A Brain-Computer Interface as Input Channel for a Standard Assistive Technology Software

Claudia Zickler, Angela Riccio, Francesco Leotta, Sandra Hillian-Tress, Sebastian Halder, Elisa Holz, Pit Staiger-Sälzer, Evert-Jan Hoogerwerf, Lorenzo Desideri, Donatella Mattia and Andrea Kübler

Key Words
Assistive Technology
Brain-Computer Interface
Electroencephalography
Information and Communication Technology
P300 Event-Related Potential
Usability
Users with Disabilities

ABSTRACT
Recently brain-computer interface (BCI) control was integrated into the commercial assistive technology product QualiWORLD (QualiLife Inc., Paradiso-Lugano, CH). Usability of the first prototype was evaluated in terms of effectiveness (accuracy), efficiency (information transfer rate and subjective workload/NASA Task Load Index) and user satisfaction (Quebec User Evaluation of Satisfaction with assistive Technology, QUEST 2.0) by four end-users with severe disabilities. Three assistive technology experts evaluated the device from a third person perspective. The results revealed high performance levels in communication and internet tasks. Users and assistive technology experts were quite satisfied with the device. However, none could imagine using the device in daily life without improvements. Main obstacles were the EEG-cap and low speed.

INTRODUCTION
Following the concept of e-Inclusion, the aim to “leave no-one behind” when enjoying the benefits of information and communication technology (ICT), we report on the evaluation of a recent integration of a brain-computer interface (BCI) technology with the commercial accessibility software QualiWORLD (QualiLife SA, Lugano, Switzerland).1,2 We collected data from assistive technology (AT) end-users and AT-experts outside of a laboratory environment.

AT-products have been developed to improve the user’s access to the environment which may otherwise be difficult or impossible. Despite progress in AT there are still a large number of people with severe motor disabilities due to motor degeneration (e.g., amyotrophic lateral sclerosis (ALS), Duchenne muscular dystrophy) or brain damage (e.g., stroke, traumatic brain injury) who cannot fully benefit from AT because it requires muscular control. Here, BCIs offer a solution as they allow users interaction and communication with the environment on the basis of classification and translation of brain signals into commands for applications.3 In the proof-of-concept-phase it has been shown that it is possible to use BCI for controlling keyboards,4-6 browsing the internet,7 playing games,8 painting9 or using a predictive spelling program.10 By now, the time has come to move BCI out of the lab into real life and develop practical BCI controlled ATs that offer access and speed comparable to that of muscular controlled standard software.

P300 BCI
For this new AT-solution the P300 BCI was chosen as input signal to the software program. The P300 is a positive deflection in the electroencephalogram (EEG) that occurs 200-700 ms after stimulus onset. It is evoked when participants attend to a rare stimulus (target) in a random series of stimulus events.5,11 It has been shown repeatedly that people with severe motor disabilities can control a P300 BCI.10,11,12

QualiWORLD interface
Sellers and colleagues13 reported that a BCI system was combined with standard software, such as Microsoft Office, although the specific implementation was not described. The goal of the current study was to combine BCI technology with a commercial AT-software that offers a complete set of communication and control functions for users with disabilities. Off-the-shelf user interfaces offer powerful functions. However, users with disabilities may not benefit from all interface features, such as many small icons or multiple open windows on a complex desktop. Features that are not adapted to specific users may render information technology challenging and frustrating. Common AT-solutions such as screen magnifiers or text-to-speech software, whilst providing great benefit to some users, are simply bolt-on solutions and, therefore, cannot solve all the limitations inherent in standard user interfaces.14 Thus, a solution may be to design a simpler user interface built bottom up with desired interactions of specific users in mind. “QualiWORLD” (QW) is an example of such a commercial AT-ICT-product. It provides users with functions for communication, such as word processing with built-in speech generation and email sending, for entertainment, such as music and video player, games, and photo browser and for environmental control, such as access to TV. In addition, QW offers an embedded internet browser. The QW-software can be controlled via a variety of different input channels, such as a scanning interface or a head-mouse, thereby providing people with disabilities the possibility of using the same software in different stages of the disease by simply changing the input channel.

