RSVP Keyboard with Inquiry Preview: Mixed performance and user experience with an adaptive, multimodal typing interface combining EEG and switch input

Betts Peters[1], Basak Celik[2], Dylan Gaines[3], Deirdre Galvin-McLaughlin[1,4], Tales Imbiriba[2], Michelle Kinsella[1], Daniel Klee[5], Matthew Lawhead[6], Tab Memmott[1,5], Niklas Smedemark-Margulies[2], Jack Wiedrick[7], Deniz Erdogmus[2], Barry Oken[5], Keith Vertanen[3], and Melanie Fried-Oken[1]

[1] Institute on Development & Disability, Department of Pediatrics, Oregon Health & Science University, Portland, Oregon, USA

[2] Department of Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts, USA

[3] Department of Computer Science, Michigan Technological University, Houghton, Michigan, USA

[4] Speech-Language and Hearing Sciences Department, Boston University, Boston, Massachusetts, USA

[5] Department of Neurology, Oregon Health & Science University, Portland, Oregon, USA

[6] Oregon Clinical and Translational Research Institute, Oregon Health & Science University, Portland, Oregon, USA

[7] Biostatistics & Design Program, Oregon Health & Science University, Portland, Oregon, USA

Corresponding author: Betts Peters, petersbe@ohsu.edu

ORCIDs:

Betts Peters: 0000-0002-7019-079X

Basak Celik: 0000-0002-0912-5243

Dylan Gaines: 0000-0002-2747-7680

Deirdre Galvin-McLaughlin: 0000-0002-7596-9209

Tales Imbiriba: 0000-0002-2626-2039

Michelle Kinsella

Daniel Klee: 0000-0003-2992-7662

Matthew Lawhead: 0000-0003-0736-3587

Tab Memmott: 0000-0001-6143-5057

Niklas Smedemark-Margulies: 0000-0002-4364-0273

Jack Wiedrick

Deniz Erdogmus: 0000-0002-1114-3539

Barry Oken: 0000-0002-3781-4273

Keith Vertanen: 0000-0002-7814-2450

Melanie Fried-Oken: 0000-0002-5075-8392

Abstract

Objective. The RSVP Keyboard is a non-implantable, event-related potential-based brain-computer interface (BCI) system designed to support communication access for people with severe speech and physical impairments. Here we introduce Inquiry Preview, a new RSVP Keyboard interface incorporating switch input for users with some voluntary motor function, and describe its effects on typing performance and other outcomes. **Approach.** Four individuals with disabilities participated in the collaborative design of possible switch input applications for the RSVP Keyboard, leading to the development of Inquiry Preview and a method of fusing switch input with language model and electroencephalography (EEG) evidence for typing. Twenty-four participants without disabilities and one potential end user with incomplete locked-in syndrome took part in two experiments investigating the effects of Inquiry Preview and two modes of switch input on typing accuracy and speed during a copy-spelling task. **Main results.** For participants without disabilities, Inquiry Preview and switch input tended to worsen typing performance compared to the standard RSVP Keyboard condition, with more consistent effects across participants for speed than for accuracy. However, there was considerable variability, with some participants demonstrating improved typing performance and better user experience with Inquiry Preview and switch input. Typing performance for the potential end user was comparable to that of participants without disabilities. He typed most quickly and accurately with Inquiry Preview and switch input and gave favorable user experience ratings to those conditions, but preferred standard RSVP Keyboard. **Significance.** Inquiry Preview is a novel multimodal interface for the RSVP Keyboard BCI, incorporating switch input as an additional control signal. Typing performance and user experience and preference varied widely across participants, reinforcing the need for flexible, customizable BCI systems that can adapt to individual users. ClinicalTrials.gov Identifier: NCT04468919.

Keywords

Brain-computer interface, multimodal access, augmentative and alternative communication, locked-in syndrome, electroencephalography, assistive technology, event-related potential

1. Introduction

People who are unable to meet all of their face-to-face communication needs through natural speech alone because of severe speech and physical impairments (SSPI) often use augmentative and alternative communication (AAC) systems for message generation. AAC may include anything from gestures to handwriting to computer-based devices with dedicated software and synthesized speech output.(1) Individuals with SSPI may have difficulty using traditional computer access methods such as a mouse or touchscreen to control an AAC device, creating further barriers to effective communication. A variety of alternative access methods have been implemented to support AAC use with limited movement, including eye tracking, head tracking, and switch scanning. However, these are not viable options for individuals with total locked-in syndrome who have no voluntary motor function and may not work well for people who have residual movement that is weak, inconsistent, or unreliable. The development of new alternative access options is critical to ensuring that AAC systems can meet the needs of all individuals with communication impairments.(2,3) Brain-computer interface (BCI) systems, which use brain signals rather than muscle activity for device control, have emerged as a promising option for individuals with SSPI.

