



# HHS Public Access

Author manuscript

*Brain Comput Interfaces (Abingdon)*. Author manuscript; available in PMC 2017 January 01.

Published in final edited form as:

*Brain Comput Interfaces (Abingdon)*. 2016 January 1; 3(1): 47–58.

## SOLICITING BCI USER EXPERIENCE FEEDBACK FROM PEOPLE WITH SEVERE SPEECH AND PHYSICAL IMPAIRMENTS

Betts Peters<sup>a</sup>, Aimee Mooney<sup>a</sup>, Barry Oken<sup>b</sup>, and Melanie Fried-Oken<sup>a,c</sup>

<sup>a</sup>Institute on Development & Disability, Oregon Health & Science University, Portland, OR, USA

<sup>b</sup>Departments of Neurology, Behavioral Neuroscience, and Biomedical Engineering, Oregon Health & Science University, Portland, OR, USA

<sup>c</sup>Departments of Neurology, Pediatrics, Biomedical Engineering, and Otolaryngology, Oregon Health & Science University, Portland, OR, USA

### Abstract

Brain-computer interface (BCI) researchers have shown increasing interest in soliciting user experience (UX) feedback, but the severe speech and physical impairments (SSPI) of potential users create barriers to effective implementation with existing feedback instruments. This article describes augmentative and alternative communication (AAC)-based techniques for obtaining feedback from this population, and presents results from administration of a modified questionnaire to 12 individuals with SSPI after trials with a BCI spelling system. The proposed techniques facilitated successful questionnaire completion and provision of narrative feedback for all participants. Questionnaire administration required less than five minutes and minimal effort from participants. Results indicated that individual users may have very different reactions to the same system, and that ratings of workload and comfort provide important information not available through objective performance measures. People with SSPI are critical stakeholders in the future development of BCI, and appropriate adaptation of feedback questionnaires and administration techniques allows them to participate in shaping this assistive technology.

### Keywords

Brain-Computer Interfaces; User Feedback; Communication Aids for Disabled; Patient Outcome Assessment; Quadriplegia; Assistive Technology; Augmentative and Alternative Communication

---

Corresponding author: Dr. Melanie Fried-Oken, Oregon Health & Science University, Institute on Development & Disability, Department of Neurology, Pediatrics, Biomedical Engineering, Otolaryngology, 707 SW Gaines Street, Room 4224, Portland, OR USA 97239, (503) 494-7587, [friedm@ohsu.edu](mailto:friedm@ohsu.edu).  
Aimee Mooney, Oregon Health & Science University, Institute on Development & Disability, 707 SW Gaines Street, Room 1290, Portland, OR USA 97239, (503) 494-2291, [mooneya@ohsu.edu](mailto:mooneya@ohsu.edu)  
Betts Peters, Oregon Health & Science University, Institute on Development & Disability, 707 SW Gaines Street, Room 1290, Portland, OR USA 97239, (503) 494-2732, [petersbe@ohsu.edu](mailto:petersbe@ohsu.edu)  
Barry Oken, Oregon Health & Science University, Departments of Neurology, Behavioral, Neuroscience, Biomedical Engineering, 3181 SW Sam Jackson Park Road, CR-120, Portland, OR USA 97239, (503) 494-8873, [oken@ohsu.edu](mailto:oken@ohsu.edu)

## 1. Introduction

Brain-computer interface (BCI) is a relatively new assistive technology (AT) which shows great promise as an access method for people with severe speech and physical impairment (SSPI). Without relying on neuromuscular activity, BCI users can control communication systems, power wheelchairs, prosthetic limbs, and other AT using only their brainwaves.[1–3] Unfortunately, people with SSPI who may benefit from BCI use have largely been excluded from the research and development process. In a review of BCI systems for people with disabilities, Pasqualotto, Federici, and Belardinelli noted that few research groups had followed a user-centered approach.[4] BCI research papers often focus on engineering developments and describe trials with healthy users; among the 186 abstracts presented at the 2013 International BCI Meeting, only 31 included participants with disabilities.[5] The time has come to involve potential users in decisions about system design and eventual clinical implementation.[6] Including people with SSPI as study participants is a necessary step, but it is not sufficient. These individuals, who have the most to gain from this technology and the most to lose if it does not meet their needs, must be asked for their opinions and given the opportunity to shape the development of BCI.

Individuals with SSPI present with a range of diagnoses that result from neurodegenerative diseases (such as amyotrophic lateral sclerosis, muscular dystrophy, or Parkinson's Plus disorders) or chronic conditions (including traumatic brain injury, brainstem cerebrovascular accident, and spinal cord injury). Commonly, individuals who could benefit from BCI have been described according to their loss of function, with conditions such as locked-in syndrome, complete motor paralysis, tetraplegia, or motor disabilities. Those with neurodevelopmental disorders that interfere with speech and physical function, such as severe cerebral palsy, or rare neurodegenerative disease, such as Friedrich's ataxia, should also be included in this population and given the opportunity to participate in BCI research.

Several recent studies have queried potential users about their opinions and priorities regarding the future development of BCI technology.[7–11] Researchers have shown increased interest in soliciting user feedback on workload, satisfaction, and other aspects of the user experience (UX) for specific BCI systems, after years of focusing on objective outcome measures such as classification accuracy and selection speed.[12] Lorenz and colleagues called for a holistic approach to evaluation of the UX for BCI, including both pragmatic (usability-related) and hedonic (pleasure-related) UX attributes.[13] In addition to calculating objective task performance measures such as selection accuracy and selection duration, they asked participants to complete the NASA Task Load Index (TLX) [14] and the User Experience Questionnaire (UEQ) [15] after using several BCI systems. However, their participants were all able-bodied individuals. Another BCI research group has used the TLX, a modified version of the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST),[16] and the Assistive Technology Device Predisposition Assessment (ATDPA) [17] with people with disabilities,[18–20] but included only those who “could speak or use ‘conventional’ AT for communication and give clear and non-ambiguous feedback”.[18,p. 237] They described no details of how these instruments were administered or how responses were obtained. Felton and colleagues used a computer-based version of the TLX to evaluate workload during BCI training for users both with and without disabilities.[21]