User evaluation
The success of an AT is determined by how well it suits its purpose and meets the needs of the targeted users, in other words: the usability of the product.15 In the 1980s the term “user-centered design” was established and has become widely used.16 Today the philosophy of
user-centered design has found its standardization in the ISO 9241-210 (Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems). It defines the following principles for a human-centered design process: (a) understand the user, the task and environmental requirements, (b) encourage early and active involvement of users, (c) be driven and refined by user-centered evaluation, (d) include iteration of design solutions, (e) address the whole user experience, (f) encourage multi-disciplinary design.

The present study reports on the user-centered evaluation of the first prototype of a P300 BCI combined with the commercial AT-software QW. Evaluation focused on the whole AT-device and its control with the input channel “P300 BCI.” The goal was to compile user requirements that will support and guide further development and enhancement.

User and task

Participants in a human-centered design process should be chosen to match the expected user population as closely as possible. In the process of product development the possibility to receive feedback from users is of utmost importance. Therefore, we (1) included people with severe disabilities due to degenerative diseases who could speak, (2) were familiar with the needs and requirements of users with severe motor disabilities, (3) if possible having some experience with a P300 BCI. Thus, we chose spelling tasks in the word-processing function, an email writing and sending task, and an internet browsing task.

In accordance with the ISO 9241-210, usability of the prototype was tested in terms of its effectiveness, efficiency and satisfaction.

METHODS

Four end-users and three AT-experts evaluated the usability of a commercial AT-software (QW) controlled by the P300 BCI. BCI stimulation was superimposed on the QW-interface thereby allowing the use of all functions of QW including an internet browser. In contrast to the classic P300 stimulation where rows and columns are flashing, the stimulation on the QW-interface and the web pages was performed by dots appearing directly next to the icon or link. Effectiveness, efficiency and user satisfaction were assessed for evaluation in two spelling tasks, an email sending and an internet browsing task. Based on the ISO 9241-210, effectiveness refers to how accurate and complete users accomplish the task. Efficiency relates the invested costs, i.e., users’ effort and time, to effectiveness. It was operationalized as information transfer rate (ITR) and subjective workload. Satisfaction is the perceived comfort and acceptability while using the product and was assessed for end-users and AT-experts.

Participants

Four potential users (4 male, age 55, 50, 39, 37 years) with severely restricted motor function participated in the study. Although none of the users were in the locked-in state, in which only few muscles are available for interaction, they were severely disabled and fully dependent on AT-input devices for computer access. Three of them had prior experience with P300 BCI (see Table 1). Thus, 3 users fulfilled all three inclusion criteria; 1 user fulfilled two inclusion criteria.

Three AT-experts (expert A: age 57, male; expert B: age 27, female, expert C: age 55, male) working at the Information Center for Supported Communication (BUK), Bad Kreuznach, Germany, with educational backgrounds in occupational therapy, software programming, and communication education participated as observers during sessions of users B and C and evaluated the device from an expert third person perspective. Two AT-experts had personal experience using the classic P300-speller introduced by Farwell and Donchin.

Data acquisition

EEG was measured using a passive 8-electrode cap (Fz, Cz, Pz, Oz, P3, P4, P7, P8) based on the International 10-20 system. All channels were referenced to the right earlobe and grounded to the right mastoid. Impedances were kept below 5 kΩ and the EEG was amplified using a g-tec g.USBamp amplifier (Graz, Austria). The EEG signals were band-pass filtered between 0.1 and 30 Hz, sampled at 256 Hz, and recorded by the BCI2000 software. The EEG signal was fed into the BCI2000 signal processing pipeline for online classification.