A variety of communication BCI (cBCI) systems using different brain signals and user interfaces have been developed and trialed with people with disabilities.(4,5) To take advantage of residual motor function and improve BCI performance, researchers have proposed multimodal or hybrid systems that incorporate signals generated by eye movements or other body movements, combining these inputs with brain responses for system control.(6) One such signal is switch activation.(7) There are many types of adaptive switches, ranging from physical buttons (which can be pressed by a hand, foot, or other body part), to sip-and-puff switches (controlled by the user’s breath), to sensors that can detect eye blinks, small muscle twitches, or electrical signals from nearly imperceptible muscle activations via electromyography (EMG). Many switches have binary output, similar to a mouse click, while others produce a continuous signal based on the level and duration of detected muscle activity. Switch input has been used for a variety of purposes in BCI systems. Some studies have used button-press modes to establish a baseline or to ensure sustained attention during a task, primarily in BCIs for surveillance applications.(8) Some cBCI systems have used EMG switch input for error correction (e.g. deletion of the previously selected character).(9,10)

RSVP Keyboard is a non-implantable, wearable cBCI system designed for text entry by people with SSPI.(11,12) Like other BCI systems involving rapid serial visual presentation of stimuli, it uses event-related potentials (ERPs), collected via electroencephalography (EEG), for system control.(8) RSVP Keyboard combines these ERPs with evidence from an integrated language model (LM) to inform which characters are presented to the user and which are selected for typing.(13) Previous exploration of RSVP Keyboard use by people with SSPI revealed wide variations in typing accuracy and speed,(12) indicating that it may not be the optimal cBCI option for some users. Other studies have demonstrated that individuals with SSPI who are unsuccessful typing with one cBCI system may have better performance with another system that relies on different control signals or a different user interface,(14–16) suggesting a need for a variety of cBCI options that can be customized for individual users. Our research team is developing a software suite, BciPy,(17) that can accommodate a variety of movement- and brain-based control signals, as well as a selection of user interface designs.

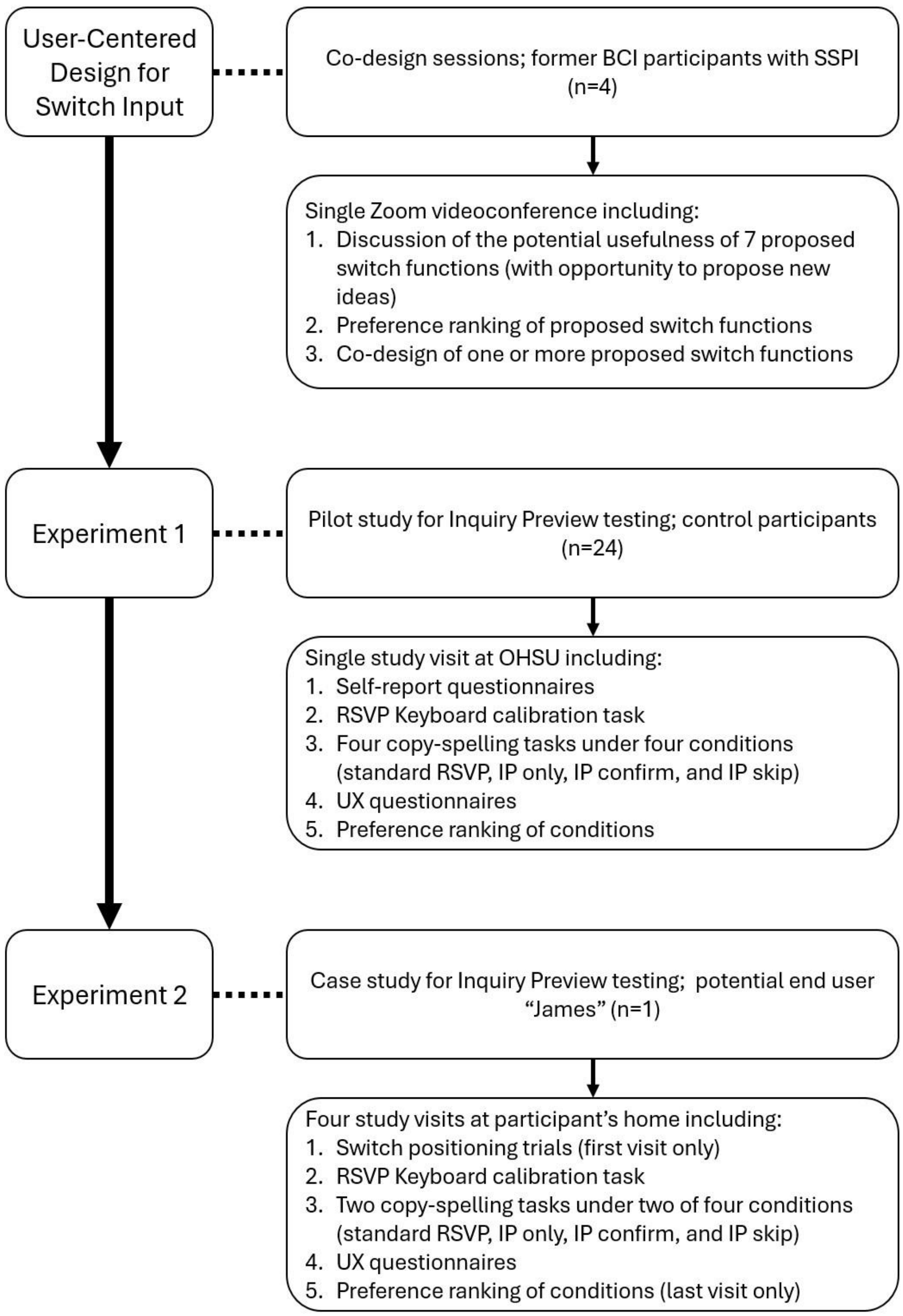
Here we introduce Inquiry Preview (IP), a new interface for RSVP Keyboard that adds a multimodal option with switch input as another control signal. We will present the user-centered design process that led to the development of IP, describe its design and functioning, and report results from two experiments: 1) a pilot study with control participants, and 2) a case study of a potential end user with SSPI. Our goal was to explore how the use of IP affected typing accuracy and speed, classifier performance, physical effort, and user experience (UX).

