User-centered design in brain–computer interfaces—A case study

Martijn Schreuder [a,*,1], Angela Riccio [b,c,1], Monica Risetti [b], Sven Dähne [a],
Andrew Ramsay [d], John Williamson [d], Donatella Mattia [b], Michael Tangermann [e]

[a] Machine Learning Lab, Berlin Institute of Technology, Marchsstraße 23, 10587 Berlin, Germany
[b] Neuroelectrical Imaging and Brain–Computer Interface Lab, Fondazione Santa Lucia, Via Ardeatina 306, 00142 Rome, Italy
[c] Department of Psychology, University of Rome Sapienza', Via dei Marsi 78, 00183 Rome, Italy
[d] School of Computing Science, University of Glasgow, G12 8QQ Glasgow, Scotland, United Kingdom
[e] BrainLinks-BrainTools Excellence Cluster, University of Freiburg, Albertstraße 23, 79104 Freiburg, Germany

ARTICLE INFO

ABSTRACT

Objective: The array of available brain–computer interface (BCI) paradigms has continued to grow, and so has the corresponding set of machine learning methods which are at the core of BCI systems. The latter have evolved to provide more robust data analysis solutions, and as a consequence the proportion of healthy BCI users who can use a BCI successfully is growing. With this development the chances have increased that the needs and abilities of specific patients, the end-users, can be covered by an existing BCI approach. However, most end-users who have experienced the use of a BCI system at all have encountered a single paradigm only. This paradigm is typically the one that is being tested in the study that the end-user happens to be enrolled in, along with other end-users. Though this corresponds to the preferred study arrangement for basic research, it does not ensure that the end-user experiences a working BCI. In this study, a different approach was taken; that of a user-centered design. It is the prevailing process in traditional assistive technology. Given an individual user with a particular clinical profile, several available BCI approaches are tested and – if necessary – adapted to him/her until a suitable BCI system is found.

Methods: Described is the case of a 48-year-old woman who suffered from an ischemic brain stem stroke, leading to a severe motor- and communication deficit. She was enrolled in studies with two different BCI systems before a suitable system was found. The first was an auditory event-related potential (ERP) paradigm and the second a visual ERP paradigm, both of which are established in literature.

Results: The auditory paradigm did not work successfully, despite favorable preconditions. The visual paradigm worked flawlessly, as found over several sessions. This discrepancy in performance can possibly be explained by the user's clinical deficit in several key neuropsychological indicators, such as attention and working memory. While the auditory paradigm relies on both categories, the visual paradigm could be used with lower cognitive workload. Besides attention and working memory, several other neurophysiological and -psychological indicators – and the role they play in the BCIs at hand – are discussed.

Conclusion: The user's performance on the first BCI paradigm would typically have excluded her from further ERP-based BCI studies. However, this study clearly shows that, with the numerous paradigms now at our disposal, the pursuit for a functioning BCI system should not be stopped after an initial failed attempt.

1. Introduction

Brain–computer interfaces (BCI) hold the promise of allowing those with (near-)complete paralysis another chance at communication or environmental control. Such paralysis is called locked-in syndrome (LIS) if all but ocular-motor functions are compromised. If this last voluntary control is also lost, the condition is called total- or completely locked-in syndrome [CLIS, 1]. For both conditions, a BCI may be the only form of independent expression. But even for patients suffering from incomplete locked-in syndrome, those

remaining residual motor functions may fatigue quickly, and they may thus be augmented by BCI.

So far, BCIs have mainly been targeted at end-users suffering from late stage amyotrophic lateral sclerosis (ALS) [2–5], a neuro-degenerative disease with short life-time expectancy after diagnosis. The median survival time after tracheostomy, the time were a BCI becomes relevant, has been reported at 21 months [6]. For ALS, the course of progression varies widely between patients, but the symptoms are rather similar. In the final state, the patient loses all reliable, voluntary muscle control [7]. It was long assumed that cognitive functions remain largely untouched, however, several recent findings have challenged this concept [8–10]. They state that a range of cognitive changes are associated with ALS, although this is difficult to assess systematically for patients in the locked-in state. Nevertheless, the symptoms in ALS largely overlap.

Interestingly, the largest cause of LIS is brainstem damage, such as a brainstem stroke or physical injury as in traumatic brain injury (TBI, [11]). There is a large population of people suffering from the consequences of such strokes or TBIs for which a BCI may be desirable [12]. The yearly mortality rate for TBI patients that survive the first six post-trauma months is estimated to be only 1.33 times higher than that of the general population [13]. For first-time stroke patients that survive the first 30 post-incident days the mortality rate is estimated at about 2.3 times that of the general population [14]. It is highest for immobile patients, due to an increase in secondary causes of death such as circulatory problems; still the life expectancy is on the order of decades once the chronic phase has been entered. In the light of this it seems odd that end-users with locked-in syndrome due to stroke or TBI have thus far played only a minor role in BCI research.

TBIs and strokes manifest themselves in very diverse ways; the level of functional and cognitive impairment depends on the locus of the trauma and the extend of the damage. Symptoms can be completely different between (1) an isolated brain-stem stroke, which mostly impairs motor-functions and can lead to vigilance and awareness deficits, and (2) a diffuse stroke, where a wide area of cognitive functions are compromised. With such a heterogeneous target group, it can be expected that a BCI may be of highest practical use for the subset that has intact cognitive abilities. The applied complexity of the BCI control and the chosen interaction paradigm should largely be determined by the extent of the loss of cognitive abilities. For example, a simple binary BCI like in [15] may – at first glance – be slower than multi-class approaches such as described in [16], but the added complexity of the second approach may require a greater mental effort. The same consideration should be made for the modality used, be it visual, tactile or auditory. For most TBI and stroke patients, a certain set of neurophysiological and -psychological tests is part of their routine post-trauma assessment. They typically test for attention-, memory- and other cognitive abilities. The results of such tests could be a viable starting-point for finding a BCI system that matches the user's abilities.

Practically, this pursuit of an appropriate BCI can be realized in a user-centered design process, which is formally described in the ISO 9241-210. In short, it states that the software should be designed with deep understanding of the end-user, that end-users should be involved in every step of the design and that the design process should have several iterations where end-user feedback is incorporated. Furthermore, any testing should be done with the potential end-user. Taking the prior knowledge of neurophysiological and -psychological tests of an end-user into consideration, such deep understanding can be gained and quantified. Using this, one or more BCI paradigms and modalities are then screened for their applicability with the particular end-user. During the screening of each such paradigm, an iterative process is followed to adapt the settings of the BCI system to the end-user's needs and abilities.

The present case-study reports on a single female end-user (FD) with severe motor- and communication deficits after a brain stem stroke. She participated in two different BCI studies. Even though both were based on event-related potentials (ERP) measured by electroencephalography (EEG), the BCI performances were substantially different. To investigate potential reasons for this difference, FD's existing history of repeated neurophysiological and -psychological test results was re-examined.

2. End-user profile

2.1. Case report

The end-user (FD) is a 48-year old Italian woman. At age 44 she suffered from an ischemic stroke in the area of the basilar artery, after which she showed a clinical picture characterized by tetraplegia and severe dysarthria. Her impairment lead to a severe lack of communication. A magnetic resonance imaging (MRI) scan, acquired 10 days after the ischemic event, showed an altered signal intensity in the infero-posterior area of the left cerebellar hemisphere, in the upper area of the right cerebellar hemisphere, in the cerebellar vermis, and in the midbrain with greater extension on the left. It also showed a small hemorrhagic rift within the left cerebellar hemisphere lesion and in the central pontine area. Twenty days after the ischemic stroke, FD was alert and able to localize sound stimuli by turning her eyes towards the sound source and reasonable changes in facial expression were present. Her motor disability was characterized by motor tetraplegia with hypotonia and symmetrical generalized hyperreflexia. She was thus diagnosed with the locked-in syndrome. Immediately after the diagnosis, a binary model of communication exploiting eye gaze was set up. FD was trained to communicate by focusing her gaze on an alphanumeric communication board, which she still uses to date. Before the event, FD worked in the field of graphic arts and played drums in a band; she was confident with using computers and other technology. At the time of testing, FD had the ability to perform inaccurate movements with the right arm and the head, had preserved facial expressions and precise eye movements. The communication of primary needs was only possible with the support of her communication board on which she pointed to letters. In addition, she could acknowledge requests by a button press.