Calibration and signal classification

For calibration, users had to select 5 pre-set letters each from four different matrices (3x3 matrix (Figure 1a): ACEGI, 4x4: ACFLO, 5x5: AGMSY, 6x6: RADIO). Red dots (radius 9 pixels) presented on the upper left corner of each letter in the matrix flashed row and column wise for stimulation. A grey frame around the matrix cell indicated the target letter. Pre-tests had shown that the calibration with the four different matrix sizes worked well such that the same coefficients could be used for all applications (see section BCI QW-Application). Data acquired during calibration were loaded into an off-line analysis software (based on Matlab environment; the MathWorks Inc. Massachusetts) for classifier training. Stepwise Linear Discriminant Analysis (SWLDA) was used for signal classification.
The SWLDA adds and removes variables from a multi-linear model following their statistical significance in the discriminant function. It starts with an initial model and then performs a series of forward and backward regression procedures in discrete steps. During the forward regression the variable, not yet included in the model, showing the lowest p-value below the tolerance threshold (0.10) is added to the model. During the backward regression the variable, already included in the model, showing the highest p-value above the tolerance threshold (0.15) is removed from the model. The backward regression procedure stabilizes addition of new variables which modify the model, potentially making an already included variable no longer suitable, which is then removed. This selection process was repeated until no further predictor variables satisfied the inclusion/exclusion criteria or until the predefined number of features (60) was reached.24

Online, the signals were fed into a linear classifier whose coefficients were obtained by SWLDA. The classifier assigned each stimulus a score. This score was then used to choose the command to execute.

Analysis of EEG data

All EEG-data were analyzed using Brain Vision Analyzer Version 1 (BrainProducts GmbH, Munich, Germany) and filtered (1-20 Hz). Data segments were extracted from the onset markers of each “flashing” period until 1000ms after the end marker of each period. An interval of 100ms before the start marker was used for baseline correction. The EEG was averaged across target and non-target segments separately for each of the tasks and participants.

The scalp distribution of the P300 is known to be centro-parietal with highest amplitudes over midline scalp sites.25 Therefore, the P300 amplitudes per task and participant were calculated for the Cz-electrode. A P300 was defined as the difference between baseline (-100-0ms) and the most positive peak between 200-500ms post stimulus.

To provide a measurement of accuracy under conditions matching the “use in daily life” as close as possible (in which an environment free of artifacts is unlikely) EEG-data was not corrected for eye-blinks and other artifacts.

BCI QW-application

The application consisted of three components: the QW-software by QualiLife, the BCI200026 component and the QW-stimulation component. The QW-software and the QW-stimulation component ran on the same mid-range laptop as two separate processes. Communication between the QW-stimulation component and the QW-software was performed using an application programming interface (API) provided by QualiLife. The BCI2000 component ran on a second laptop. The end-user saw only the screen with the QW-interface.

The QW-software automatically expands to the whole screen, not allowing for another graphical interface to be run at the same time. Had we used only one laptop, it would have been impossible for the investigator to monitor the BCI2000 without disturbing the participant during task performance. For this reason we ran the BCI2000 on a separate laptop, although a system with one laptop only is desirable in the future.

The BCI2000 component sent the P300 stimulation sequence and the classification result to the QW-stimulation component via network connection. The QW-stimulation component was responsible for superimposing the stimulation (see below) over the QW-interface and sending commands to the QW-software. For each selection, the QW-stimulation component received from the QW-software the list of commands (icons, letters) available on the QW-interface. Commands in the QW-software do not follow a matrix layout. Therefore, a mapping operation between these commands and the positions of the BCI2000 matrix was performed by the QW-stimulation component. A red dot was assigned automatically to all available commands of the QW-software. Thus, superimposition was entirely flexible. The red dots corresponded to the cells of the P300 matrix of the BCI2000 software and appeared directly next to the selectable letters, icons or links in the QW-interface (Figure 1). For reasons of standardization, all users were presented with red dots. However, the program allows the user to choose the color individually.