However, due to unspecified “difficulties in administering the TLX battery”, [21,p.528] they obtained TLX responses from only seven of 12 participants with disabilities. It is clear that the use of UX questionnaires in their original forms leads to the exclusion of some people with SSPI from the conversation about the development of BCI technology. Other existing instruments for measuring outcomes related to rehabilitation and AT may be similarly inadequate for soliciting the opinions of this population. People with SSPI, despite their frequent use of AT and rehabilitation products and services, are often unable to give their opinions on these interventions, potentially leading to frustration and technology abandonment. [22, 23]

The difficulty of communicating with research participants with SSPI has been addressed in the context of ethical concerns, such as obtaining informed consent, [24, 25] and interview-based qualitative research. [10] However, there has been little discussion within the BCI community about communication methods for facilitating UX feedback. Concepts and techniques from the field of augmentative and alternative communication (AAC) may be useful in adapting feedback scales for use with people with SSPI. Respondents who are unable to verbally produce an answer or mark an X on a sheet of paper, as required by most instruments, can communicate their answers in other ways with the help of high-tech or low-tech AAC. [26–28] The use of augmented input, [27] with questions and response options presented in both auditory and visual formats, can help overcome obstacles related to visual, hearing, or language comprehension impairments. Question presentation may be modified to allow the use of nonverbal signals as a response modality, with or without assistance from the examiner. In one BCI study, for example, researchers administered the TLX to participants with SSPI by moving a finger along each rating scale until a participant gave a predetermined signal to indicate the desired rating. [29] Finally, non-essential questions and steps can be omitted where possible to reduce administration time and respondent fatigue.

In order to effectively assess the UX for BCI and other AT, questionnaires must be appropriate not only for the respondent, but for the task being studied. The TLX was initially designed by NASA for use in evaluating pilot workload during aviation tasks, but has become increasingly common in diverse fields including medicine and human-computer interaction (HCI). [30] After a literature search on workload assessment in human factors research revealed an “explosive” increase in use of the TLX, de Winter [31] proposed that its popularity was due to a Matthew effect (“the rich get richer” [32,p.62]). In essence, he suggested that the TLX grows in popularity primarily because it is already popular, and has now become the “obvious choice available to researchers and practitioners”. [31,p.293] While this idea in itself is not a criticism of the TLX, it may serve as a cautionary tale for the fields of rehabilitation and AT, including BCI. Work by Bierton and Bates [33, 34] found that the 20-interval, partially labeled scale used in the TLX did not allow respondents to easily discriminate among levels of difficulty for HCI tasks, indicating that the TLX may be unsuitable for the assessment of workload in this area. HCI is an important aspect of many modern ATs, including BCI systems, speech-generating devices (SGDs), and power wheelchairs, which should raise red flags about the widespread use of the TLX for evaluation of workload related to these devices.

Modifications have been proposed to improve the discriminability of the TLX for HCI and AT applications. Bierton and Bates [33] compared fully and partially labeled scales with a variety of intervals and found that a seven-point, fully labeled scale allowed for the clearest discrimination among different levels of HCI task difficulty. They also identified optimal quantifiers for use in labeling scaled response options.[33] Bates applied these results to create a modified version of the TLX, using a seven-point, fully labeled response scale for each question.[34] He cited previous work showing that the weighted scoring of the TLX is unnecessary and that the Effort component is redundant and inappropriate for use in evaluating HCI tasks, and omitted both of these elements from his version to allow for simpler and faster administration. In addition to the modified TLX for evaluating workload, Bates created similarly structured questionnaires on comfort and ease of use. The three resulting instruments (Workload Assessment Questionnaire, Comfort Assessment Questionnaire, and Ease of Use Questionnaire) were validated for use in assessing head- and eye-controlled HCI tasks.

In this article, we describe the adaptation and administration of Bates' questionnaires for evaluation of a BCI typing system for people with SSPI. We propose modifications to make the questionnaires more suitable for both the BCI task and the target population, including the use of AAC techniques for administration and response collection. Finally, we present results from the administration of the modified questionnaire to individuals with SSPI after trials with the RSVP Keyboard™ BCI typing system.[35, 36]

## 2. Methods

### 2.1 Participants

Participants included 12 adults with SSPI (10 men, 2 women) who were part of an ongoing research study on the RSVP Keyboard™ BCI system. Their mean age  $\pm$  standard deviation (SD) was  $51.8 \pm 12.1$  years. Information about their diagnoses, yes/no responses, and primary communication methods is displayed in Table 1. All participants presented with tetraplegia or tetraparesis, but had at least some reliable muscle movement (e.g. eye blink or muscle twitch) for yes/no communication. Nine presented with severe dysarthria or anarthria, while three had intelligible speech in conversation (though one had respiratory and positioning challenges that prevented him from using speech to meet all of his communication needs throughout the day). Two participants with severe dysarthria continued to use speech for specific purposes such as yes/no responses or short, simple phrases. Ten participants used high-tech and/or low-tech AAC for communication beyond yes/no responses.[26–28] Individuals with total LIS who had no reliable means for yes/no communication were excluded. This study was approved by the university Institutional Review Board and all participants provided informed consent.

### 2.2 Questionnaire Revision and Adaptation

Due to their high discriminability, simplified administration and scoring, and success in evaluating HCI tasks, Bates' Workload Assessment, Comfort Assessment, and Ease of Use Questionnaires [34] formed the basis for our instrument. The questionnaires were revised and adapted for use with people with SSPI, and for evaluation of the BCI tasks. We added a

question on overall satisfaction with the device, and omitted several questions that were deemed redundant, unlikely to elicit useful information, or unrelated to the RSVP Keyboard™ tasks. Questions about comfort were rephrased in an attempt to focus on discomfort specifically related to BCI use rather than general discomfort. For example, “Do you have headache pain of any kind?” became “*After using the system*, do you *now* have headache pain of any kind?” To reduce confusion among physical comfort and mental or emotional comfort, the overall comfort question was changed to, “Overall, during the task did using the system make you feel *physically* comfortable or uncomfortable?” Bates’ original seven-point, fully labeled rating scales were maintained, and a seven-point scale ranging from “1: Extremely satisfied” to “7: Extremely unsatisfied” was created for the new overall satisfaction question. The resulting questions were combined into a single questionnaire with four sections: Workload, Comfort, Ease of Use, and Overall Satisfaction (see Appendix). Due to space considerations, only endpoint labels were displayed on the data collection form, but all scales were fully labeled in the visual analogues seen by participants, described below.