FD was curious and motivated to test BCI. She initially joined for testing the auditory BCI prototype AMUSE (**a**uditory **mu**lti-class **s**patial **E**RP, see Section 3.2.1) when she was 46 years old and, one year later, she joined the testing of a vision-based BCI prototype (Photobrowser see Section 3.2.2). Her motor deficit did not change between the studies. Around the time of the testing of each prototype, FD underwent assessment of her cognitive functions. The results of both evaluations are described in Section 4.3.

2.2. Neuropsychological assessment

In the context of a standard clinical diagnosis process, FD was subjected to a neuropsychological assessment of general cognitive impairment, attention, memory, working memory and executive functions twice (see Fig. 1). The first assessment took place about 4 months before the first AMUSE trial, the second assessment around 6 months after the last Photobrowser trials. The test battery was administered in a quiet room and over several sessions to prevent fatigue. During both assessments, FD was motivated and cooperative. Her general cognitive level was tested by means of the mini mental state examination (MMSE; [17]). The two subtests of the scale that were not applicable due to FD's physical condition (spontaneous writing subtest and constructive praxis ability
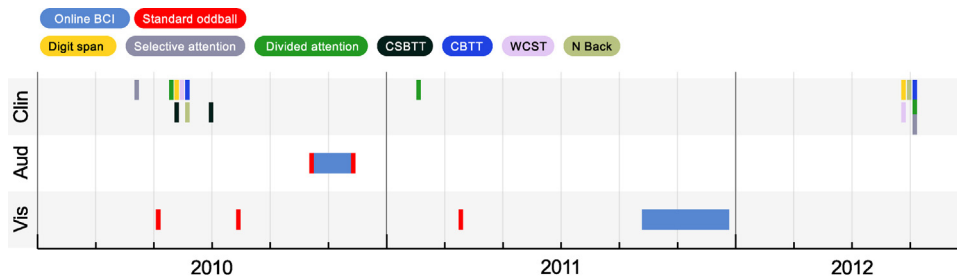
Fig. 1. Timeline for the neurophysiological and -psychological assesments ('Clin') and auditory ('Aud') and visual ('Vis') EEG sessions. Two clusters of assessments are seen, one before the first AMUSE trials and one after the last Photobrowser trials. Stroke incident of FD occurred late 2008. WCST: Wisconsin card sorting test, CBTT: Corsi's block-tapping test, CSBTT: Corsi's supraspan block-tapping test.

subtest) were not presented to her. Missing tests were scored with the maximum score. The attention capacities were evaluated using tests of divided attention and selective attention, taken from a computerized test battery [18]. In the latter tests, FD was asked to provide an answer by pressing a key. The forward and backward digit-span tasks [40] were utilized to assess verbal short-term memory and working-memory, respectively, and the Corsi's block-tapping test and supraspan block-tapping test [19,20] were used to assess visuo-spatial working memory. FD was alert and able to pay attention even for prolonged periods of time. Also administered were the Wisconsin card sorting test [21,22] to assess executive functions, and the N-back test [18] to assess the control of information flow and the updating of information in working memory.

The computerized tests provide normative values for reaction times as well as for accuracies. However, due to the slowness of motor response (FD had residual, but very weak movements of the right index finger), only scores of accuracy were considered in drawing the neuropsychological profile. As FD was unable to give a verbal response, the backward and forward digit span tests were administered by asking her to point to a series of numbers on her communication board.

2.3. Neurophysiological profile

Four types of neurophysiological screening methods were carried out, eliciting cortical evoked potentials: visual evoked potentials (VEP), brainstem auditory evoked responses (BAER), and ERPs from auditory and visual stimulation. VEP were recorded in order to detect possible lesions or inflammations of the optic nerve, and BAER were recorded for the assessment of the brainstem function and for the evaluation of hearing ability. Both were within the normal range. In order to screen the cognitive ERPs, FD was presented twice with a slow, two-class attended auditory and visual oddball paradigm (see Section 3). Fig. 2A shows the averaged temporal responses for the EEG channel Cz for both modalities, target stimuli (deviants) and non-target stimuli (standards). Cz is taken to visualize both early and late components of both modalities. Shown is data from the first and the last available recording. The scalp maps in Fig. 2B give the averaged spatial patterns for the intervals indicated in Fig. 2A. The ERPs from the two recording sessions of each modality do not differ substantially. For both modalities, typical early negative (N1/N2) and late positive (P3b) components could be evoked. Both components were class-discriminative between target and non-target stimuli, which is a prerequisite for the use by discriminative machine learning methods in a BCI paradigm. A particularly strong positive deflection (P2) was found for the auditory modality. Though the visually evoked ERPs generally have higher class-discriminability, the class-discriminability of the auditory ERPs for FD was at least as high as for the visual ERPs.

3. Methods

Apart from the aforementioned neurophysiological assessments, FD took part in two separate online BCI studies (AMUSE and Photobrowser, see below). For a timeline of all data see Fig. 1.

3.1. Standard auditory and visual oddball screening

The auditory oddball consisted of a two-tone discrimination task (deviant: 1000 Hz, 80% probability. standard: 500 Hz, 20% probability). A total of 400 tones were recorded, with a stimulus onset asynchrony (SOA) of 1000 ms. The user was instructed to focus on and count the high pitched deviants. Deviant and standard tones came from the same front-left speaker.

The visual oddball consisted of a set of symbols (X and O) that were serially presented in the middle of a computer screen (white on black), with an SOA of 2100 ms. The subject was told to focus on one particular symbol (X) and count its occurrence, whilst ignoring the other symbol (O). Target probability was the same as for the auditory oddball (20% deviant, 80% standard). A total of 500 stimuli were recorded. Results are presented under the end-user profile (see Fig. 2).

3.2. Description of BCI paradigms

As the user had good discriminative information in both the visual and the auditory standard oddball task, two different paradigms were tested. First, an auditory ERP paradigm based on spatially located sound sources [23,16] was used to drive a text entry system. Second, a visual ERP paradigm was investigated. Contrary to the classical speller application [24], a photo application was controlled, and a novel set of optimized visual stimuli was exploited [25].

Both paradigms are based on external stimulation in the respective modality which elicit ERPs. By focusing on a particular target stimulus, the EEG response to that stimulus differs from responses to non-attended stimuli (non-targets). The stimuli were presented with an SOA of 300 and 220 ms, respectively (see Table 2 for stimulus details). Compared to the standard oddball tasks, this shorter SOA is typical for BCI applications, as it increases the communication bandwidth.