The matrix size of the BCI2000 was automatically chosen according to the number of available commands. For selection of icons in the QW-interface a 3x3 matrix in the BCI2000 component was sufficient, with two exceptions: the 8x7 spelling-matrix of the word-processing function (Figure 1b) required an 8x7 matrix in the BCI2000 component, and for the web pages used in the internet task (see Procedures) the number of available links required either 5x5 or 7x7 matrices.

In the spelling-matrix of the word-processing function the end-user was presented with red dots flashing row and column wise (Figure 1b). For the other selections the user was presented with flashing single dots next to the selectable icons or links (see Figure 1c and 1d). We refer to the latter as “single dot stimulation.”

Procedure

Users evaluated the BCI QW-device at home (users A and D) or at the BUK (users B and C) with caregivers, family members or AT-experts present. No effort was made to reduce daily life distractions such as telephone ringing, people entering and leaving the room, or noisy construction sites directly outside the BUK.

During the measurements, all users sat in their wheelchair at a distance of approximately 1m from the computer screen on which the QW-interface was displayed. Users performed four tasks with progressive levels of difficulty on four days (one task per day).

The first session started with calibration (see above). In the following copy spelling task participants were required to spell three pre-set words “COMPUTER”, “EMAIL”, and “INTERNET” using the 8x7 spelling matrix in the QW-word-processing function (Figure 1b) with a short break after each word. The words were displayed on a piece of paper fixed at the top of the screen. The selected letters appeared next to the matrix on the “page” in the word-processing function (see Figure 1b “HALLO”). In case of an error, subjects were instructed not to correct it but to proceed with the next selection. For the second session, users opened word processing, selected the 8x7 matrix and typed a self-selected 20-character sentence (including spaces and punctuation). In the third session users had to choose the word-processing function and the 8x7 matrix, write the word “HALLO,” close the matrix and then perform nine selections in the menu to send the text in an email to a predefined address (Figure 1c). In the fourth session users had to choose the “internet browser-“icon” where a pre-selected internet page appeared (webpage of the German society of people with muscular diseases, http://www.dgm.org/; Figure 1d). Users were asked to select specified links and perform tasks such as zoom in/out and scroll up/down.

The total number of selections (provided no errors to be corrected) was 21 in the copy spelling, 22 in the free spelling, 17 in the email and 15 in the internet task. Errors had to be corrected in all tasks except copy-spelling. A short break was provided after every 10 selections. In the email and internet tasks subjects were told and shown by the investigator which icon to choose next. Before each task, subjects were introduced to the program and could practice. They were asked to focus their attention on the item to be selected by counting the number of times the “stimulation dot” next to the item was flashing.
For all users and all sessions (incl. calibration) one symbol selection consisted of 20 flashes (10 flashes per row and column in the BCI2000 component = 10 sequences). The duration of each flash was 125ms with 125ms inter-stimulus interval. Before and after each block of 10 sequences a pre-interval of 5s and a post-interval of 1360ms were inserted. The long pre-interval was chosen to allow users to become orientated in the applications that they — except for a short practice phase at the beginning of the session — were using for the first time. Selection of a letter in the 8x7 matrix of the word-processing function required 43.86s. The links on the webpage that had to be selected in the internet browsing task had a 5x5 and 7x7 matrix in the BCI2000 component with selection times of 31.36s and 41.36s respectively. The selection of an icon on the QW-interface with a 3x3 matrix in the BCI2000 component took 21.36s.

Questionnaires and user feedback

After every task, subjective workload was assessed with the NASA Task Load Index (TLX). After the free spelling, the email and the internet task the extended Quebec User Evaluation of Satisfaction with assistive Technology (QUEST 2.0) and a visual analogue scale (VAS) “overall device satisfaction” were administered. After the last task end-users gave feedback in a short interview.

