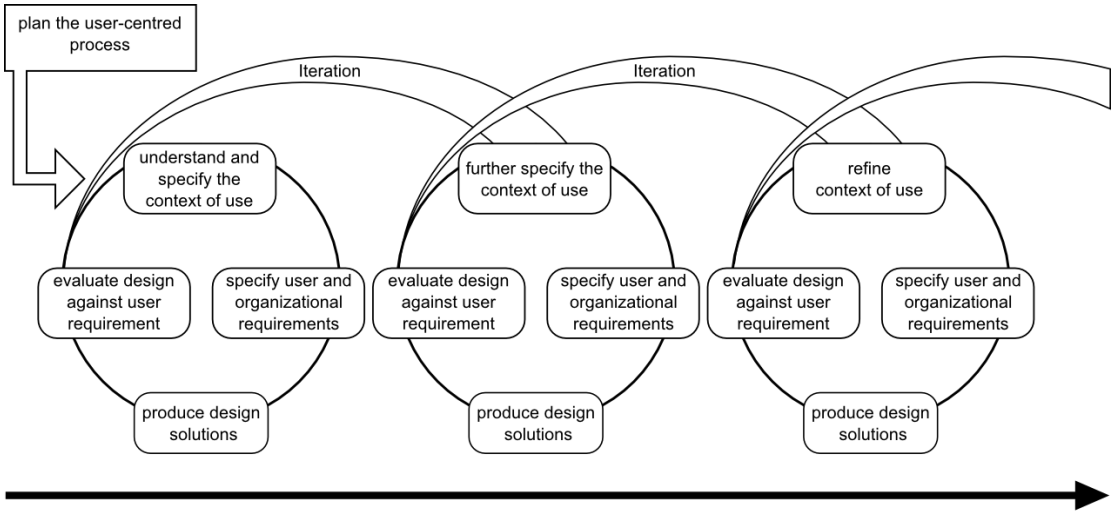

Figure 1 – the User-Centred-Design Process.

3.1 Implementing the UCD process for BNCIs

An important aspect in the UCD approach is the definition of **valid evaluation metrics**. Generally, these metrics should be as reliable as possible, but care should be taken not to sacrifice external validity. In addition, perceived performance of a BNCI-application might strongly depend on the task and software ecosystem (see D2.2) so application and user specific information can be gathered even using simple face valid measures. Following the definition of usability, the next section present possible metrics for effectiveness, efficiency, and satisfaction (see Table 5).

Table 5 – Evaluation metrics (from Kübler et al., 2014)

Aspect of usability	**Application to BNCIs**	**Example metrics**
Effectiveness	Accuracy	% correct response
Efficiency	Information transfer rate	bits/min
	Utility metrics	correct responses per unit of time
	Workload questionnaire	NASA-TLX[1]
Satisfaction	perceived reliability, learnability, speed, aesthetic design	single item measures
	Match between product and user	ATD-PA[2]

[1]: Hart & Staveland (1988). [2]: Corradi, Scherer, & Lo Presti (2012)

3.1.1 Effectiveness

Effectiveness refers to measures of how accurate and complete users can accomplish a given task using a BNCI, i.e. how often the intended output can be achieved. Thus, accuracy, as a measure of effectiveness, can be calculated by relating the number of successful selections to the total number of attempted selections.

3.1.2 Efficiency

As highlighted in D2.2, measures of effectiveness do not address the frequent need to balance the trade-off between accuracy and speed. Therefore, measures of efficiency relate the costs, i.e. effort and time, invested by the user to effectiveness. An **objective measure of efficiency** is the information transfer rate (ITR) and its modifications with regards to error probability, accuracy, and practicality. However, even systems showing a high information transfer rate can be impractical to use if the number of errors is high. Thus, more global measures, such as **utility metrics** (e.g. number of correctly spelled letters per unit of time) have emerged. In addition, **subjective measures of efficiency**, e.g. workload (the perceived "costs incurred by a human operator to achieve a particular level of performance", Hart & Staveland, 1988, p. 140) should be used.

3.1.3 Satisfaction

User satisfaction is defined with reference to the perceived comfort and acceptability while using a product. Depending on the context of use, different metrics, e.g. referring to aspects of a device, or face-valid questions on overall satisfaction may be used. However, the ultimate proof of user satisfaction may lie in its **actual daily use**. Unfortunately, few institutions have enough equipment available for extended home use so this requirement often remains unmet.