2. Methods

User-centered design for switch input

To ensure that the multimodal RSVP Keyboard with switch input would meet the needs and preferences of potential BCI end users, we implemented a user-centered design approach.(18,19) In October 2020, four former BCI study participants were invited to act as consultants in collaborative design sessions (see overview in figure 1). All consultants had SSPI and had previous experience with RSVP Keyboard and commercially-available (non-BCI) AAC systems. See supplemental table S1 for additional information on consultant demographics, diagnoses, communication methods, and device access methods. Each consultant participated in an individual co-design session with two or three researchers. Sessions were conducted via the Zoom videoconference platform and recorded. Interview topics and questions were shared in advance of the meeting so consultants could prepare responses if desired, and consultants used their existing communication methods to participate. All study activities were approved by the Oregon Health & Science University (OHSU) Institutional Review Board. Consultants provided informed consent and received monetary compensation for their participation.

Figure 1. Diagram summarizing participant involvement and study visit procedures for the user-centered design process with potential end users, a pilot study with control participants, and a case study with a potential end user. SSPI = severe speech and physical impairments; IP = Inquiry Preview; UX = user experience.



During each session, the consultant was presented with a list of potential functions for switch integration in RSVP Keyboard, generated by the research team. They were encouraged to propose their own ideas for functions not already listed, but none did so. Consultants were asked to consider whether each proposed function might be helpful to themselves or other RSVP Keyboard users, and to rank their top three functions in order of preference for further development. Finally, they were invited to participate in collaborative design exercises for one or more of their top-ranked functions. During co-design, the research team asked questions about the consultant's preferences for functionality, interaction style, appearance, and feedback for switch integration in RSVP Keyboard. One researcher (author BP) shared her screen via the videoconference platform and used PowerPoint (Microsoft, Redmond, WA) to create and edit interface mockups as the consultant and research team discussed potential design options. Table 1 lists the proposed functions for switch input, consultant feedback, and functions chosen for collaborative design.

Table 1. Proposed functions for switch input in RSVP Keyboard and the number of end user consultants (n=4) who identified each function as a potentially useful RSVP Keyboard feature, as their most preferred feature idea, and as a target for collaborative design. One participant indicated a tie between his two most preferred feature ideas.

Proposed switch function	Potentially useful feature	Most preferred feature idea	Chosen for co-design
Backspace (delete previously typed character)	4	1	1
Access a list of stored phrases	4	0	1
Abbreviation expansion (expand an abbreviation to a full phrase, e.g. ‘TY’ -> ‘Thank you’)	3	0	0
Call bell (make a sound to alert a caregiver)	4	1 (tie)	0
Pause RSVP stimulus presentation	4	1	1
Speak typed message	4	1 (tie)	2
Inquiry Preview (described below)	4	1	1

End user consultants responded positively to all proposed switch functions. There was wide variety in the functions consultants identified as their most preferred, and in those they chose for collaborative design. Given the lack of a clear consensus among consultants, the research team discussed which switch function should be the first to be integrated into RSVP Keyboard based on both consultant preferences and the novelty and potential results of each proposed function. They selected IP, which had been described to consultants as follows: “The RSVP Keyboard shows 10 letters in each sequence. (A sequence is the set of quickly flashing letters that follows each red +.) That means the letter you want to type may not be in the sequence. We could display those 10 letters all at once before the sequence starts, so you would know in advance whether your letter is going to appear. If your letter is not in the sequence, you could activate a switch to skip to the next sequence." IP was selected because it is a novel interface feature and multimodal input option that had not been previously explored for cBCI systems, and because of its potential effects on typing performance. All consultants felt that IP would be a potentially useful feature in RSVP Keyboard, and the consultant with the most previous RSVP Keyboard experience rated it as his most preferred switch function and chose it for his co-design session. He expressed that IP might help users feel more in control and avoid the frustration he had felt in the past when the system did not provide enough opportunities to select the target letter, especially in cases where the system then made an incorrect selection. The interface design mockup resulting from his collaborative design session is shown in Figure 2. Consultants' feedback on the other features will be considered in future development of the RSVP Keyboard interface.