Effectiveness

As effectiveness refers to how accurate and complete users accomplish the task we measured accuracy per task by taking the number of correct selections and dividing this value by the total number of selections.

Efficiency

Information transfer rate

The quantification of successful BCI use can be derived from computing the accuracy (percent of correctly selected targets), the information content within each selection (number of different targets) and the time needed per selection. Together these determine the information transfer rate (ITR) in bits/min incorporating speed and accuracy in a single value. The ITR was calculated to measure the...
speed of command selection according to Wolpaw and colleagues:

\[ B = \log_2 N + P \log_2 P + (1 - P) \log_2 \left( \frac{1 - P}{N - 1} \right) \]

with \( N \) being the number of possible selections in the matrix and \( P \) being the accuracy of the participant. Every time an error was made, the required correct backspace and the selection of the intended item were included in the calculation of the ITR.

Subjective workload

Users' subjective workload was assessed with the NASA Task Load Index (TLX), which identifies (1) the overall workload in the different tasks and (2) the main sources of workload. Workload in the TLX is defined as a “hypothetical construct that represents the cost incurred by a human operator to achieve a particular level of performance.” (p. 140). The TLX is specifically adequate when interested in detecting the sources of workload. Workload is estimated with six subscales (mental, physical, and temporal demand and performance, effort, and frustration). Participants rated subjective workload for each dimension on twenty step bipolar scales with scores from 0 to 100. A weighting procedure was used to combine the six individual ratings into a global score. To do so, the six scales were combined to 14 pairs and subjects had to indicate which scale of the pair contributed more to their workload. A weighted average technique was then used to compute an overall measure of workload (between 0 and 100) and the relative contribution of each subscale to overall workload.

Satisfaction

Extended QUEST 2.0

The Quebec User Evaluation of Satisfaction with assistive Technology (QUEST 2.0) is the only standardized satisfaction assessment tool that was designed specifically for AT-devices. It was used to assess satisfaction with different aspects of the BCI-controlled AT-device. The questionnaire consists of 12 items. Four items (durability, service delivery, repairs/servicing, follow-up services) were not adequate for the evaluation of a prototype and were thus, removed from the questionnaire (see Table 3). Validity is ensured if no more than six items are omitted. As the QUEST 2.0 was developed to assess satisfaction with a wide range of technology, the authors invite researchers to add a few items to render the questionnaire more suitable for a specific piece of technology. However, such data cannot be added to the total score of the QUEST.

Four items (reliability, speed, learnability, and aesthetic design) were added to render the QUEST more suitable for evaluation of BCI controlled AT. The whole questionnaire will be referred to as “extended QUEST 2.0.” Items were to rate on a scale from 1 (not satisfied at all), 2 (not very satisfied), 3 (more or less satisfied), 4 (quite satisfied) to 5 (very satisfied). Users were asked to comment if they were not “very satisfied.” A total satisfaction score can be obtained by summing across the QUEST items and dividing the sum by the number of items. In addition, a separate total score was calculated for the four added items. Finally, users were asked to indicate the three most important items.

VAS

Overall device satisfaction was assessed with a visual analogue scale (VAS), ranging from 0 (not at all satisfied) to 10 (absolutely satisfied).

After the final session, users were asked in an open interview about their suggestions for improvement and their requirements for using the device in daily life.

Feedback from AT-experts

After the last session, AT-experts were asked to give written feedback on the different aspects of the BCI QW-device according to the extended QUEST 2.0.

Statistical analysis

As the group of users was too small for inferential statistical analysis, individual data will be reported descriptively.

RESULTS

P300 amplitudes

All users completed all four tasks. Users' individual P300 amplitudes in the four tasks are depicted in Figure 2. Except for user

---

**Figure 2.** Average EEG-amplitudes as a function of time for \( N = 4 \) end-users depicted for all tasks (see figure bottom right for scaling).
C, who had very low amplitudes in all tasks, the highest amplitudes were found in the copy and free spelling tasks.