Augmented input (also known as multimodal communication) [27] was incorporated into the administration procedures to improve communication with respondents with potential visual, hearing, or language comprehension impairments. To allow presentation of questions in both auditory and large-print visual formats, a visual analogue was created for each question. Each visual analogue consisted of a standard 8.5x11” page in landscape orientation, with the question printed at the top and a fully labeled Likert response scale below (see Figure 1). All text was in black letters on a white background, with a minimum font size of 36-point. Visual analogues were displayed either as printed, laminated 8.5x11” sheets, or as full-screen PDF files on a 17” laptop monitor, resulting in a displayed size of approximately 8x10.5”. Place figure 1 about here

### 2.3 BCI Tasks

Participants completed one session with the RSVP Keyboard™, a noninvasive, electroencephalography (EEG)-based BCI designed as a communication access method for people with SSPI.[35, 36] All data collection sessions took place at participants’ residences. Two researchers were present at each session; one focused primarily on managing the BCI hardware and software, allowing the other to concentrate on interacting with the participant, explaining the tasks, answering questions, and soliciting feedback. BCI tasks included two to four calibration sequences and a copy-spelling task with five levels of difficulty. These tasks, as well as procedures and equipment for EEG recording, are described in detail in previous work by Oken and colleagues.[36]

The primary dependent variable for the BCI tasks was the highest area under the curve (AUC) obtained during calibration. AUC is a measure of system classifier accuracy based on true positive versus false positive rate for the calibration target versus non-target classification. Possible AUC scores range from 0 to 1, with 0.5 indicating accuracy at chance level and higher scores indicating a better likelihood of success in typing with the RSVP Keyboard™. For each participant, the calibration sequence with the highest AUC score was used to create a system classifier for the copy-spelling task. The number of calibration

sequences was determined by the performance, preference, and energy level of individual participants.

Copy-spelling performance, as determined by the highest level completed on the copy-spelling task, served as an additional dependent variable. Levels of difficulty in the copy-spelling task were determined by the selection of target phrases which provided varying amounts of support from the RSVP Keyboard™'s integrated language model (LM). The LM assigns a probability to each character based on the five previously typed characters, and the Keyboard selects the character to be typed based on a Bayesian fusion of LM and EEG evidence. At level 1 of the copy-spelling task, users copied common words in common contexts, with all target characters being assigned high LM probabilities. At each successive level, the LM assigned lower probabilities to target characters, requiring stronger EEG evidence from the user and increasing the difficulty of the task. Level 5 is the most difficult level, with the lowest probabilities assigned to target characters. For example, in Level 1, the participant was asked to copy the word NOT in the sentence I DO NOT AGREE. In Level 5, the word FLEA was copied from the phrase A CAN OF FLEA POWDER. Each participant attempted the copy-spelling task once, ending the task when he or she successfully completed level 5, failed to complete a lower level, or asked to stop the session.

## 2.4 Questionnaire Administration

Immediately after finishing the copy-spelling task, participants completed the feedback questionnaire. For each question, the researcher displayed the corresponding visual analogue (on a printed page or laptop screen) at the participant's eye level, approximately two to three feet away, and confirmed via yes/no response that the participant could read the text well in that position. After reading each question aloud, the researcher then read aloud the number and label of each response option, while pointing to the matching text on the visual analogue. Participants with intelligible speech spoke the number associated with their choice, while those with speech impairments used partner-assisted scanning [27] as a response modality. Before beginning partner-assisted scanning, the participant and researcher first agreed on a designated response signal, such as a vocalization, eye blink, or other small movement. The researcher read and pointed to each response option in turn, keeping the rate of presentation constant and pausing after each option to allow adequate time for the participant to respond. For each option, the participant either gave the signal to indicate a selection, or did nothing to indicate that the researcher should proceed to the next option. After each response selection, the researcher confirmed the participant's choice with a yes/no question, e.g. "Is your answer '3'?"

## 2.5 Narrative Feedback

In addition to completing the questionnaire, participants were encouraged to ask questions and provide narrative feedback about the RSVP Keyboard™ during data collection sessions. Natural opportunities for soliciting and recording this type of feedback occurred in each session (for example, during system setup and take-down, system adjustments between tasks, or AUC calculation). One research team member would focus on communicating with the participant while another managed the system hardware and software. After completing the quantitative questionnaire, participants were asked if they wished to provide any

explanation or additional detail related to their responses. Team members ensured that the participant had access to his or her primary communication method (e.g. SGD) when narrative feedback was requested. Participants could choose to have a family member or care provider present to assist with communication, if desired. For those who used AAC, preferred communication procedures were established at the beginning of the session. For example, participants using SGDs were asked whether they wished to use the device's synthesized speech output or have an investigator read the text on the screen, and those using low-tech AAC (e.g. partner-assisted scanning or a letter board) were asked whether the investigator should guess words as they were spelled or wait for the participant to complete the entire message. Participants with email access were encouraged to follow up using that medium if they had additional comments or questions in the days following a session.

### 3. Results

#### 3.1 Questionnaire

All twelve participants completed the questionnaire, using either speech or partner-assisted scanning with signals involving minimal muscle movement. Signal modalities included vocalizations, eye blinks, and hand, head, or eye movements. All participants indicated understanding (via yes/no response) of partner-assisted scanning, and responses were generally easy for researchers to interpret. Any confusion or misunderstanding was easily resolved with yes/no questions. The two researchers who administered the questionnaire estimated that administration required less than five minutes for all participants.