As evoked responses of the EEG are not visible with the bare eye, data driven filtering algorithms are used to enhance the signal-to-noise ratio (SNR) and machine learning methods for classification of EEG features make it possible to estimate the attended vs. unattended stimulus on the basis of a single epoch [26,27]. Using these methods, the EEG signals are analyzed for the presence of target ERPs in real-time. As the brain signals differ between individuals, the analysis is tailored to each BCI user, and the classifier is calibrated on his/her individual data. In order to take a reliable decision about the intended target, several stimulation repetitions
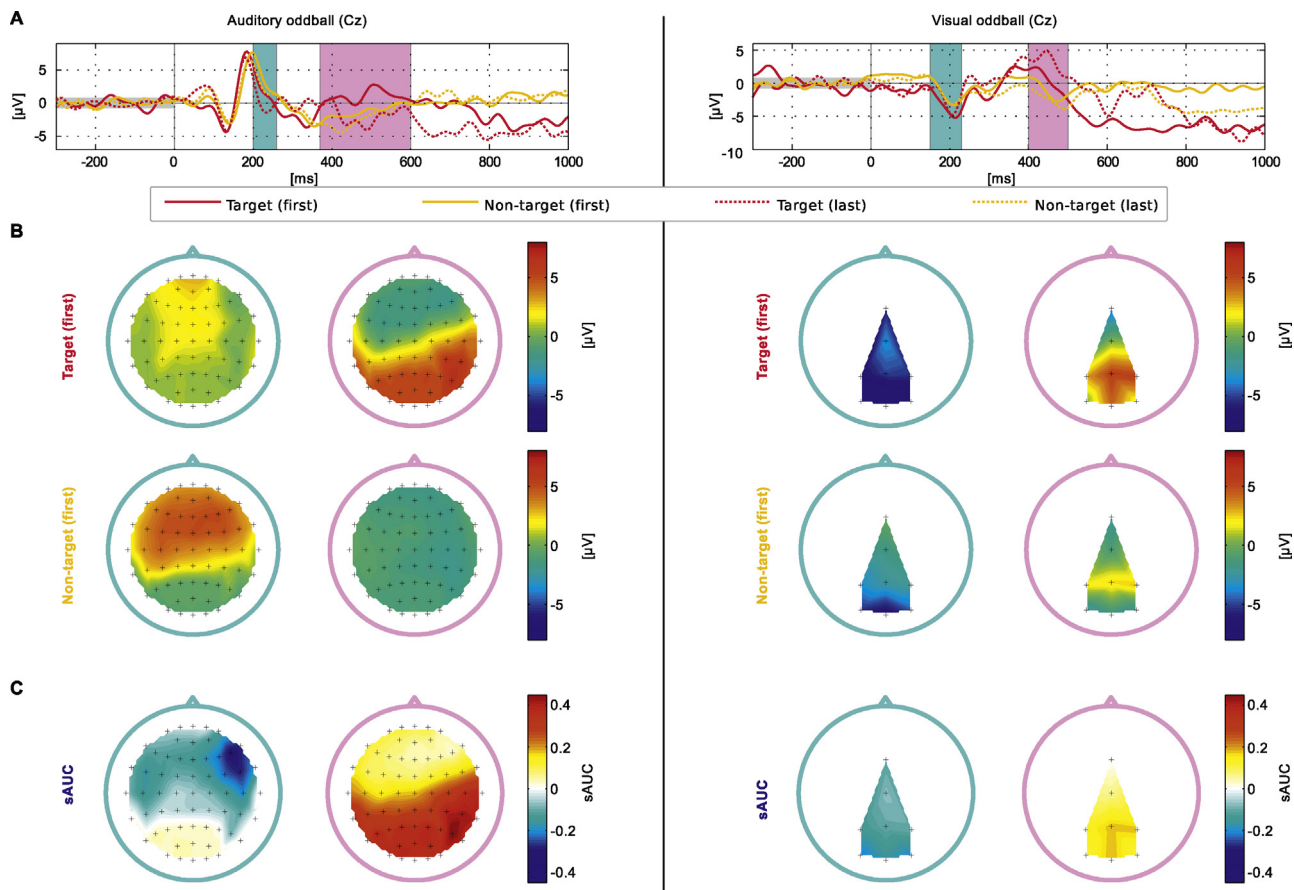
Fig. 2. Standard oddball results. The left panel is for the auditory oddball, whereas the right panel displays the visual oddball. (A) For each modality, two ERP timeseries are plotted per class for channel Cz, displaying the first and last recorded standard oddball run (see Fig. 1). They clearly show a similar pattern, thus indicating only marginal changes in FD's neurophysiology. (B) For both intervals marked in (A), the spatial distribution is plotted for both classes. (C) The discriminability between targets and non-targets is given in terms of sAUC, for both intervals marked in (A). Areas with class-discriminative information are rather typical for healthy users. (B) and (C) are based on data from the first recording.

(e.g. 3–15) can be presented, and the collection of the resulting classifier outputs leads to an accumulation of evidence. This is the basic principle that drives every ERP-based BCI.

3.2.1. AMUSE – an auditory paradigm

AMUSE is an auditory BCI paradigm that uses spatial location as a discriminative cue [23]. As spatial hearing is an innate human ability, its exploitation allows an auditory BCI to provide more than the usual two classes. An increased number of classes allow the paradigm to transfer more information per selection and in principle any application can then be controlled more quickly. In practice however, the increased number of classes has to be balanced against a lower SNR compared to the usual pair of classes. In AMUSE, a multi-class setup is realized in free-field, with one loudspeaker for each direction. Loudspeakers were placed around the subject in a circle at about 65 cm from the center of the head (see Fig. 3). Every loudspeaker presents one unique tone only, which leads to a double-cued paradigm: pitch and direction both code for a stimulus class.

The AMUSE principle was applied to drive a speller with six classes, e.g. six spatial locations, with healthy subjects in [16]. As the minimum number of classes in a spelling task is the number of letters in the alphabet, spelling was realized in a two step process. First, the user selects a group of letters by focusing on the corresponding direction. In the second step, the selected group of letters is re-distributed over the locations, such that a letter

can be selected. Apart from the differences mentioned above, this paradigm was equivalent to that described in [16], where further details about the AMUSE paradigm and its application in a spelling scenario are given.

The AMUSE paradigm was designed exclusively with the end-user in mind. For end-users in the (C)LIS state it is difficult, if not impossible, to control eye-gaze. Any auditory BCI alleviates this need. AMUSE was thoroughly tested with healthy users for feasibility and robustness [23,16]. Several aspects, however, remained to be tuned for end-user application, which was the goal of the study that FD participated in. The speaker distribution was no longer equidistant, to compensate for FD's wheelchair mounted head-rest and to reduce the front-back errors due to the cone of confusion [28]. Stimulus loudness and SOA were adapted to the user's comfort. The amount of recorded calibration data was reduced, to account for the reduced operation time. Insights gained in this study will go into the next iteration of development.

During a first preparation session, informed consent was given by FD and she was introduced to the paradigm. Furthermore, this first session was used to record an initial amount of calibration data, which was utilized to derive system parameters for following sessions. Feedback was not given during this first session. In the five following sessions, calibration data was recorded at the beginning, before a classifier was trained on that data, which could be used for online runs. The online runs comprised copy-spelling tasks of text entry.
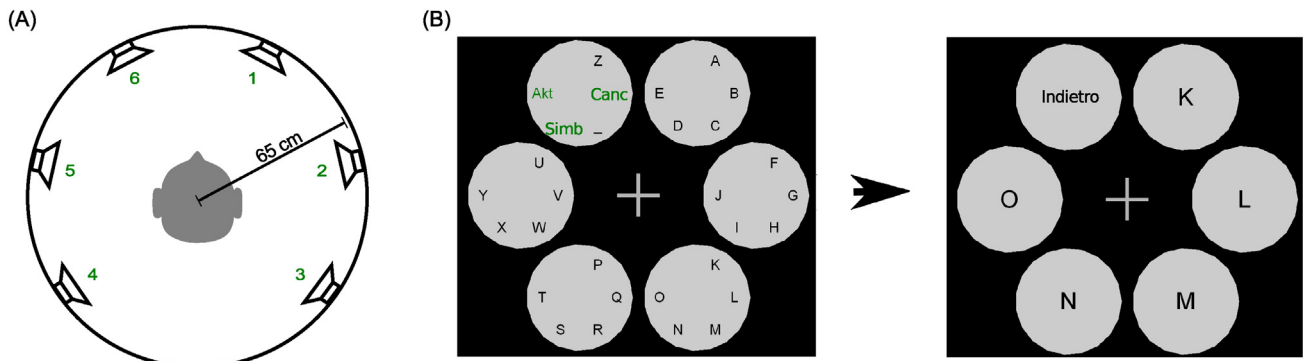
Fig. 3. (A) top view of the AMUSE setup. The user is surrounded by 6 speakers, all of which are activated with a unique tone and in random order. (B) AMUSE was used to drive a two-step spelling interface. The left panel shows the first step, in which a subset of the alphabet can be selected. That subset is reassigned to the speakers in the second step, and an individual letter can be selected.