Effectiveness

Accuracy was above 70% for all users in all tasks (Table 2). In the copy spelling task three users (A, C, D) achieved 100%. Best performance in the other tasks was 95.8% (free spelling), 100% (email), and 93.8% (internet) (user A).

Efficiency

Information transfer rate

Highest ITRs were achieved by user A with 8.56 bits/min in the email task. In the copy spelling task users A, C and D achieved an ITR of 7.95 bit/min (ITR for all users and tasks in Table 2).

Subjective workload

Users A and C rated their subjective workload comparably high in the different tasks (Table 2), while user D rated his subjective workload highest in the internet task (46). User B who had used a BCI for the first time indicated that his subjective workload decreased considerably with every session (from 49 to 15). Mental demand or temporal demand contributed most to workload for all users in almost all tasks. Physical demand (all users), frustration (users A, B, and D) and effort (users B, C and D) ratings were low.

Satisfaction

Extended QUEST 2.0

According to the total scores of the QUEST 2.0 and the added items, users A and C were “quite satisfied” (4) to “very satisfied” (5) in all tasks, while users B and D were “more or less satisfied” (3) with the BCI QW-device. We first report on the six items that were related to aspects of the device that were the same in all tasks and second on the items related to aspects that could vary in the different tasks.

All users were “quite satisfied” (4) to “very satisfied” (5) with the "safety," two users (A and C) were also “very satisfied” (5) with the “dimensions" and the "weight," while the other two were only “more or less” (3) or “not very satisfied” (2) judging the device as “too big” (user B), and consisting of “too many parts” (user D). None of the users were “very satisfied” (5) with the “adjustment” as adjusting the “electrodes takes too long” (user C), “checking the impedances is time-consuming” (user A), the “gel is unpleasant” (user B), there were “too many different parts that have to be attuned to each other” (user D) and the setup of the software was judged as being “very technical” (user C). Likewise none of the users were “very satisfied” (5) with “comfort”, and "aesthetic design." Three users (A, B, D) indicated that they did not find the cap “comfortable” as it felt like a “foreign body.” Due to the cables of the electrodes they felt that their “mobility is restricted” (users A, B) such that they could neither change their sitting position in the wheelchair (without help) nor move around in their wheelchair. Two users (A, C) would not feel comfortable using the device in every environment because the cap “looks very medical” (user C). The users indicated that the device was not “aesthetic” and especially the EEG-cap was “eye-catching” (users C, D).

All users were “quite satisfied” (4) to “very satisfied” (5) with the “professional services (instructions)” and the “learnability” in all tasks. With regard to “ease of use,” “effectiveness," and "reliability" users A and C were “quite” (4) or even “very satisfied” (5) in all tasks, while the other users showed lower ratings in at least one task. Except for user A in the free spelling and the internet tasks no one was “very satisfied” (5) with the “speed” (see Table 3). With regard to the selection of an item, speed was judged as “too slow” by three users (A, C, D) and should be “twice as fast” (user A), “3-4 times as fast” (user D). To underline his reasons for dissatisfaction user C compared the BCI-device with the AT (Wergen-system) he is currently using, indicating that “with my own AT I can write 90 characters per minute.”

Users were further asked to indicate the three most important items each time the QUEST was administered. All users rated at least in one task “effectiveness” (N=10) as one of the three most important aspects, followed by “reliability” (N=8) that was rated by user A, B, and C in all