Table 2 displays RSVP Keyboard™ task performance and selected questionnaire responses for all participants. The table reports each participant's highest AUC score, as only the calibration session with the best score was used to create a classifier for the copy-spelling task. The mean highest AUC  $\pm$  SD was  $.699 \pm .080$ , and the mean  $\pm$  SD for highest copy-spelling level completed was  $1.3 \pm 1.4$ . Only one participant completed all five levels of the copy-spelling task. Participant 4 had a high AUC (.85), but was unable to pass the first copy-spell level. Significant electrical interference was observed via real-time monitoring of EEG signals during this participant's session, and may have affected system performance more during the copy-spelling task than during calibration.

Table 3 presents the range, mean, and SD of responses to each question for all 12 participants. Ranges reflect the highest and lowest response numbers received for each question, and mean rating label is the rating scale label associated with the integer nearest to the mean response number. For example, participants' ratings of the time pressure associated with RSVP Keyboard™ use ranged from 1 ("extremely low") to 5 ("somewhat high"). The mean response for that question was 2.75, which is rounded up to 3 to find the mean rating label ("somewhat low"). One participant responded that his frustration was extremely low (1) for the copy-spelling task, but extremely high (7) for the calibration task. These responses were averaged (4) to get a rating of his overall frustration during the session. Response data showed that, on average, participants found the RSVP Keyboard™ BCI to require "somewhat low" physical effort, but "somewhat high" mental effort. They reported that it required only "a little" hard work, and caused little to no discomfort of any kind. Participants rated the system as "neither accurate nor inaccurate", "somewhat fast", and

“somewhat satisfying” overall. All questions except the one related to headache discomfort elicited a variety of responses, with ranges spanning four to seven points on a seven-point scale.

Spearman rank-order correlations were used to examine the relationships between objective measures of BCI performance (AUC and highest copy-spelling level completed) and participants’ ratings of overall workload, typing accuracy, and overall system satisfaction. There was no significant correlation between overall workload and AUC ( $r_s = -.343$ ,  $p = .138$ ), while the correlation between overall workload and highest copy-spelling level approached significance ( $r_s = .426$ ,  $p = .084$ ). Better ratings of typing accuracy were significantly correlated with completion of higher copy-spelling levels ( $r_s = -.507$ ,  $p = .046$ ), and approached a significant correlation with higher AUC scores ( $r_s = -.414$ ,  $p = .090$ ). Greater overall satisfaction showed a significant correlation with both higher AUC scores ( $r_s = -.568$ ,  $p = .027$ ) and completion of higher copy-spelling levels ( $r_s = -.562$ ,  $p = .029$ ).

### 3.2 Narrative Feedback

All participants asked questions and provided narrative feedback during data collection sessions, using their primary communication method(s) (see Table 1). Only the participant who used partner-assisted scanning requested the presence of a care provider to assist with communication of narrative feedback. One participant followed up with additional comments via email the next day. Participants commented on how they completed the task (e.g. “Before [during a trial at an earlier visit], I found myself reading each letter. This time I tried to only focus on [the target character].”), provided additional context regarding their performance (e.g. “I was more reclined in the chair during the first calibration, so I was too relaxed.”), and gave their opinions about the user interface and stimulus presentation (e.g. “It’s easier to respond to letters at the beginning or end of a [sequence].”; “I think I would be more accurate [with slower RSVP presentation].”). One participant provided feedback about the questionnaire itself, calling the overall comfort question “irrelevant”, since “using a BCI system would never make someone *more* comfortable.”

## 4. Discussion

Soliciting BCI UX feedback from participants with SSPI can be challenging, as their speech and physical impairments can prevent effective use of existing tools. Yet these individuals are crucial members of the BCI development team. In fact, they are experts in living with disabilities, and are vital informants on how to make BCI technology meet their needs. Results of this study indicate that people with SSPI are able to provide valuable feedback about AT interventions, including BCI systems. Modified administration techniques based on AAC concepts, including partner-assisted scanning as a response modality and simultaneous visual and auditory presentation of questions and response options, were helpful in ensuring effective communication and response collection with this population. Administration of the adapted questionnaire was fast and easy, and required minimal physical exertion by participants with SSPI. The presence of two researchers at each data collection session was crucial, as it ensured that one researcher could devote sufficient time and attention to soliciting participant feedback.

The range of questionnaire responses was often large, with responses for some questions spanning the full range of options (1 through 7). This indicates a high degree of variability in individual users' responses to the same AT. These differences may be related to factors unique to each user, such as physical function, health status, attitudes or expectations about the AT, or emotional state. Questionnaire responses could be used to compare potential AT interventions for an individual with SSPI, to ensure provision of the most appropriate and effective option and reduce technology abandonment. Such subjective assessments might also be used to explore AT adjustment or customization options to minimize workload and maximize comfort and ease of use.

User ratings of workload may or may not correlate with performance on RSVP Keyboard™ tasks, but they provide new information not available through objective performance measures. This supports the notion that subjective measures of UX should be considered along with objective performance measures when choosing an AT. For example, if two BCI typing systems offered similar speed and accuracy, the system with lower subjective workload would be preferred. Some users may choose a low-workload system with slightly slower typing speed over a faster system that requires a greater workload. Use of the TLX as a measure of subjective workload in AT research and clinical implementation is a positive step. Adaptation of the TLX following Bates's methods and those proposed above, including a modified response scale and AAC-inspired administration techniques, will make this tool more suitable for use with people with SSPI in evaluating HCI-related AT tasks.

The Ease of Use section of the questionnaire may have been misnamed, as its component questions represent subjective measures of system performance rather than ease of use. The correlation between better copy-spelling performance and more favorable ratings of typing accuracy shows that our participants had a realistic view of their performance on the RSVP Keyboard™ tasks. These results indicate that subjective measures of system performance tend to replicate data obtained through objective measures of success on a given task.