3.2.2. Photobrowser – a visual paradigm

The most widely investigated and most successful BCI application to date is the visual matrix speller, first introduced by [24]. In this classic spelling application, the characters are displayed in a grid (see Fig. 4A) on the screen. Rows and columns are (pseudo) randomly highlighted by simply increasing their brightness for a short time (though more complex patterns have been successfully investigated, [5]). The user's task is to attend to the symbol he wants to select and count the number of times the attended symbol is highlighted. Although it is also known as the P300 speller, which refers to the cognitive positive deflection in the EEG around 300 ms post-stimulus, this type of BCI is actually found to strongly rely on earlier ERP components in addition [29]. These earlier components are particularly strong when the user is allowed to gaze at the symbol, in which case the cognitive P300 component may get less decisive. Though recent studies show that these earlier ERP components may not be applicable when directed gaze is not possible [30,31], these paradigms certainly have the best track record when used for moderately affected end-users. Several studies have adapted this traditional speller to further optimize it for end-user use by making them gaze independent [32], or eliciting additional ERP components [33].

Here, the principles of the matrix speller were used for driving an application that allows the user to receive, view and share photos with friends via the Internet. It is a social application which puts the user in the center of his own group of significant persons. For this purpose, the characters were exchanged for objects such as images, folders and functional symbols (see Fig. 4B). Since the Photobrowser

was operated by the BCI to share photos between the end-user and her family and friends, the application became highly personal and engaging. The social aspect of the Photobrowser application was inspired by recent findings that about 20% of end-users rank 'independent participation in social life' amongst the three most important things that would increase their quality of life [34].

The Photobrowser underwent many iterations with healthy users to improve both the paradigm and the functionality. On top of the simple brightness highlighting from [24], more complex stimuli were available that were found to be particular salient in a prior comparison study [25]. Which of the three available stimulus types was finally applied in the Photobrowser was decided together with the end-user. The SOA was set to the liking of the end-user, and a user optimal EEG channel set was used.

FD was able and allowed to direct her gaze at the target image. Two preparation sessions and six online EEG sessions were performed. Informed consent was given by FD during the first preparation session, and aspects of quality of life and her use of assistive technology were assessed by questionnaires. The second preparation session was used to introduce her to the paradigm and to collect initial calibration data.

3.3. Recordings

For AMUSE and the auditory oddball, EEG was recorded using a fixed set of 61 Ag/AgCl electrodes (see Fig. 5B left) and BrainAmp amplifiers (Brain Products, Munich, Germany). The signals were sampled at 1 kHz and filtered by a hardware analog bandpass filter
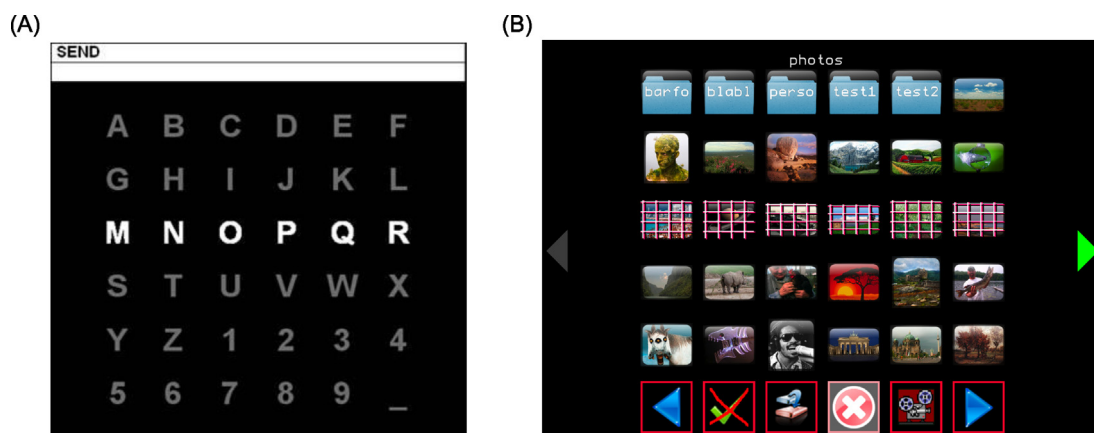

Fig. 4. (A) The traditional spelling interface, in which rows and columns of letters are consecutively highlighted. (B) The interface of the Photobrowser, which adheres to the same basic principle. It is not a spelling application, but rather a photo application which can be used to browse and view photos, and share them over the internet. Note the grid overlayed stimulation, which was neuro-ergonomically optimized in a previous study [25].
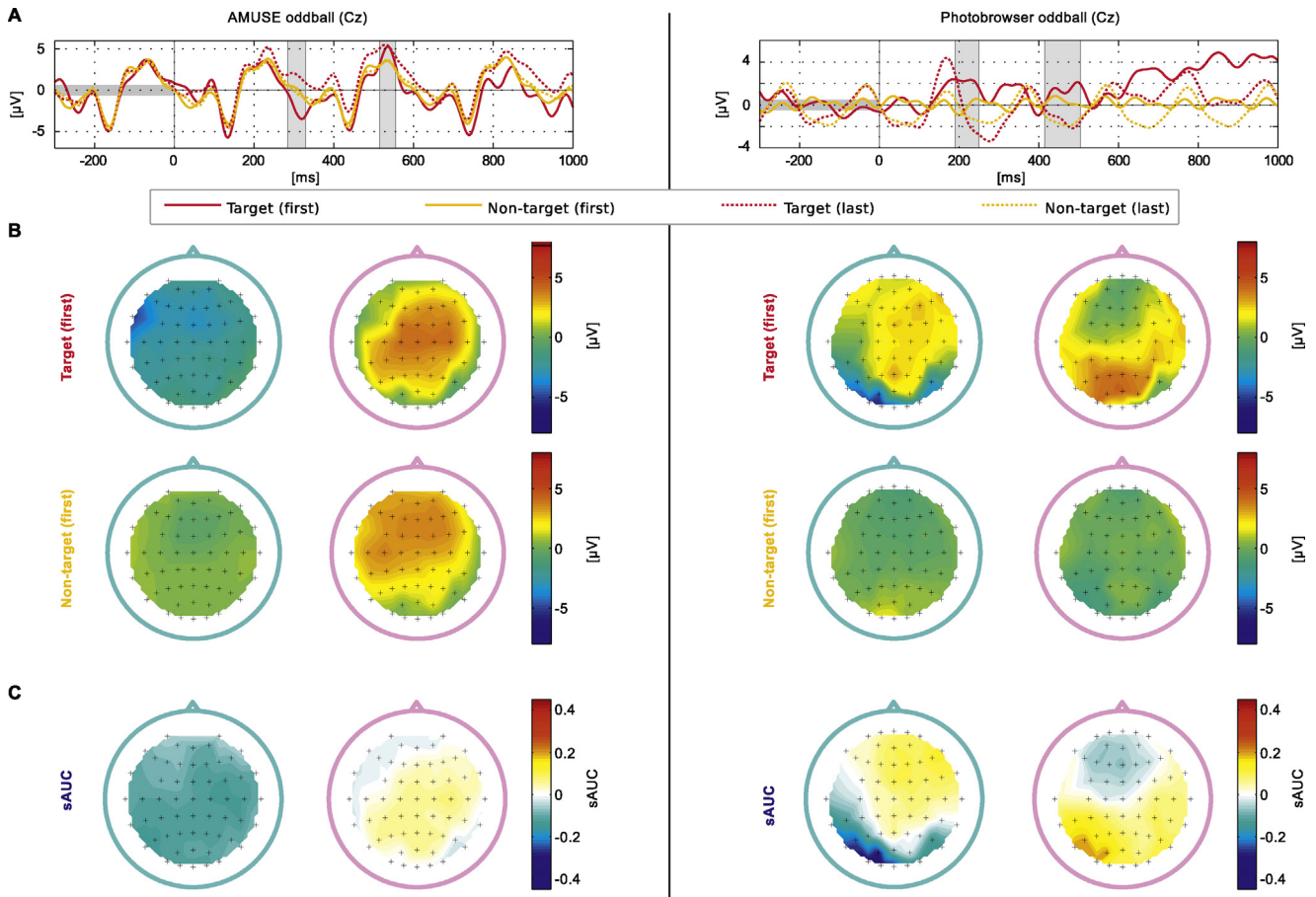
Fig. 5. ERP traces for both BCI paradigms. The left panel indicates the responses to the AMUSE paradigm, whereas the right panel displays the ERP responses to the Photobrowser. (A) The time domain responses for the first and last session on channel Cz are shown. For AMUSE, there is very similar response for targets and non-targets and over sessions. (B) The scalp distributions for the intervals indicated in (A) are given. The class difference in terms of sAUC are given in (C). The clear lack of class discriminative information in the AMUSE paradigm explains why FD was unable to operated the auditory BCI. Class-discriminative information for the Photobrowser is found over the occipital electrodes, as is typical in healthy users. All responses are based on the calibration data. (B) and (C) are based on the first recording session.