<table>
<thead>
<tr>
<th>Task</th>
<th>End-user A</th>
<th>End-user B</th>
<th>End-user C</th>
<th>End-user D</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>100 (21/21)</td>
<td>85.7 (18/21)</td>
<td>100 (21/21)</td>
<td>100 (21/21)</td>
</tr>
<tr>
<td>FS</td>
<td>95.8 (23/24)</td>
<td>76.5 (26/34)</td>
<td>72.7 (32/44)</td>
<td>88.9 (26/34)</td>
</tr>
<tr>
<td>Email</td>
<td>100 (17/17)</td>
<td>90.0 (19/21)</td>
<td>90.5 (19/21)</td>
<td>76.0 (19/25)</td>
</tr>
<tr>
<td>Internet</td>
<td>93.8 (15/16)</td>
<td>88.2 (15/17)</td>
<td>70.6 (24/34)</td>
<td>73.3 (22/30)</td>
</tr>
</tbody>
</table>

Note. CS=copy spelling task, FS=free spelling task, Email=email task, Internet=internet task. NASA TLX: M=mental demand, Ph=physical demand, T=temporal demand, P=performance, E=effort, F=frustration, S=total score
tasks as most important. Furthermore, “speed” (N=6, users C and D) and “comfort” (N=6, users B and D) were indicated. Interestingly, user B who used a BCI for the first time indicated “learnability” as most important in the free spelling task, but no longer in the email and internet tasks (see Table 3).

**Feedback from end-users**

Users were asked about their requirements for using the device in daily life. All users indicated the need for a better solution for the EEG-cap. Adjustment should be easier and there should be no cables. The hardware should be reduced to one compact device. Environmental and wheelchair control should be integrated. Finally, all users indicated that higher speed in selecting items would be crucial for use in daily life. None of the users could imagine using the BCI-controlled QW-interface in its current state, but could imagine doing so if there was improvement in equipment or disease progression.

**Feedback from AT-experts**

Three AT-experts indicated that the device consisted of too many different parts, interfaces and cables, and that the setup time for the whole equipment was too long. Furthermore, it was stated that “there is a need for careful training” (expert B) to learn to adjust the EEG-cap and the electrodes correctly. Expert C indicated that the EEG-cap and the electrodes might have a negative effect on the reliability of the device as the adjustment was a “source of errors.” Setting up the software required “taking care of different submenus” which was not a satisfactory solution (expert B). However, AT-experts indicated that “learnability” with regard to use by the end-user was not a major problem (experts A, B, C).

With regard to “ease of use” and “effectiveness” experts (A, B) indicated the long selection times as a disadvantage, particularly when the 8x7 matrix for text entry was used. Furthermore, the “standard of 10 sequences during the study took long” and added to the selection time (expert B). With regard to “comfort” of the device, the AT-experts criticized the restricted mobility during the use of the device. Also, chewing, swallowing, or other movements of the facial muscles might lead to disturbing artifacts. Some users might experience a feeling of discomfort due to the cap, and dislike of the gel and the need to have their hair washed after using the device (experts B, C).
AT-experts were asked for a general feedback on the BCI QW-device. They indicated that for the future, the BCI QW-device was an option to provide users with an AT that needs no muscular control. Compared to a classic P300 speller (known from other studies) the QW-interface with its different functionalities was a clear improvement. In addition, the QW-software was comparable to standard software, so that users recognized it as familiar and could resort to previous knowledge. However, the future of the BCI-device was considered dependent on a better solution for the EEG-cap and the electrodes.

**DISCUSSION**

The outcome of the evaluation was first, a proof of feasibility of the combination of BCI with commercial AT-software and second, a description of user requirements for establishing priorities for further BCI development, in particular for independent use at home. While reliability and learnability were judged to be very good, the EEG cap, system operability and speed clearly require substantial improvement.

**Effectiveness**

Performance was high in the four tasks and always above 70% accuracy, thereby fulfilling the criterion for satisfactory communication. This result is supported by the fact that users were rather satisfied with the effectiveness of the device and — except for low speed — did not mention problems with performance levels. For further development of BCI it is important to take into account that two users showed highest accuracies in the copy spelling and lowest in the internet task and that P300 amplitudes were higher in the spelling than the email and internet tasks. This may be due to the larger matrix size in the spelling tasks. It has been reported that P300 amplitudes increase with lower probability of the target item. The small size of the stimulation dot might also have contributed to lower accuracies in the tasks with “single dot” stimulation (email and internet) as compared to flashing of dots in rows and columns. The statement of two users about the small size of the dots supported this assumption. For this reason, the next prototype will provide the possibility to change the dot size individually or to choose an “inverse” stimulation where the icons available for selection will flash and be depicted in inverse colors.