Figure 2. Design mockup of the Inquiry Preview interface from a co-design session with a consultant. The top line of text is the target phrase for copy-spelling, with the target word in green. Beneath it is the typed string. In the center of the screen, in the same location in which the letter stream will appear, is the preview of the 10 letters that will be included in the next inquiry.

Inquiry Preview

User interface

In the RSVP Keyboard interface, the user is presented with a stream of symbols quickly flashing in the center of a computer monitor. The appearance of the user's desired symbol in this rapid serial visual presentation elicits an ERP, detected via EEG. Presentation is divided into inquiries, each containing a set of symbols (often nine letters plus a backspace symbol) determined based on letter probabilities assigned by the integrated LM and on EEG responses to previous inquiries. Typically, EEG data from multiple inquiries are required before the system has enough information to make a confident decision regarding the user's intended target.

When typing with RSVP Keyboard in standard mode, the user does not know which characters will be included in an upcoming inquiry. In the IP interface, the system displays a preview prior to each inquiry showing the letters that will be included (see figure 2). The preview is followed by a fixation cross and letter stream, as in standard mode. IP typing mode can be configured to add switch input as another source of evidence for determining the characters displayed in each inquiry and the characters selected by the system. When the preview box is on the screen, the user can activate a switch to indicate either that their target letter is present in the preview (in which case the previewed inquiry will be displayed), or absent from the preview (in which case that inquiry will be skipped and the system will preview the next inquiry with a different set of characters; this is the version originally described to collaborative design consultants). Switch evidence is combined with LM and EEG evidence to guide character presentation and selection, as described below. Because multiple inquiries may be presented for each character selection, and each inquiry is preceded by a preview, there may be more than one switch activation per selection.

Inquiry preview signal modeling

Each symbol selection when typing with RSVP Keyboard is based on evidence from a series of inquiries. After each inquiry, a recursive Bayesian update of symbol probabilities is performed, incorporating three sources of evidence: LM, switch, and EEG. Before the first inquiry of a series, the LM provides prior probabilities for each symbol, conditioned on any previously typed characters. When IP typing mode is configured to incorporate switch input, a switch input model provides multiplicative updates based on the set of symbols displayed in a preview and the user's response. Finally, after each inquiry, an EEG signal model provides multiplicative updates based on the user's EEG response to each displayed symbol. These modeling inputs are described in greater detail below.

Language Model

Before the first inquiry in a series, the LM generates evidence as to which characters are most likely to come next based on any previously typed text (left context). The LM in the current study was built on top of the 125M parameter version of the GPT-2 model.(20) Given a left context, we queried GPT-2 for a probability distribution over the tokens in its vocabulary (approximately 50K). These tokens are sequences of characters, ranging in length from a single character to a full word, that frequently appeared in the model's training data. Since the RSVP Keyboard types one character at a time, our LM needed to provide evidence for each series of inquiries by generating a probability distribution over individual characters.

To produce a distribution over characters (instead of over tokens), RSVP Keyboard first removes the last token from the left context. By removing this token, we allow the model the opportunity to predict tokens that extend it, instead of only tokens that follow it. For example, if the left context is "yeste" (in the word "yesterday"), GPT-2 splits this into two tokens: "y" and "este". If we do not remove the final token “este”, the model cannot consider the possibility of the token sequence “y” followed by “ester”. That is, it could only consider “y”, “este”, followed by a new token. After removing the final token, we generate a probability distribution over GPT-2's token vocabulary. We consider the 20 most probable tokens that start with the text of the removed token (“este” in our example), and sum the probability mass over the next character in each token (e.g. “r” in “ester”, or “d” in “ested”). We then normalize to sum to one, producing a probability distribution over the next character. To prevent any character from having a probability of zero, we mix GPT-2's probability distribution (weight 0.8) with the distribution given by a letter unigram model (weight 0.2). The letter unigram model was based on the frequencies of each letter in a set of phrases created by people with amyotrophic lateral sclerosis.(33) We took the resulting mixed distribution to the power of 0.5 and renormalized to smooth the distribution towards uniform.

Switch Input Model

As described previously, IP mode includes an option for users to provide switch input as an additional source of evidence. Depending on system configuration, switch activation during presentation of an inquiry preview has two possible functions: it can indicate either that the user’s desired symbol is present in the preview (IP confirm), or that the desired symbol is absent from the preview (IP skip). To simplify modeling and avoid introducing extra parameters, user behavior is modeled under two assumptions. First, user switch input was assumed to be independent of previous history, such that a single switch activation is fully determined by the currently desired symbol and the currently displayed preview. This assumption is typically referred to as the Markov assumption.(21) Second, users are assumed to make symmetric errors with a fixed probability of 5%; in other words, if the desired letter is displayed, the user is assumed to have a 95% probability of correctly pressing the switch in IP confirm mode, and to have a 95% probability of correctly not pressing the switch in IP skip mode. Based on these assumptions, symbol probabilities were updated as follows: based on the switch input mode (IP confirm or IP skip) and user response (whether or not the switch is activated), each symbol in a confirmed (or unskipped) inquiry is given a multiplicative update of 0.95, while each symbol in a skipped (or unconfirmed) inquiry is given a multiplicative update of 0.05.