between 0.1 and 250 Hz before being digitized and stored for offline analyzes. The same setup was used for the initial session of the Photobrowser, to determine a discriminative channelset. For the visual paradigms, a g.USBamp EEG amplifier (g.Tec, Graz, Austria) was used, with 8 channels for the visual oddball (see Fig. 5B right) and 16 channels for the Photobrowser (Fz, Cz, CP5, 1, z, 2, 6, P5, 1, 2, 6, PO7, z, 8, O1, 2). The signals were sampled at 1.2 kHz and filtered by a hardware analog bandpass filter between 0.1 and 100 Hz before being digitized and stored for offline analyzes. In both cases, channels were referenced to the earlobes.

For online use, the ongoing signal was low-pass filtered below 40 Hz, downsampled to 100 Hz and streamed to the Berlin BCI system for online processing. The stimulus presentation as well as the online Berlin BCI system and the offline analysis were implemented in Matlab (Mathworks), making use of the Psychophysics Toolbox [35] for multi-channel audio presentation. A multi-channel, low-latency Firewire soundcard from M-Audio (M-Audio Firewire 410) was used to individually control the active, off-the-shelf computer speakers (type Sony SRS-A201).

3.4. Signal processing and binary classification

Both the AMUSE and Photobrowser paradigm operated online. Given that the ERP responses are hidden in the background EEG, machine learning methods are typically applied to be able to find the selection that was intended by the user [26]. The following processing chain was used to retrieve features from the (ongoing) EEG and to train and apply a classifier to them. First, the EEG is segmented into epochs starting from 150 ms pre-stimulus up to 1000 ms post-stimulus, and baselined on the pre-stimulus interval. Then, *N* post-stimulus intervals were selected per channel, either based on their discriminative information content or on prior knowledge (see Table 2), and the samples within those intervals were averaged. The *N* features of each channel were then concatenated and feature-wise normalized. The normalization factor was stored for online feature extraction. Based on this feature vector, a linear discriminant analysis (LDA) classifier was trained. In order to prevent over-fitting, the classifier was regularized using shrinkage regularization of the covariance matrix [36,26].

For offline analysis, the data were first bandpass filtered between 0.4 and 15 Hz, applying the filter in forward and backward direction. Though this is an acausal operation and thus not transferable to an online setting, it prevents phase-shifts in the resulting ERPs. This makes the physiological interpretation of the components more robust. During the online use of the BCI system, a causal filter was used instead.

3.5. Multi-class decision process

During the online use of the BCI system, the single binary classifier outputs had to be transformed into a one-out-of-many decision. For this reason, classifier outputs were collected over several stimulus iterations in a decision matrix *D*. The classifier was trained to assign negative scores to target epochs and positive scores to

non-target epochs. Let $D \in \mathbb{R}^{C \times J}$ be a matrix of classifier scores of a trial, where C is the number of classes, and J the number of performed iterations. If $c = \{1, \ldots, C\}$, then let $\tilde{d}$ be a row vector, where $\tilde{d}_c$ denotes the median value of classifier scores for class c. The winning class c^* can be described as $c^* = \text{argmin}_c \ \tilde{d}_c$, i.e. the class with the lowest median value. This was evaluated after a fixed amount of trials.

For the Photobrowser, this was slightly modified due to the matrix structure of the feedback. D was defined by $D_{row} \in \mathbb{R}^{C_{row} \times J}$ and $D_{col} \in \mathbb{R}^{C_{col} \times J}$, where *row* and *col* denote the number of rows and columns, respectively. As above we calculate $\tilde{d}_{row}$ and $\tilde{d}_{col}$, on the respective matrix. The winning class is the intersection of the winning row and column: $c^*_{row} = \text{argmin}_c \ \tilde{d}_{row_c}$ and $c^*_{col} = \text{argmin}_c \ \tilde{d}_{col_c}$.

3.6. Questionnaires

In both of the BCI studies and before each session, FD's motivation was rated by means of the questionnaire for current motivation in BCI QCMBCI [2], consisting of 18 statements to be rated on a 7-point Likert scale and including four factors: mastery confidence, incompetence fear, interest, and challenge. After each session, FD also filled out the multidimensional NASA-TLX questionnaire [37], assessing perceived workload on a scale from 0 to 100 by means of six factors: mental demand, physical demand, temporal demand, performance, effort and frustration.

4. Results

4.1. Neurophysiology

The temporal evolution of the Cz channel for the first and the last session of both the paradigms is depicted in Fig. 5. No clear differences are visible between the target and the non-target ERPs in the AMUSE average. In the first session of the AMUSE paradigm, the negative deflection between 250 ms and 350 ms, which can be identified as an N2 component, presents a slightly greater amplitude for target epochs compared to non-target epochs. Such difference is even less visible in the positive deflection identifiable between 450 ms and 600 ms after the stimulus presentation. In the last session no differences between target and non-target are visible in the early components of the ERPs, but still a very weak difference in the late component (P300) can be identified.

In contrast, the visual ERPs detected in the Photobrowser paradigm are characterized by strong differences between target and non-target epochs, especially in the early components of the ERPs. The temporal evolution of the Cz channel for the Photobrowser paradigm shows that differences between targets and non-targets in the P2 and P3 components in the first session and in the P2 and N2 components in the last session are strongly marked. The topography of values for the class-wise signed area under the receiver-operator curve (sAUC, [26]) shows highly discriminative components at the occipital channels in the interval between 200 ms and 250 ms.

Though a strong N2 is usually found for visual BCI paradigms, FD additionally shows a remarkably strong and discriminative P1 component. This is probably due to the optimized stimulation, which was shown to elicit stronger cortical responses [25].