Correlations of greater overall satisfaction with higher AUC and better copy-spelling performance show that, unsurprisingly, RSVP Keyboard™ users were more satisfied with the system when they performed well. This suggests that a subjective rating of overall satisfaction is not very informative, at least in the case of evaluation of a single AT, as in the present study. A measure of overall satisfaction may prove more useful for comparing two or more different AT options or tasks, though a multiple-question satisfaction assessment such as the QUEST [16] would provide more detailed information about an individual's reasons for preferring one AT over another.

This was our first attempt at adapting a UX questionnaire for users of the RSVP Keyboard™. Based on the results presented here, as well as discussion among the authors and feedback from a research team member with SSPI, we plan to make several modifications to the questionnaire in the future. For example, some participants found the questions in the Comfort section unclear. One man rated his level of eye fatigue/strain/pain as a 2 (scarcely tired, strained, or painful), but then said it was "the same as before" beginning the task. In this case, his discomfort was unrelated to his use of the RSVP Keyboard™. We will revise the comfort questions to ask more clearly and specifically about *increases* in discomfort

during the task. For instance, the headache question might be rewritten as, “Did you experience any increase in headache pain while using the system?”, with response options ranging from “1: No pain/no increase in pain” to “7: Extreme increase in pain”. One participant (also a member of the research team) pointed out that the wording of the overall comfort question was strange, as using a BCI is unlikely to make someone feel physically comfortable. In addition, the bipolar response scale for this question contrasted with the unipolar scale used for the other comfort questions, which could potentially lead to confusion when giving responses or scoring the questionnaire. We plan to re-word the overall comfort question to ask only whether system use made the respondent feel uncomfortable, and change its response scale from bipolar to unipolar, with options ranging from “1: Not at all uncomfortable” to “7: Extremely uncomfortable”. Finally, we will omit the misnamed and uninformative Ease of Use section from the revised version of the questionnaire, and rely on objective measures of system performance. Instead, we may ask participants to rate their satisfaction with typing speed and accuracy. Some potential BCI users, particularly those whose other communication options are limited due to SSPI, may be satisfied even with low typing rates, or may expect high levels of accuracy.[6]

The questionnaire modifications and adapted administration techniques described here may be beneficial for evaluating UX for other AT, including other BCI systems. Questions in the comfort section may be modified, added, or omitted to fit the tasks being evaluated. For example, assessment of a BCI system using auditory or tactile stimuli would require questions about discomfort related to those stimuli. Other UX and healthcare outcomes assessment tools may be modified for use with people with SSPI following the techniques described here. The QUEST [16] features a fully labeled, 5-point response scale that could easily be adapted with visual analogues, augmented input, and partner-assisted scanning. The UEQ [15] and item banks from the Patient-Reported Outcomes Measurement Information System [37] would also lend themselves easily to these modifications. Regardless of the tool implemented, researchers should adhere to recommended guidelines for soliciting UX feedback from individuals with SSPI, available in Figure 2. More detailed responses and opinions may be obtained using the interview procedures recommended by Andresen and colleagues.[10]

#### 4.1 Limitations

The reported analysis is limited by a small sample size, which makes it difficult to interpret the Spearman correlation results. A larger sample of individuals with a variety of diagnoses, ages, and ability levels is needed to determine the validity of these questionnaire modification and administration techniques for obtaining feedback on AT interventions from people with SSPI. AUC and copy-spelling performance were fairly low for the sample described in this study, which may have affected participants’ UX. Although electrical interference was only observed in the EEG signals of a single participant, it may have affected his copy-spelling performance. Another regrettable, though unavoidable, limitation was the exclusion of people with total LIS from the study. While the AAC techniques described above will allow questionnaire administration to anyone with a reliable yes/no response, they are inadequate for use with people with total LIS. Individuals who have no voluntary motor function, and therefore no way to communicate a yes/no response, continue

to be excluded from decisions regarding AT and other healthcare options. BCI technology offers hope for effective communication by people with total LIS. As it continues to develop and improve, it may eventually serve as an access method allowing people with total LIS to complete UX questionnaires and other healthcare outcomes instruments.

## 5. Conclusions

These questionnaire modifications and administration techniques are specifically designed for obtaining UX feedback from people with SSPI for HCI-related AT tasks. The questionnaire used in this study was easy to administer in five minutes or less, and required minimal effort on the part of respondents with severe disabilities. Modified administration techniques including augmented input for question and response presentation and partner-assisted scanning as a response modality allowed use of the questionnaire with any individual who had a reliable yes/no response. Future versions of the questionnaire will include minor revisions to question wording, omit the Ease of Use section, and request participant ratings of satisfaction with system performance.

When selecting an instrument for the assessment of UX with AT, researchers and clinicians should consider whether it is 1) appropriate for evaluation of a particular device or task, and 2) administrable to the target population, regardless of communication or physical abilities. Although it is desirable to use standardized tools to allow comparison among different studies, “everyone else is using it” should not be sufficient justification for choosing an instrument. Further research is needed into efficient, valid, and practical methods for soliciting the opinions and input of individuals with SSPI on AT and healthcare options, including BCI. Even people who cannot speak should be able to make their voices heard.

## Acknowledgments

Financial support in part from NIH Grant #1R01DC009834. A portion of this publication was developed under a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant #H133E140026) to the Rehabilitation Engineering Research Center on Augmentative and Alternative Communication (RERC on AAC). NIDILRR is a Center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this publication do not necessarily represent the policy of NIDILRR, ACL, HHS, and you should not assume endorsement by the Federal Government. The authors wish to thank Dr. John Simpson for his insightful feedback and suggestions, and Amy Golinker for her assistance with preparation for publication.

## Biographies

**Betts Peters, M.A., CCC-SLP** is an ASHA-certified speech-language pathologist specializing in augmentative and alternative communication and assistive technology. She has been a Research Associate with REKNEW (Reclaiming Expressive Knowledge in Elders With communication impairments) Projects since 2012, working with Dr. Melanie Fried-Oken on research aiming to improve communication for individuals with severe speech and physical impairments. Her primary research focus is on the development and refinement of a brain-computer interface system that can be used for communication by people with locked-in syndrome. Before joining the REKNEW team, Betts worked as an assistive technology specialist for people with ALS (Lou Gehrig’s disease).