EEG Signal Model

EEG signals are classified in RSVP Keyboard using a multi-stage pipeline as described previously by Memmott and colleagues,(17) and below in “Signal Acquisition and Processing”. After each stimulus presentation, a posterior probability update is made by using a multistage pipeline to compute two likelihood terms: the probability of observing the recorded EEG response given that the trial displayed the desired symbol *p(data | target)*, and the probability of observing the response given that the trial did not display the desired symbol *p(data | non-target)*. In the first stage of this pipeline, the data are compressed and zero-variance directions are removed using a linear dimensionality reduction method, channel-wise principal components analysis (PCA). Next, a regularized discriminant analysis (RDA) (22) classifier is used to further project the EEG responses into a one-dimensional representation. RDA is an intermediate between Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA), where the shrinkage and regularization hyperparameters are adjusted so that the class covariance estimates are closer to the overall data covariance. This is a useful method in cases where the sample size is small and the number of highly correlated features in the multivariate dataset is large. At this stage, calculation of the area under the curve (AUC) of the receiver operating characteristic (ROC) curve gives a suitable representation of classifier performance and helps determine these hyperparameters. AUC varies between [0,1], where 1 indicates perfect detection (no miss & no false alarms) and a value of 0.5 indicates random chance. After the selection of regularization and shrinkage parameters by cross validating the AUC, covariance and mean estimates for each class are used to generate a scalar feature that minimizes expected risk under the Gaussianity assumption of class distributions. Using RDA, the log-likelihood ratio (i.e. score) of target and non-target classes is computed as follows:

$$\delta_{RDA}(x) = log \frac{f_N(x;\ \widehat{\mu_1}, \widehat{\Sigma_1}(\lambda, \gamma))\widehat{\pi_1}}{f_N(x;\ \widehat{\mu_0}, \widehat{\Sigma_0}(\lambda, \gamma))\widehat{\pi_0}}$$

where $\widehat{\mu_c}$, $\widehat{\pi_c}$ are the class mean and prior estimates for target and non-target classes (1 and 0, respectively), $\boldsymbol{x}$ is the data to be classified, λ and γ are the shrinkage and regularization parameters, and $f_N(\boldsymbol{x};\ \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the multivariate normal (Gaussian) distribution probability density function. This score gives a 1-dimensional representation of the data $\boldsymbol{x}$. Finally, gaussian kernel density estimation (KDE) is used to estimate the conditional probability density functions of RDA scores given the class labels, $P(\delta_{RDA}(x_s)|\ c_s = c)$, which we use to fuse with LM evidence and perform posterior updates. For further details, see (23).

Evidence fusion and symbol selection

After the multiplicative update based on the user's EEG response to each symbol in an inquiry, the symbol probabilities are re-normalized to sum to 1. Any symbol with a posterior probability above 80% is selected and typed. If no symbol reaches this decision threshold after a predetermined number of inquiries (eight in the current study), the symbol with the highest probability is selected by default. A new series then begins, with new prior probabilities assigned by the LM based on the updated left context.

Experiment 1: Pilot study with control participants

To test the efficacy of RSVP Keyboard with IP and switch input, we conducted a single-visit experimental pilot study at Oregon Health & Science University (OHSU) in Portland, Oregon, USA. Data were collected between July and September, 2022.

Participants

Thirty-one participants without severe speech or physical impairments were recruited for pilot testing. All participants met the following inclusion criteria: 1) 21 to 89 years of age; 2) self-reported the ability to see words on a computer screen, read and communicate in English, and pay attention to a computer-based task for up to 3 hours; 3) no known or reported communication impairment or physical impairment affecting the ability to use a computer; 4) lack of cognitive impairment as indicated by a score of 32 or higher on the Telephone Interview for Cognitive Status (24); and 5) RSVP Keyboard calibration AUC of 0.70 or higher at the start of the data collection session. Participants were recruited from local universities and colleges, the OHSU Study Opportunities web page, and word of mouth. This study was performed in accordance with the Declaration of Helsinki and was reviewed and approved by the OHSU Institutional Review Board (protocol #23625). All participants provided informed consent and received a monetary incentive.