4.2. Online BCI performance

Online performance for both BCI paradigms can be found in Fig. 6. Performance for AMUSE was above the chance level of 16% after the first session, reaching a peak average performance of 39% in session 4. This was, however, not sufficient for FD to successfully
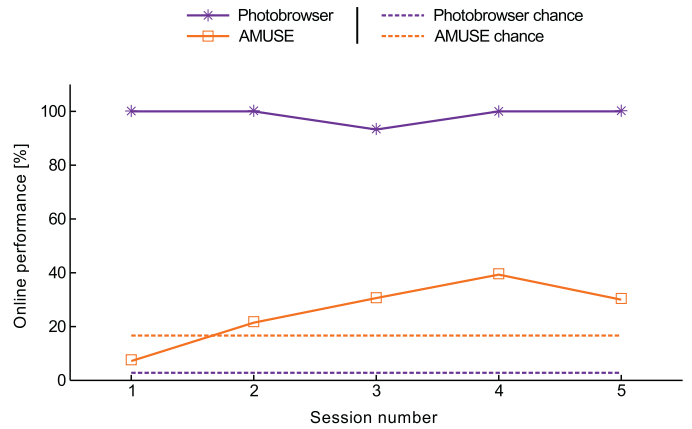

Fig. 6. Online BCI performance. All data for a single session are collapsed over online conditions, and a single performance score is given per session. For most sessions, FD performs above chance with both paradigms. However, the difference between AMUSE and the Photobrowser is striking, with the Photobrowser performance being 100% on all but one session. Striped lines indicate chance level as calculated by the number of classes. AMUSE- and Photobrowser data are from 2010 and 2011, respectively (see Fig. 1).

use the spelling interface. Conversely, the Photobrowser online performance was 100% in four out of five sessions, with the remaining session reaching 93.2% (with chance level below 3%). The subject had near-perfect control over the Photobrowser application. This was mainly accomplished by early visual components over the occipital cortex. The P3 response which is typical for oddball tasks seems to be missing completely, something which is also reflected in the (offline) binary cross-validation loss.

4.3. Neuropsychological assessment

The scores obtained by FD in both the neuropsychological evaluations and the cut-off for each test are reported in Table 1. In both of the evaluations FD was alert, fully cooperating, and oriented in time and space. She performed normally in the mini mental state examination, showing a normal general cognitive level. With regard to attention abilities, her level of performance was comparable between the two assessments when considering the accuracy scores. The scores obtained in the tests were reported in percentiles, and to detect the presence of an attention impairment, a fifth percentile cut-off point was used [38,39].

FD performed above the fifth percentile for both of the selective attention tasks, showing retained capacity to select target stimuli and to inhibit the response to non-target stimuli. On the contrary, she performed below the cut-off on the tasks of divided attention, which is required to process a visual and an auditory task in parallel. Indeed she presented a large number of omissions in the visual task. With regard to the memory abilities, both her verbal short term span (forward digit-span task; [40]) and her verbal working

Table 1
Results of the neuropsychological assessment, and their respective cut-off values. WCST: Wisconsin card sorting test, CBTT: Corsi's block-tapping test, CSBTT: Corsi's supraspan block-tapping test.

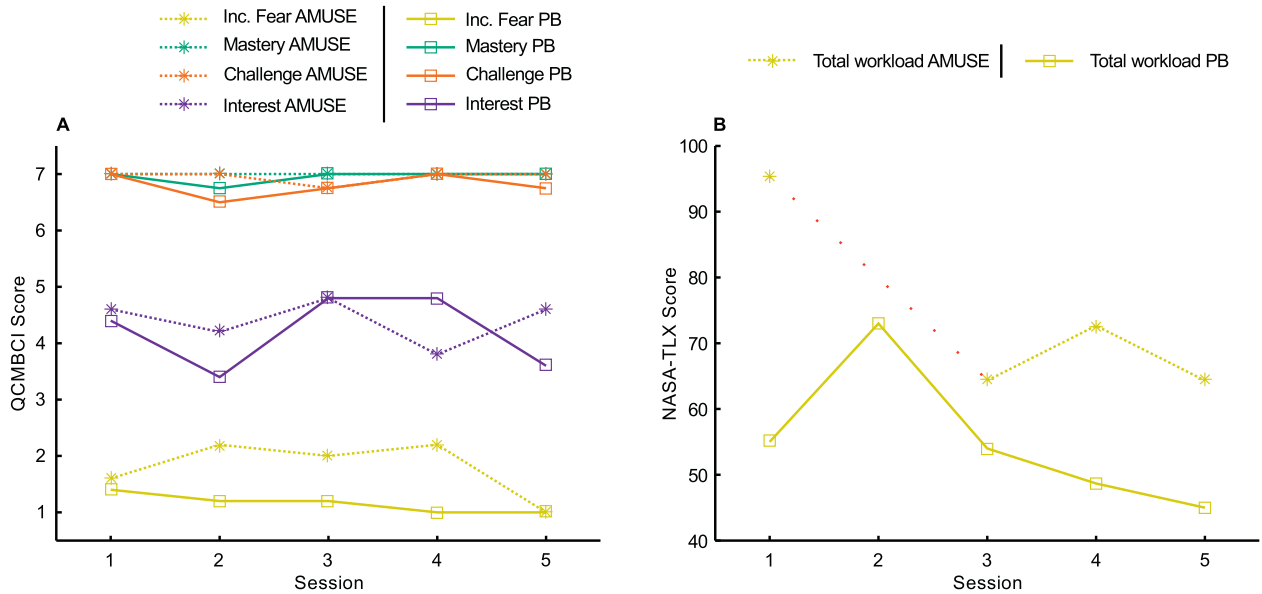|	First	Second	Cut-off
Mini mental test attention	30	30	22
Selective attention	45	45	5
Divided attention Memory Digit span	3	3	5
Forward	8	6	5
Backward	10	10	5
CBTT	2	0	0
CSBTT	0	0	0
N back (2 back)	1	1	5
WCST	+	+	

Fig. 7. Questionnaire results. (A) Scores for the QCMBCI questionnaire, as adapted to BCI, were very similar for both paradigms. Only the incompetence fear for the Photobrowser was found to significantly decrease over sessions. (B) Scores for the NASA-TLX (workload) were generally higher for AMUSE; FD perceived more workload for the AMUSE paradigm. NASA-TLX score for session two of the AMUSE paradigm was missing, and it is interpolated from session 1 and 3 as shown by the red, dotted line. PB refers to the Photobrowser.

Table 2
Paradigm parameters. Stimulation, recording and analyzes parameters are given. Nr. Features indicates the number of features per channel, as selected for the online classification.

	SOA [ms]	Stimulus length [ms]	Nr. channels	Nr. features	Nr. iterations	Online sessions
Auditory oddball	1000	50	61	–	–	–
Visual oddball	2100	93.75	8	–	–	–
AMUSE	300	40	61	3–5	15	5
Photobrowser	220	100	16	21	12	5

memory span (backward digit-span tasks; [40]) were within the normal range. FD performed normally on the visuo-spatial task (Corsi's block tapping test,[19,20]) in the first evaluation. In the second evaluation, a decrease of performance was observed, leading to a pathological score. A similar decrease in performance from the first to the second evaluation was observed in the verbal working memory task, but the score of the second evaluation was still in the normal range. However, even if the scores obtained by FD in both the forward and the backward digit span task did not fall in the pathological range, they were in the lower limit of the normal range (between the fifth and tenth percentile). Furthermore, FD showed an impairment in the Corsi's supraspan block tapping test [19,20] and in the 2-back test. No pathological scores were obtained in the Wisconsin card sorting test.

In summary, the cognitive profile of FD was characterized by a deficit of divided attention, of visuo-spatial learning ability and of the capability to control the information flow and to update information in working memory. In the second assessment, a deficit of visuo-spatial memory span was also found.

4.4. Questionnaire results

Fig. 7A shows that FD scored a comparable degree of motivation over the 5 sessions of the two paradigms, as measured by the four factors of the QCMBCI. As can be seen, the difference between the scores obtained by FD never exceeded 1.2 points (out of 7). Though FB finds both paradigms highly challenging, here confidence in mastering them is equally high and fear of incompetence is absent for both. FB is moderately interested in both

paradims. These results thus give no systematical explanation for the performance difference. Fig. 7B shows that the perceived overall workload, as measured by the NASA-TLX, is scored higher for the AMUSE paradigm than for the Photobrowser, and this is true for all sessions.