**Efficiency**

With 4.03-8.57 bits/min. ITRs in the four tasks were comparable to other P300 BCI studies with end-users. All users indicated that improvement of speed was crucial for use in daily life. While for two users a two- to four-fold improvement of speed would be satisfying, user C indicated higher requirements and compared the BCI with his current AT where he achieved 90 characters per minute; a speed which constitutes a challenge for BCI as electro-physiological responses may be slower or require averaging.

Measures with the NASA TLX identified that for three users (A, C, D) subjective workload was about equal in the different tasks. Main sources of workload were “mental demand” (“mental and perceptual activity required”) and “temporal demand” (“time pressure felt due to the pace of the task”), but “effort” (“how hard did you have to work to accomplish your level of performance”) was rated low in all tasks by three users. Interestingly, user B who was a novice had expected that controlling a P300 BCI would be “very exhausting.” His subjective workload including “effort” declined with each task and he stated that his concerns had “not become true.” In addition, it proved that he learned to use the BCI QW-device very fast: in session two he indicated “learnability” in the extended QUEST 2.0 as one of the three most important aspects of an AT. Already in sessions three and four this aspect was no longer mentioned. This is an encouraging result as it shows that learning to use even a more complicated P300 BCI controlled device, such as an internet browser, can be achieved in a short time.

**Satisfaction**

In general, end-users and AT-experts were quite satisfied with the device. AT-experts stated that the device with its different functionalities was a clear improvement compared to the classic P300 speller. Users indicated that it compensated for deficits in motor system and speech and proved to be “usable.” Both, experts and users requested a better solution for the EEG-cap. Therefore, a solution with a fast, reliable and simple setup and without cables or the need of gel is urgently needed. New dry or water-based electrodes and wireless signal transmission may provide such solutions. Improvement in speed was requested, and users and experts asked for a compact hardware solution.

A further requirement was a simple setup procedure for the software. AT-experts also indicated the need of training for adjustment. Simple procedures and training that is “ease of use” are essential needs of caregivers so that BCI use does not add to their burden.

Finally, users suggested an interface that integrates BCI with other input channels. These so-called hybrid BCIs are in the focus of ongoing work of different research groups.

**CONCLUSION**

For the first time a P300 BCI was integrated with commercial AT-software that offers a whole set of communication and environmental control functions and its usability was tested with several tasks in a standardized evaluation procedure by potential end-users in a real world environment. The results revealed (1), high performance levels and (2), ITRs comparable to other P300 BCI studies with end-users; (3), that control of the BCI was not tiring for the users and (4), that learnability of the complex BCI controlled AT-software was high. (5), Users and AT-experts were quite satisfied with the device. However, none of them could imagine the use of the device in daily life without improvements. Main obstacles were the EEG-cap and the low speed. (6), Future evaluation of BCI as an additional channel to access AT-devices requires the inclusion of caregivers, family members, and personal assistants as this user group might have other requirements that are important for use in daily life.

The successful integration of a BCI with the QW-software shows that it is possible to provide BCI users with AT-solutions that allow them to enjoy the benefits of a broad range of ICT-functions. First steps toward use under real world conditions have been made, but further improvement is needed.

**ACKNOWLEDGMENT**

This work is supported by the European ICT Programme Project FP7-224631(TOBI). This paper only reflects the authors’ views and funding agencies are not liable for any use that may be made of the information contained herein.

**DISCLOSURE AND CONFLICT OF INTEREST**

REFERENCES