Procedure

Each participant attended a single study visit (see overview in figure 1). After providing informed consent, participants completed a series of questions about their demographic information, medications, and sleep habits, administered via Qualtrics online survey tools (Qualtrics, Provo, UT). They then completed an RSVP Keyboard calibration task without IP in order to generate a predictive EEG model for target classification. Those who successfully completed calibration and obtained AUC scores of 0.70 or above attempted the RSVP Keyboard copy-spelling task under four conditions: standard RSVP typing mode (without IP); IP only (i.e. the user was shown which symbols would be in each inquiry but did not use switch input and had no way to confirm or skip an inquiry); IP confirm (i.e. switch activation during preview presentation confirmed that the target symbol was displayed and was followed by presentation of the inquiry for EEG input);

and IP skip (i.e. switch activation during preview presentation skipped the displayed inquiry because the target was not present, and a new inquiry was displayed). Condition order was counterbalanced using a Latin square, such that the four unique condition orders occurred equally across the sample. Instructional videos (see supplemental materials) were presented prior to the calibration task and all four unique experimental conditions. Before calibration and before each copy-spelling session, participants rated their sleepiness on the Karolinska Sleepiness Scale (25,26) and their degree of headache pain, if any, on a 5-point Likert scale. After each copy-spelling session, they completed a UX questionnaire about the condition they had just tried.

Signal acquisition and processing

All data were collected using a Lenovo Legion 5 Pro laptop with Windows 11, an Intel Core i7-11800H @ 2.30 GHz, 16 GB DDR4 RAM, and an NVIDIA GeForce RTX 3050. Trigger fidelity on the experiment laptop was verified using a photodiode. The results of this timing test were used to determine static offsets between stimulus event marker time codes and the true physical appearance of stimuli onscreen to prevent experimentation with timing violations greater than +/- 10ms. All experiment software was written by the research team and is freely available on GitHub or PyPi using BciPy version 2.0.0rc1 or above.(17)

EEG data were acquired at 300 Hz using a DSI-24 dry electrode cap (Wearable Sensing, San Diego, CA, USA). All data were pre-processed using an inquiry-based approach for training and online usage: data for each two-second-long inquiry were collected with a two-second buffer on each end. Inquiries were notch-filtered at 60 Hz with a quality factor of 30 and then filtered at 1-20 Hz with a second-order bandpass filter before being downsampled by a factor of two (from 300 Hz to 150 Hz). Pre-processed inquiry data were subsequently segmented into 500 ms epochs (trials) synchronized with the onsets of individual symbol stimuli presented during each inquiry.

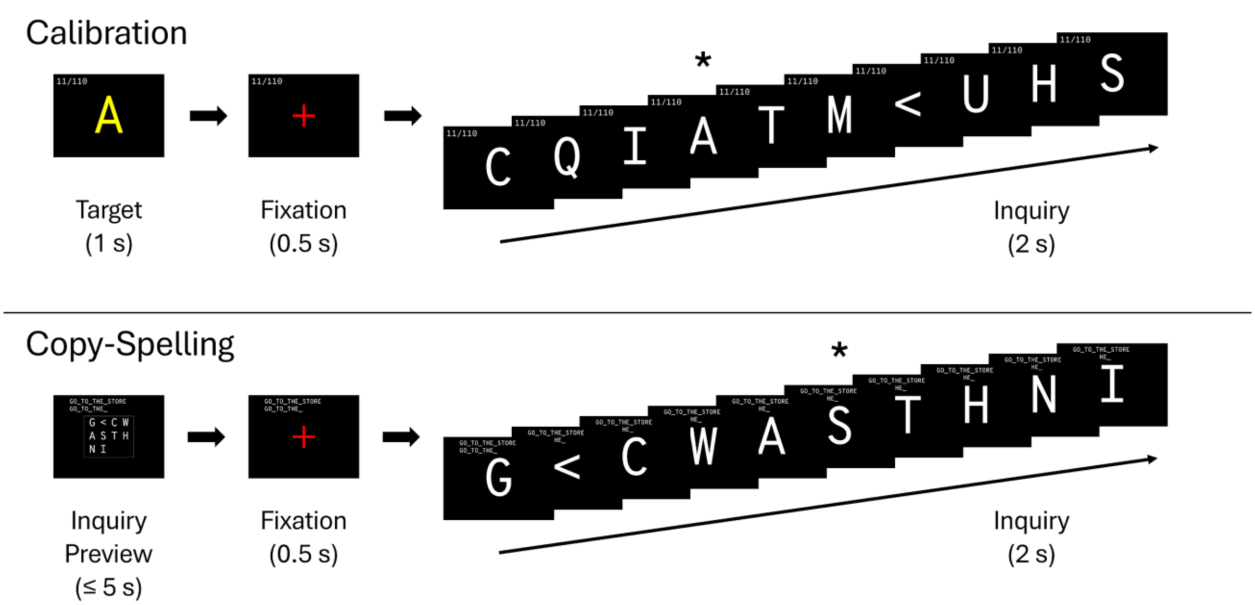
EEG data were semi-automatically inspected using MNE (27) offline with the following voltage criteria: +/- 50 uV or below 0.5uV. Blinks were labeled using the MNE find_eog_events using channels FP1/2 and F7/8. All labels were manually inspected for correctness and artifacts missed by auto-labeling were added. If greater than >50% of epochs or channels were contaminated with noise, the participant was flagged for sensitivity analysis (see below). Data from four participants exceeded this threshold, with overlapping EOG and movement artifacts being the primary drivers.