**Aimee Mooney, M.S., CCC-SLP** has practiced as a Speech and Language Pathologist for 24 years, specializing in adult survivors of neurological injury/illness and the application of cognitive rehabilitation with these populations. Her clinical research focuses on cognitive and communication changes in aging and dementia and investigate ways that Augmentative and Alternative Communication (AAC) support adults with degenerative neurological disease as they experience complex communication impairments. She is an Adjunct Faculty in the Speech and Hearing Sciences Department at Portland State University where she is a professor of Cognitive Rehabilitation and a clinical supervisor.

**Melanie Fried-Oken, Ph.D., CCC-SLP** is a certified speech-language pathologist, Professor of Neurology, Pediatrics, Biomedical Engineering, and Otolaryngology at the Oregon Health & Science University. As a leading international clinician and researcher in the field of Augmentative and Alternative Communication, she provides expertise about assistive technology for persons with acquired and developmental disabilities. She has a number of federal grants to research communication technology for persons with severe speech and physical impairments, including a translational research grant on BCI. She is a partner in the Rehabilitation Engineering Research Center on Communication Enhancement ([www.lerc-aac.com](http://www.lerc-aac.com)) and a practicing clinician in the Augmentative Communication Clinic at Oregon Health & Science University.

**Barry Oken, M.D., M.S.** received a B.A. degree in math from the University of Rochester and an M.D. degree from the Medical College of Wisconsin. Then he completed a residency in Neurology at Boston University Medical Center and fellowship in EEG and evoked potentials at Massachusetts General Hospital and Harvard Medical School. Since 1985, he has been a member of the faculty at Oregon Health & Science University, where he is currently Professor in the Departments of Neurology, Behavioral Neuroscience and Biomedical Engineering. He has had continuous grant funding from NIH since 1990. He also has a long interest in quantitative data analysis and recently obtained an M.S. in Systems Science from Portland State University to further his knowledge in that area.

## References

1. Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain-computer interfaces for communication and control. *Clin Neurophysiol.* 2002; 113(6):767–91. [PubMed: 12048038]
2. Akcakaya M, Peters B, Moghadamfalahi M, Mooney A, Orhan U, Oken B, Erdogmus D, Fried-Oken M. Noninvasive brain computer interfaces for augmentative and alternative communication. *IEEE Rev Biomed Eng.* 2014; 7:31–49. [PubMed: 24802700]
3. Daly JJ, Huggins JE. Brain-computer interface: Current and emerging rehabilitation applications. *Arch Phys Med Rehabil.* 2015; 96(3):S1–7. [PubMed: 25721542]
4. Pasqualotto E, Federici S, Belardinelli MO. Toward functioning and usable brain-computer interfaces (BCIs): A literature review. *Disabil Rehabil Assist Technol.* 2012; 7(2):89–103. [PubMed: 21967470]
5. Millán, JdR; Gao, S.; Müller-Putz, GR.; Wolpaw, JR.; Huggins, JE., editors. Proceedings of the fifth international brain-computer interface meeting: Defining the future; 2013 Jun 3–7; Pacific Grove, CA USA. 2013. Available from: [http://castor.tugraz.at/doku/BCIMeeting2013/BCIMeeting2013\\_all.pdf](http://castor.tugraz.at/doku/BCIMeeting2013/BCIMeeting2013_all.pdf)
6. Peters B, Bieker G, Heckman SM, Huggins JE, Wolf C, Zeitlin D, Fried-Oken M. Brain-computer interface users speak up: The virtual users' forum at the 2013 international brain-computer interface meeting. *Arch Phys Med Rehabil.* 2015; 96(3):S33–7. [PubMed: 25721545]