3.2 Summary

The UCD perspective in the development of computer-based interactive systems provides a theoretical framework which can guide the design of mandatory translational studies on how to transfer BNCI-controlled applications from the laboratories of developers to the end-users' homes. Appropriate measures for evaluation of usability are now available and have been shown to be applicable even in severely paralyzed and locked-in potential BNCI end-users (Kübler et al., 2014).

4 Discussion

D4.4 provides an overview of the requirements of different BNCI stakeholders concerning applicability, broad spectrum of end users, and commercialization. In addition, we provide a methodology to evaluate and compare BNCI solutions.
Analysis of focus group discussion and interviews indicated a broad spectrum of topics relevant for BNCI's success (see section 2.1). On the most abstract level (see Table 2), current challenges include the improvement of the technology proper, fostering interdisciplinarity between providers and potential (tertiary) users of BNCI technology, assessing the marketability of BNCI related products, and improving their usability (see also section 3).

This deliverable focused on user requirements relating to the use of the technology and/or applications proper. Nevertheless, users were also highly conscious of the financial aspects of BNCI technology. Thus, we would like to refer the reader to D3.3, which is devoted to aspects of transfer and exploitation of technology. A comprehensive statement linking technological requirements and aspects of commercialization will be delivered in D4.5 (Contribution to the roadmap).

5 References

Corradi, F., Scherer, M. J., & Lo Presti, A. (2012). Measuring the Assistive Technology Match. In S. Federici & M. J. Scherer (Eds.), *Assistive Technology Assessment Handbook* (pp. 49–65). Boca Raton: CRC Press.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139–183). Amsterdam: North Holland Press.

Haselager, P., Vlek, R., Hill, J., & Nijboer, F. (2009). A note on ethical aspects of BCI. *Neural Networks*, 22(9), 1352–1357. doi:10.1016/j.neunet.2009.06.046

Holz, E. M., Höhne, J., Staiger-Sälzer, P., Tangermann, M., & Kübler, A. (2013). Brain-computer interface controlled gaming: Evaluation of usability by severely motor restricted end-users. *Artificial Intelligence in Medicine, 59*(2), 111–120. doi:10.1016/j.artmed.2013.08.001

ISO 9241-210. (2008). *Ergonomics of human system interaction: Part 210* Human centred design for interactive systems (formerly known as 13407). Switzerland: International Organization for Standardization (ISO).

Kübler, A., Holz, E. M., Riccio, A., Zickler, C., Kaufmann, T., Kleih, S. C., . . . Federici, S. (2014). The User-Centered Design as Novel Perspective for Evaluating the Usability of BCI-Controlled Applications. *PLoS ONE, 9*(12), e112392. doi:10.1371/journal.pone.0112392

Riccio, A., Leotta, F., Bianchi, L., Aloise, F., Zickler, C., Hoogerwerf, E.-J., . . . Cincotti, F. (2011). Workload measurement in a communication application operated through a P300-based brain-computer interface. *Journal of Neural Engineering, 8*(2), 25028. doi:10.1088/1741-2560/8/2/025028

Schettini, F., Riccio, A., Simione, L., Liberati, G., Caruso, M., Frasca, V., Calabrese, B., Mecella, M., Pizzimenti, A., Inghilleri, M., Mattia, D., Cincotti, F. (2015). An assistive device with conventional, alternative and Brain-computer interface inputs to enhance interaction with the environment for people with Amyotrophic Lateral Sclerosis: a feasibility and usability study. *Archives of Physical Medicine and Rehabilitation, in press*

Vlek, R. J., Steines, D., Szibbo, D., Kübler, A., Schneider, M. J., Haselager, P., & Nijboer, F. (2012). Ethical issues in brain-computer interface research, development, and dissemination. *Journal of Neurologic Physical Therapy, 36*(2):94-99.

Zickler, C., Halder, S., Kleih, S. C., Herbert, C., & Kübler, A. (2013). Brain Painting: usability testing according to the user-centered design in end users with severe motor paralysis. *Artificial Intelligence in Medicine*, *59*(2), 99–110. doi:10.1016/j.artmed.2013.08.003

Zickler, C., Riccio, A., Leotta, F., Hillian-Tress, S., Halder, S., Holz, E., . . . Kübler, A. (2011). A brain-computer interface as input channel for a standard assistive technology software. *Clinical EEG and Neuroscience*, *42*(4), 236–244. doi:10.1177/155005941104200409