A Jelly Bean switch (AbleNet, Roseville, MN, USA) was used for binary switch input; participants could choose to leave the switch on the table and press it with either hand or to hold it in one hand and squeeze to activate it.

RSVP Keyboard configuration

During the calibration and copy-spelling tasks (see figure 3), stimuli were presented at 5 Hz in inquiries of 10 symbols, with the backspace symbol included in each inquiry. In IP copy-spelling mode, the preview box was displayed for five seconds or until the switch was activated, whichever came first. A maximum of eight inquiries were displayed in each series; if no symbol's posterior probability reached the decision threshold of 0.80 after eight inquiries, the symbol with the highest probability was selected.

Figure 3. Task schematic outlining the structure of single inquiries during the RSVP Keyboard calibration and copy-phrase tasks (letter sizes are not to scale). For calibration, each inquiry began with a target prompt, followed by a fixation cross and then 10 letter stimuli presented sequentially over 2 s (0.2 s per character). A total of 110 inquiries were presented; the target (marked with *) was present in 100/110 inquiries. The calibration task was used to generate an EEG model for use in all copy-phrase conditions. During Inquiry Preview (IP) copy-spelling tasks, each inquiry began with a preview box presented for up to 5 s, followed by the fixation cross and then 10 letter stimuli presented sequentially over 2 s. In the IP only condition, the preview would remain on screen for a full 5 s. In the IP confirm condition, participants could activate the switch during preview box presentation to proceed immediately to fixation. In the IP skip condition, participants could activate the switch during preview box presentation to skip immediately to the next preview with an updated character list. In all cases, the order of characters in the preview box matched the order of the characters presented in the letter stream. In the standard RSVP Keyboard condition, no preview box was presented and each inquiry began with fixation.



Tasks

Calibration

The RSVP Keyboard calibration task consisted of 110 trials, each beginning with a target symbol displayed in the center of the screen, followed by a fixation cross and an inquiry of 10 symbols. No inquiry previews were displayed during calibration. Participants were instructed to look for the target symbol in each inquiry and respond mentally when it appeared (see supplemental materials for instructional videos). An EEG classifier was trained on each participant’s calibration data for use in the copy-spelling tasks.

Copy-spelling

Under each condition, participants attempted to use RSVP Keyboard to copy-spell four five-letter words (see supplemental materials for instructional videos). To ensure an equal number of switch activation opportunities in the IP confirm and IP skip conditions, each word set included two “easy” words (in which all target letters had LM-assigned prior probabilities that ensured they would appear in the first inquiry of the series) followed by two “hard” words (in which one target had a lower prior probability and did not appear until the third inquiry or later). For each participant, a different set of target words was selected at random from a predetermined list for each condition. Only five series were presented for each word, so participants had five symbol-selection opportunities when attempting to spell each five-letter word, for a total of twenty opportunities under each condition.

User experience questionnaires

After each copy-spelling session, participants completed a short questionnaire about their workload and satisfaction (28) as well as what they liked and disliked about typing under that condition. After the final copy-spelling session, participants were asked to rank the four conditions in order of preference and provide a brief explanation for their ranking. All UX and preference questions were administered via Qualtrics online survey tools (Qualtrics, Provo, UT).

Data management & analysis

Session files from BciPy were processed using RStudio (2022.02.3, Build 492) and a custom R ((29) version 4.2.0) script was used to extract and calculate outcome measures related to typing and classifier performance. Typing performance measures and Qualtrics questionnaire responses were analyzed with Stata (StataCorp LLC, College Station, TX, USA; versions 17 and 18) and visualized with R (version 4.2.1), RStudio (version 2022.07.1, Build 554), and Excel 2019 (Microsoft, Redmond, WA, USA).

Primary dependent variables included measures of accuracy and speed, for both typing performance and classifier performance. To assess typing performance, accuracy was calculated as the percentage of target letters that were correctly copied, and speed was calculated as the number of correctly copied target letters per minute. To assess classifier performance, accuracy was calculated as the number of correct selections out of the total number of selections, and speed was calculated as the number of correct selections per minute. Differences in the calculation of typing performance variables and classifier performance variables reflect the fact that many correct selections were of the backspace character (correctly selected to delete an error) rather than of target characters. Secondary dependent variables included the number of switch activations per selection (for the IP confirm and IP skip conditions) and self-reported workload and satisfaction ratings.