7. Huggins JE, Wren PA, Gruis KL. What would brain-computer interface users want? Opinions and priorities of potential users with amyotrophic lateral sclerosis. *Amyotrophic Lat Scler.* 2011; 12(5): 318–24.
8. Blain-Moraes S, Schaff R, Gruis KL, Huggins JE, Wren PA. Barriers to and mediators of brain-computer interface user acceptance: Focus group findings. *Ergonomics.* 2012; 55(5):516–25. [PubMed: 22455595]
9. Zickler C, Di Donna V, Kaiser V, Al-Khodairy A, Kleih S, Kübler A, Malavasi M, Mattia D, Mongardi S, Neuper C. BCI applications for people with disabilities: Defining user needs and user requirements. *Assistive Technology from Adapted Equipment to Inclusive Environments, AAATE. 25 Assistive Technology Research Series.* 2009:185–9.
10. Andresen EM, Fried-Oken M, Peters B, Patrick DL. Initial constructs for patient-centered outcome measures to evaluate brain-computer interfaces. *Disabil Rehabil Assist Technol.* 2015; 0:1–10. [PubMed: 25806719]
11. Huggins JE, Moinuddin AA, Chiodo AE, Wren PA. What would brain-computer interface users want: Opinions and priorities of potential users with spinal cord injury. *Arch Phys Med Rehabil.* 2015; 96(3 Suppl 1):S38–45. [PubMed: 25721546]
12. Plass-Oude Bos D, Gürkök H, Van de Laar B, Nijboer F, Nijholt A. User experience evaluation in BCI: Mind the gap! *Int J Bioelectromagn.* 2011; 13(1):48–9.
13. Lorenz R, Pascual J, Blankertz B, Vidaurre C. Towards a holistic assessment of the user experience with hybrid BCIs. *Journal of Neural Engineering.* 2014; 11(3):035007. [PubMed: 24835132]
14. Hart SG, Staveland LE. Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Human Mental Workload.* 1988; 1:139–83.
15. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. Springer. 2008; 5298:63–76.
16. Demers, L.; Weiss-Lambrou, R.; Ska, B. Quebec user evaluation of satisfaction with assistive technology (QUEST version 2.0) - an outcome measure for assistive technology devices. Webster, NY: Institute for Matching Person & Technology, Inc; 2000.
17. Scherer M, Jutai J, Fuhrer M, Demers L, Deruyter F. A framework for modelling the selection of assistive technology devices (ATDs). *Disabil Rehabil Assist Technol.* 2007; 2(1):1–8. [PubMed: 19263548]
18. Zickler C, Riccio A, Leotta F, Hillian-Tress S, Halder S, Holz E, Staiger-Sälzer P, Hoogerwerf E, Desideri L, Mattia D. A brain-computer interface as input channel for a standard assistive technology software. *Clinical EEG and Neuroscience.* 2011; 42(4):236–44. [PubMed: 22208121]
19. Kübler A, Zickler C, Holz E, Kaufmann T, Riccio A, Mattia D. Applying the user-centred design to evaluation of brain-computer interface controlled applications. *Biomed Tech (Berl).* 2013; 58(S1)
20. Zickler C, Halder S, Kleih SC, Herbert C, Kübler A. Brain painting: Usability testing according to the user-centered design in end users with severe motor paralysis. *Artif Intell Med.* 2013; 59(2): 99–110. [PubMed: 24080077]
21. Felton EA, Williams JC, Vanderheiden GC, Radwin RG. Mental workload during brain-computer interface training. *Ergonomics.* 2012; 55(5):526–37. [PubMed: 22506483]
22. Phillips B, Zhao H. Predictors of assistive technology abandonment. *Assistive Technology.* 1993; 5(1):36–45. [PubMed: 10171664]
23. Wessels R, Dijcks B, Soede M, Gelderblom G, De Witte L. Non-use of provided assistive technology devices, a literature overview. *Technol Disabil.* 2003; 15(4):231–8.
24. Haselager P, Vlek R, Hill J, Nijboer F. A note on ethical aspects of BCI. *Neural Networks.* 2009; 22(9):1352–7. [PubMed: 19616405]
25. Nijboer F, Clausen J, Allison BZ, Haselager P. The asilomar survey: Stakeholders' opinions on ethical issues related to brain-computer interfacing. *Neuroethics.* 2013; 6(3):541–78. [PubMed: 24273623]
26. Fager S, Beukelman DR, Fried-Oken M, Jakobs T, Baker J. Access interface strategies. *Assist Technol.* 2012; 24(1):25–33. [PubMed: 22590797]
27. Garrett, KL.; Happ, MB.; Costello, JM.; Fried-Oken, MB. AAC in the intensive care unit. In: Beukelman, DR.; Garrett, KL.; Yorkston, KM., editors. *Augmentative communication strategies*

for adults with acute or chronic medical conditions. Baltimore, MD: Paul H. Brookes Publishing Co; 2007. p. 17-57.

28. Fried-Oken M, Mooney A, Peters B. Supporting communication for patients with neurodegenerative disease. *NeuroRehabilitation*. in press.
29. Combaz A, Chatelle C, Robben A, Vanhoof G, Goeleven A, Thijs V, Van Hulle MM, Laureys S. A comparison of two spelling brain-computer interfaces based on visual P3 and SSVEP in locked-in syndrome. *PLoS One*. 2013; 8(9):e73691. [PubMed: 24086289]
30. Hart, SG. NASA-task load index (NASA-TLX); 20 years later. NASA-task load index (NASA-TLX); 20 years later. Proceedings of the human factors and ergonomics society annual meeting Sage Publications; 2006 Oct 16–20; San Francisco, CA USA. 2006.
31. De Winter J. Controversy in human factors constructs and the explosive use of the NASA-TLX: A measurement perspective. *Cogn Technol Work*. 2014; 16(3):289–97.
32. Merton RK. The Matthew effect in science. *Science*. 1968; 159:56–63.
33. Bierton R, Bates R. Experimental determination of optimal scales for usability questionnaire design. *Proceedings of Human Computer Interaction*. 2000; II:55–6.
34. Bates, REA. Enhancing the performance of eye and head mice: A validated assessment method and an investigation into the performance of eye and head based assistive technology pointing devices [dissertation]. Leicester (UK): De Montfort University; 2006.
35. Orhan, U.; Hild, KE., II; Erdogmus, D.; Roark, B.; Oken, B.; Fried-Oken, M. RSVP keyboard: An EEG based typing interface. *RSVP keyboard: An EEG based typing interface*. 2012 IEEE international conference on acoustics, speech, and signal processing, ICASSP; 2012; March 25–30; Kyoto, Japan. 2012. Sponsors: Inst. Electr. Electron. Eng. Signal Process. Soc.; Conference code: 93091
36. Oken BS, Orhan U, Roark B, Erdogmus D, Fowler A, Mooney A, Peters B, Miller M, Fried-Oken MB. Brain-computer interface with language model-electroencephalography fusion for locked-in syndrome. *Neurorehabil Neural Repair*. 2014 May; 28(4):387–94. [PubMed: 24370570]
37. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, Amtmann D, Bode R, Buysse D, Choi S. Initial adult health item banks and first wave testing of the patient-reported outcomes measurement information system (PROMIS™) network: 2005–2008. *J Clin Epidemiol*. 2010; 63(11):1179. [PubMed: 20685078]

## Appendix: Questionnaire data collection form

<b>Workload</b>								
1. How much <i>physical</i> effort or activity was required to operate the system?	Low						High	
		1	2	3	4	5	6	7
2. How much <i>mental</i> effort or concentration was required to operate the system?	Low							High
		1	2	3	4	5	6	7
3. How much <i>temporal</i> or time pressure did you feel under?	Low							High
		1	2	3	4	5	6	7
4. What level of <i>frustration</i> did you experience when using the system?	Low							High
		1	2	3	4	5	6	7
5. Overall, how hard did you have to <i>work</i> during the task?	Low							High
		1	2	3	4	5	6	7