A non-parametric Spearman's rank correlation was calculated to address the relationship between time (session number) and, respectively, (i) motivational factors rated with the QCMBCI and (ii) overall workload obtained during the AMUSE and the Photobrowser evaluation, as done in [2]. Only for the Photobrowser did the results show a significant negative correlation between the time and (i) the Incompetence Fear ($\rho = -0.95$; $p < 0.05$) and (ii) the Total Workload ($\rho = -0.90$; $p < 0.05$), indicating that FD became more comfortable with using the Photobrowser over time.

5. Discussion

The BCI end-user reported on here differs from the cases usually described in the BCI literature because of the etiology of her motor- and communication deficits. FD is affected by a brainstem ischemic stroke, causing tetraplegia with severe dysarthria which lead to a serious communication deficit. She participated in the evaluation of a visual- and an auditory BCI paradigm and was, parallel to both trials, screened for attention, memory and executive capabilities. Her cognitive profile was characterized by a deficit of divided attention, visuo-spatial learning ability, the ability to control the information flow and to update information in working memory. Her scores for short-term memory and verbal working memory also fell in the lower limit of the normal range.

FD showed highly class-discriminative components for the slow visual- and in particular the slow auditory standard oddball task. Due to these promising preconditions, an auditory paradigm (AMUSE, [23,16]) was screened first. Unfortunately, FD did not gain sufficient control over the spelling application, much in contrast to the subsequently screened Photobrowser application. This current result is not uncommon, as previous literature already described reduced performance for end-users with auditory BCIs when compared to visual BCIs [41]. However, in this case it can at least partially be explained by FD's clinical profile.

FD indicated the same high level of motivation over the recording sessions for both paradigms, as shown by the motivation questionnaire scores. Therefore, motivation as a factor for the differences in the performances obtained by FD in the two paradigms can be ruled out. The factor *Incompetence Fear* of the QCMBCI questionnaire significantly decreased over sessions, but only for the Photobrowser paradigm. FD thus seemed to gain more confidence in controlling the Photobrowser over sessions. This was not observed in the AMUSE paradigm, most probably because of the difficulty that FD had with the control of the AMUSE spelling application. The latter was also reflected in the scores of perceived workload, which were higher for all the sessions of the AMUSE paradigm. The workload scores significantly decreased over sessions, but just for the Photobrowser paradigm. Clearly, FD became comfortable with handling the Photobrowser, but the low accuracy prevented her from becoming comfortable with the AMUSE paradigm.

As underlined in [42], a paradigm like AMUSE potentially has a high information transfer rate (ITR), in comparison to a binary auditory paradigm [15,43]. This higher ITR comes with a trade-off; an increased demand is put on the subject's attention and working memory to accomplish the task. Furthermore, AMUSE employs covert attention exclusively and does not require additional vision abilities. This means that the user has to keep a mapping of the target symbol to the attended direction in mind. Also, before each trial, all tones are serially presented and the user has to memorize the tone and location of the target stimulus. This requires a greater involvement of the working memory, the capacity to temporarily maintain and to manipulate the information necessary for solving such cognitive tasks. During an ongoing trial, the user is required to identify the target when it is presented in a random sequence together with five competing other tones. Though the spatial features of AMUSE reduce the difficulty of this task [23], it is still a cognitive demand. In hindsight, it appears that with the cognitive deficits that FD has, she is not able to afford a more cognitively complex task such as the AMUSE paradigm.

The Photobrowser application, on the other hand, makes use of overt attention: the user is allowed to gaze at the target directly. Though this may not always be accessible for end-users [30,31], it works very well for those in possession of gaze and eyelid control [44,4]. Allowing the subject to gaze directly at his target reduces the need for keeping the target in memory; once a target is picked the user simply keeps it locked in gaze. This relies less on those cognitive functions where FD shows clinical deficits. This circumstance, along with the optimized visual stimulation applied [25] allowed FD to have near perfect control over the visual application.

Each end-user will fall somewhere on this complexity continuum, which clearly stresses the relevance of testing and adapting BCI-based devices for a particular end-user individually and in accordance with a user-centered design principle. The feedback and the paradigm modality should be matched with the user's clinical profile, so as to reduce the trial-and-error time of the user-centered design process. But whether or not it is based on such clinical profiles, end-users should be offered a range of BCI paradigms, as great differences may exist in their outcome, even when they apparently rely on the same features.

Acknowledgments

References

[1] Bauer G, Gerstenbrand F, Rumpl E. Varieties of the locked-in syndrome. Journal of Neurology 1979;221(2):77–91.

[2] Nijboer F, Birbaumer N, Kübler A. The influence of psychological state and motivation on brain–computer interface performance in patients with amyotrophic lateral sclerosis—a longitudinal study. Frontiers in Neuroscience 2010;4(55), http://dx.doi.org/10.3389/fnins.2010.00055.

[3] Sellers EW, Donchin E. A p300-based brain–computer interface: initial tests by ALS patients. Clinical Neurophysiology 2006;117(3):538–48, http://dx.doi.org/10.1016/j.clinph.2005.06.027.

[4] Nijboer F, Sellers E, Mellinger J, Jordan M, Matuz T, Furdea A, et al. A p300-based brain–computer interface for people with amyotrophic lateral sclerosis. Clinical Neurophysiology 2008;119(8):1909–16, http://dx.doi.org/10.1016/j.clinph.2008.03.034.

[5] Townsend G, LaPallo BK, Boulay CB, Krusienski DJ, Frye GE, Hauser CK, et al. A novel p300-based brain–computer interface stimulus presentation paradigm: Moving beyond rows and columns. Clinical Neurophysiology 2010;121(7):1109–20, http://dx.doi.org/10.1016/j.clinph.2010.01.030.

[6] Vianello A, Arcaro G, Palmieri A, Ermani M, Braccioni F, Gallan F, et al. Survival and quality of life after tracheostomy for acute respiratory failure in patients with amyotrophic lateral sclerosis. Journal of Critical Care 2011;26(3):329.e7–14, http://dx.doi.org/10.1016/j.jcrc.2010.06.003.

[7] Caroscio JT, Mulvihill MN, Sterling R, Abrams B. Amyotrophic lateral sclerosis. Its natural history. Neurologic Clinics 1987;5(1):1–8.

[8] Ringholz GM, Appel SH, Bradshaw M, Cooke NA, Mosnik DM, Schulz PE. Prevalence and patterns of cognitive impairment in sporadic ALS. Neurology 2005;65(4):586–90, http://dx.doi.org/10.1212/01.wnl.0000172911.39167.b6.

[9] Sarro L, Agosta F, Canu E, Riva N, Prelle A, Copetti M, et al. Cognitive functions and white matter tract damage in amyotrophic lateral sclerosis: a diffusion tensor tractography study. American Journal of Neuroradiology 2011;32(10):1866–72, http://dx.doi.org/10.3174/ajnr.A2658.

[10] Phukan J, Elamin M, Bede P, Jordan N, Gallagher L, Byrne S, et al. The syndrome of cognitive impairment in amyotrophic lateral sclerosis: a population-based study. Journal of Neurology, Neurosurgery and Psychiatry 2012;83(1):102–8, doi:10.1136/jnnp-2011-300188.

[11] Laureys S, Pellas F, van Eeckhout P, Ghorbel S, Schnakers C, Perrin F, et al. The locked-in syndrome: what is it like to be conscious but paralyzed and voiceless? In: Laureys S, editor. The boundaries of consciousness: neurobiology and neuropathology, vol. 150 of progress in brain research. Amsterdam, The Netherlands: Elsevier; 2005. p. 495–611, http://dx.doi.org/10.1016/S0079-6123(05)50034-7.

[12] León-Carrión J, Eeckhout Pv, Domínguez-Morales MdR. Review of subject: the locked-in syndrome: a syndrome looking for a therapy. Brain Injury 2002;16(7):555–69, http://dx.doi.org/10.1080/02699050110119466.

[13] Brown AW, Leibson CL, Malec JF, Perkins PK, Diehl NN, Larson DR. Long-term survival after traumatic brain injury: a population-based analysis. Neurorehabilitation 2004;19(1):37–43.

[14] Dennis MS, Burn JP, Sandercock PA, Bamford JM, Wade DT, Warlow CP. Long-term survival after first-ever stroke: the Oxfordshire Community Stroke Project. Stroke 1993;24(6):796–800, http://dx.doi.org/10.1161/01.STR.24.6.796.

[15] Hill NJ, Lal TN, Bierig K, Birbaumer N, Schölkopf B. An auditory paradigm for brain–computer interfaces. In: Saul LK, Weiss Y, Bottou L, editors. Advances in neural information processing systems 17. Cambridge, MA: MIT Press; 2005. p. 569–76.

[16] Schreuder M, Rost T, Tangermann M. Listen, you are writing! Speeding up online spelling with a dynamic auditory BCI. Frontiers in Neuroscience 2011;5(112), http://dx.doi.org/10.3389/fnins.2011.00112.

[17] Magni E, Binetti G, Bianchetti A, Rozzini R, Trabucchi M. Mini-Mental State Examination: a normative study in Italian elderly population. European Journal of Neurology 1996;3(3):198–202, http://dx.doi.org/10.1111/j.1468-1331.1996.tb00423.x.

[18] Zimmermann P, Fimm B. Test for attentional performance (TAP). Herzogenrath: PsyTest; 1995.

[19] Corsi P. Human memory and the medial temporal region of the brain, Ph.D. thesis. Montreal, Canada: McGill University; 1972.

[20] Spinnler H, Tognoni G. Standardizzazione e taratura italiana di test neuropsicologici: gruppo italiano per lo studio neuropsicologico dell'invecchiamento, Italian Journal of Neurological Sciences. Milan, Italy: Masson Italia Periodici; 1987.

[21] Heaton RK, Chelune GJ, Talley JL, Kay GG, Curtiss G. Wisconsin Card Sorting Test Manual – Revised and expanded. Odessa, FL: Psychological Assessment Resources Inc.; 1993.

[22] Heaton RK, Chelune GJ, Talley JL, Kay GG, Curtiss G. WCST: Wisconsin Card Sorting Test (Italian). Firenze: Giunti-Organizzazioni Speciali; 2000.

[23] Schreuder M, Blankertz B, Tangermann M. A new auditory multi-class brain–computer interface paradigm: Spatial hearing as an informative cue. PLOS ONE 2010;5(4):e9813, http://dx.doi.org/10.1371/journal.pone.0009813.

[24] Farwell LA, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalography and Clinical Neurophysiology 1988;70(6):510–23, http://dx.doi.org/10.1016/0013-4694(88)90149-6.

[25] Tangermann M, Schreuder M, Dähne S, Höhne J, Regler S, Ramsay A, et al. Optimized stimulation events for a visual ERP BCI. International Journal of Bioelectromagnetism 2011;13(3):119–20.

[26] Blankertz B, Lemm S, Treder MS, Haufe S, Müller K-R. Single-trial analysis and classification of ERP components – a tutorial. Neuroimage 2011;56(2):814–25, http://dx.doi.org/10.1016/j.neuroimage.2010.06.048.

[27] Lemm S, Blankertz B, Dickhaus T, Müller K-R. Introduction to machine learning for brain imaging. Neuroimage 2011;56(2):387–99, http://dx.doi.org/10.1016/j.neuroimage.2010.11.004.

[28] Wightman FL, Kistler DJ. Resolution of front-back ambiguity in spatial hearing by listener and source movement. Journal of the Acoustical Society of America 1999;105(5):2841–53, doi:10.1121/1.426899.

[29] Bianchi L, Sami S, Hillebrand A, Fawcett IP, Quitadamo LR, Seri S. Which physiological components are more suitable for visual ERP based brain–computer interface? A preliminary MEG/EEG study. Brain Topography 2010;23(2):180–5, http://dx.doi.org/10.1007/s10548-010-0143-0.

[30] Treder MS, Blankertz B. (C)overt attention and visual speller design in an ERP-based brain–computer interface. Behavioral and Brain Functions 2010;6(1):28, http://dx.doi.org/10.1186/1744-9081-6-28.

[31] Murguialday AR, Hill J, Bensch M, Martens S, Halder S, Nijboer F, et al. Transition from the locked in to the completely locked-in state: a physiological analysis. Clinical Neurophysiology 2010;122(5):925–33, http://dx.doi.org/10.1016/j.clinph.2010.08.019.

[32] Treder MS, Schmidt NM, Blankertz B. Gaze-independent brain–computer interfaces based on covert attention and feature attention. Journal of Neural Engineering 2011;8(6):066003, http://dx.doi.org/10.1088/1741-2560/8/6/066003.

[33] Kaufmann T, Schulz SM, Grünzinger C, Kübler A. Flashing characters with famous faces improves ERP-based brain–computer interface performance. Journal of Neural Engineering 2011;8(5):056016, doi:10.1088/1741-2560/8/5/056016.

[34] Zickler C, Di Donna V, Kaiser V, Al-Khodairy A, Kleih SAK, et al. BCI applications for people with disabilities: defining user needs and user requirements. In: Emiliani PL, Burzagli L, Como A, Gabbanini F, Salminen A-L, editors. Assistive technology from adapted equipment to inclusive environments, AAATE, vol. 25 of assistive technology research series. Amsterdam, The Netherlands: IOS Press; 2009. p. 185–9, http://dx.doi.org/10.3233/978-1-60750-042-1-185.

[35] Brainard DH. The psychophysics toolbox. Spatial Vision 1997;10(4):433–6, http://dx.doi.org/10.1163/156856897X00357.

[36] Ledoit O, Wolf M. A well-conditioned estimator for large-dimensional covariance matrices. Journal of Multivariate Analysis 2004;88(2):365–411, http://dx.doi.org/10.1016/S0047-259X(03)00096-4.

[37] Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Human Mental Workload 1988;1(3):139–83.

[38] Pero S, Incoccia C, Caracciolo B, Zoccolotti P, Formisano R. Rehabilitation of attention in two patients with traumatic brain injury by means of 'attention process training'. Brain Injury 2006;20(11):1207–19, http://dx.doi.org/10.1080/02699050600983271.

[39] Zoccolotti P, Matano A, Deloche G, Cantagallo A, Passadori A, Leclercq M, et al. Patterns of attentional impairment following closed head injury: a Collaborative European study. Cortex 2000;36(1):93–107, doi:10.1016/S0010-9452(08)70839-6.

[40] Orsini A. La memoria diretta e la memoria inversa di cifre in soggetti dai 16 ai 64 anni. Bollettino di Psicologia Applicata 2003;239:73–7.

[41] Kübler A, Furdea A, Halder S, Hammer EM, Nijboer F, Kotchoubey B. A brain–computer interface controlled auditory event-related potential (p300) spelling system for locked-in patients. Annals of the New York Academy of Sciences 2009;1157(1):90–100, http://dx.doi.org/10.1111/j. 1749-6632.2008.04122.x.

[42] Riccio A, Mattia D, Simione L, Olivetti M, Cincotti F. Eye-gaze independent eeg-based brain–computer interfaces for communication. Journal of Neural Engineering 2012;9(4):045001, http://dx.doi.org/10.1088/1741-2560/9/4/045001.

[43] Halder S, Rea M, Andreoni R, Nijboer F, Hammer EMSCK, et al. An auditory oddball brain–computer interface for binary choices. Clinical Neurophysiology 2010;121(4):516–23, http://dx.doi.org/10.1016/j.clinph.2009.11.087.

[44] Münßinger JI, Halder S, Kleih SC, Furdea A, Raco VAH, et al. Brain painting: first evaluation of a new brain–computer interface application with als patients and healthy volunteers. Frontiers in Neuroscience 2010;4(182), http://dx.doi.org/10.3389/fnins.2010.00182.