<b>Comfort</b>							
1. After using the system, do you now have <i>headache</i> pain of any kind?	Not at all			Extremely			
	1	2	3	4	5	6	7
2. After using the system, do your <i>eyes</i> now feel tired, strained or painful?	Not at all			Extremely			
	1	2	3	4	5	6	7
3. After using the system, do your <i>facial muscles</i> now feel tired, strained or painful?	Not at all			Extremely			
	1	2	3	4	5	6	7
4. After using the system, does your <i>neck</i> now feel tired, stiff or painful?	Not at all			Extremely			
	1	2	3	4	5	6	7
5. Overall, during the task did using the system make you feel physically comfortable or uncomfortable?	Comfortable			Uncomfortable			
	1	2	3	4	5	6	7

<b>Ease of Use (skip if participant only completed Calibration)</b>							
1. Did you find that <i>letter selection</i> was accurate or inaccurate?	Accurate			Inaccurate			
	1	2	3	4	5	6	7
2. Did you find that the <i>speed of letter selection</i> was fast or slow?	Fast			Slow			
	1	2	3	4	5	6	7

<b>Overall Satisfaction</b>							
1. How satisfied are you with the system overall?	Satisfied			Unsatisfied			
	1	2	3	4	5	6	7

## How much *physical* effort or activity was required to operate the system?

7	Extremely high
6	Considerably high
5	Somewhat high
4	Neither high nor low
3	Somewhat low
2	Considerably low
1	Extremely low

**Figure 1.**  
Sample visual analogue.

- Use questionnaires with high discriminability and success in evaluating HCI tasks
- Omit unnecessary or inapplicable questions to reduce administration time and fatigue
- Avoid soliciting subjective ratings of selection accuracy, speed, or other outcomes that can be measured objectively
- Use fully labeled rating scales with a limited number of scale points (5-9 points is common)
- Present questions and response options in both visual and auditory formats to provide augmented input
- Use large-print visual analogues, positioned in the participant's line of sight, and ask the participant if he or she can read them
- Have at least two researchers present during each data collection session, with one focusing primarily on communicating with the participant and soliciting feedback
- Identify clear, reliable yes/no signals for each individual participant (work with a participant's family member, care provider, or speech-language pathologist if necessary)
- Ask questions in yes/no format whenever possible; "Did you like the second layout better?" can be answered with a yes/no response, but "Did you prefer the first layout or the second layout?" cannot
- Use partner-assisted scanning to obtain rating-scale responses from participants with speech impairments
- Ask for yes/no confirmation of responses, e.g. "Is your answer '3'?"
- Request narrative feedback during data collection sessions
- Encourage the participant to submit follow-up questions and comments via email when possible

**Figure 2.**  
Recommendations for soliciting UX feedback from individuals with SSPI.

**Table 1**

Participant diagnoses, yes/no responses, and primary communication methods.

Participant	Diagnosis*	Yes/no response	Primary communication method(s)*
1	LIS	Look up/look down	Partner-assisted scanning
2	ALS	Squint eyes/look side to side	Tablet-based SGD with eye tracking
3	CP	Nod head/shake head	Keyboard-based SGD with hand brace stylus
4	DMD	Speech (with impaired breath support)	Speech or tablet-based SGD with eye tracking
5	ALS	Move foot up and down/ move foot side to side	Tablet-based SGD with foot-controlled trackball
6	CP	Move arm up/move arm down	Keyboard-based SGD
7	SCA	Thumbs up/thumbs down	Letter board with finger pointing or keyboard-based SGD
8	ALS	Speech (dysarthric)	Speech or tablet-based SGD with head mouse
9	ALS	Speech (dysarthric)	Tablet-based SGD with head mouse
10	ALS	Speech	Speech
11	ALS	Speech	Speech
12	ALS	Blink once/blink twice	Tablet-based SGD with eye tracking or switch scanning

\* ALS = amyotrophic lateral sclerosis, CP = cerebral palsy, DMD = Duchenne muscular dystrophy, LIS = locked-in syndrome, SCA = spinocerebellar ataxia, SGD = speech-generating device

**Table 2**

BCI task performance and selected questionnaire responses.

Participant	Number of calibration sequences	Highest AUC	Highest copy-spelling level completed	Reason for stopping copy-spelling task	Overall workload	Overall comfort	Overall satisfaction
1	2	.78	2	Failed level 3	2	4	1
2	2	.66	2	Participant request (fatigue)	5	5	5
3	2	.62	1	Failed level 2	4	4	3
4	2	.85	0	Failed level 1*	1	1	3
5	3	.81	2	Failed level 3	4	4	3
6	2	.67	1	Participant request (time constraints)	2	2	2
7	4	.60	0	Failed level 1	5	5	6
8	3	.70	2	Participant request (time constraints)	2	2	2
9	2	.66	0	Failed level 1	3	3	3
10	3	.66	0	Failed level 1	1	3	5
11	2	.63	1	Failed level 2	3	4	3
12	2	.75	5	Completed all levels	6	4	2

\* Electrical interference was noted in the EEG signals during this visit, and may have affected this participant's copy-spell performance.

**Table 3**

Range, mean, and standard deviation (SD) of questionnaire responses (n=12).

	Range	Mean	SD	Mean rating label*
<b>Workload</b>				
Physical effort	1–7	2.83	2.08	3: Somewhat low
Mental effort	1–6	4.83	1.47	5: Somewhat high
Time pressure	1–5	2.75	1.87	3: Somewhat low
Frustration	1–7	3.75	1.91	4: Neither high nor low
Overall workload	1–6	3.17	1.64	3: A little hard
<b>Comfort</b>				
Headache	1–1	1.00	.00	1: Not at all painful
Eyes	1–5	2.33	1.30	2: Scarcely tired, strained or painful
Facial muscles	1–4	1.58	1.17	2: Scarcely tired, strained, or painful
Neck	1–4	1.67	1.07	2: Scarcely tired, stiff, or painful
Overall comfort	1–5	3.42	1.24	3: Somewhat comfortable
<b>Ease of use</b>				
Accuracy	2–7	3.75	1.71	4: Neither accurate nor inaccurate
Speed	1–5	3.08	1.08	3: Somewhat fast
<b>Overall satisfaction</b>				
Overall satisfaction	1–6	3.17	1.47	3: Somewhat satisfied

\* Mean rating label is the rating scale label associated with the integer nearest to the mean.