In order to estimate the effects of condition and condition order on the primary dependent variables within each session, we used a population-averaged Poisson model clustered by participant with clustered jackknife standard errors (robust to violations of Poisson model assumptions). For accuracy measures, the count of target letters copied (typing accuracy) or of correct selections (classifier accuracy) was predicted from the model and divided by 20 (total selections) to convert to a percentage. For speed measures, the mean session length (across 4 sessions for each participant) was included as an exposure term in the model (i.e. the natural log of mean session length was added to the model with coefficient constrained to 1), and predictions divided by 4 to account for the fact that the mean session length describes an average single session while the count of the measure represents the total across all 4 sessions. The models for each primary dependent variable included fixed effects to allow comparison of the effects of the conditions and of condition order within the data collection session. Sensitivity analyses were conducted to assess how the inclusion of the four participants with poor data quality affected the model predictions. Simple correlation analyses were used to explore the relationships between calibration and typing performance and participants' preferences across conditions. Secondary dependent variables were summarized using descriptive statistics.

Experiment 2: End user case study

After the pilot study, "James", one of the end user consultants who participated in the initial user-centered design sessions described above, was asked to try the new interface and provide feedback. He is a White, non-Hispanic man, 55 years old at the time of data collection, with incomplete locked-in syndrome secondary to a brainstem stroke. He breathes independently and uses a manual wheelchair, pushed by a caregiver, for mobility. For communication in everyday life and during study visits, James uses eye movements for yes/no responses and to signal a partner to spell messages via partner-assisted scanning. His other voluntary movements include blinking and slowly turning or lifting his head, and he experiences occasional coughing that causes movement in his head and upper body. James had been a member of the research team for approximately 14 years, providing input and feedback on RSVP Keyboard research and development, and had many hours of experience calibrating and typing with RSVP Keyboard in standard mode.

Data were collected in James's home during four visits (see overview in figure 1) over two weeks in May, 2023. Due to concerns about fatigue and the apparent effects of condition order observed in the pilot study (see Results), and to allow him more than one opportunity to try each condition before forming an opinion, he completed four data collection sessions instead of the single session used in the pilot study. During each home visit, he calibrated RSVP Keyboard, attempted copy-spelling under two of the four conditions, and completed user feedback questionnaires. The first visit also included trialing positioning options for the Jelly Bean switch. Based on James's preference and results from simple switch activation testing with Scanning Wizard (Koester Performance Research, Ann Arbor, MI, USA (30)), the switch was mounted to his wheelchair with an articulated arm mount (Manfrotto, Cassola, Italy) and positioned at the right side of his chin so that he could activate it with a slight head turn. Conditions were presented in pseudorandom order, with each condition occurring twice (once as the first copy-spelling condition within a visit and once as the second condition). Signal acquisition and processing, RSVP Keyboard configuration, and calibration, copy-spelling, and UX questionnaire tasks were identical to those in the pilot study, with the following exceptions: 1) due to LM issues discovered during the pilot study (see below), James used a different version of the LM code than the control participants; 2) he was allowed to re-attempt copy-spelling of individual words when visual analysis revealed signal interference caused by electrodes pressing or rubbing against his wheelchair headrest during typing (in these cases, data from the initial attempt were omitted from analysis); and 3) he provided questionnaire responses and narrative feedback via partner-assisted scanning. Descriptive statistics were used to summarize typing performance, classifier performance, and UX, with the same dependent variables as for control participants.

3. Results

Experiment 1: Pilot study with control participants

Of 31 participants who met initial inclusion criteria, seven (22.6%) were excluded due to low AUC (< 0.70). Twenty-four participants without disabilities completed data collection under all four conditions and were included in the analysis; their demographic characteristics are summarized in supplemental table S2. Although four participants had poor EEG data quality (determined by visual analysis after data collection), sensitivity analysis revealed that including their data had no significant impact on the findings (see below), so the results presented here include all 24 participants.

After completing data collection, we discovered a number of issues in the language model code that decreased its overall performance in Experiment 1. For example, minor issues included not conditioning on the sentence start symbol and not summing over both upper- and lower-case letters when calculating the probability distribution for the next character. The major issue was related to how GPT-2 segments text into a sequence of subword tokens. Our algorithm in Experiment 1 removed only the last subword token and then searched for all possible following subword tokens consistent with the user's current text. For certain texts, this led to poor predictions for the next character. A better approach (and the one used in Experiment 2), is to remove all subword tokens for the user's currently partially written word and search forward from there. These issues were present in all conditions of Experiment 1. Since stimuli text was chosen at random for each participant and for each condition, we have no reason to believe that these issues systematically affected any particular condition.

Typing performance

The distributions of percentage of target letters copied and number of targets copied per minute are summarized in figure 4, with participant-level data displayed in supplemental figure S1. Table 2 presents results of the analysis quantifying the effects of condition (each IP condition compared to standard RSVP) and condition order (the second, third, and fourth conditions attempted by each participant compared to the first condition they attempted) on typing performance. Negative average marginal effects indicate that IP conditions tended to worsen performance overall compared to standard RSVP, but did so more consistently (across participants) for speed than for accuracy, and more severely for IP only than for IP confirm or IP skip. Fatigue appeared to impact performance, with the largest detrimental effects observed during the final condition of the data collection visit, after prolonged system use. However, there was considerable variability, and some participants demonstrated improved performance with one or more of the IP conditions compared to standard RSVP, or at the end of the session compared to the beginning.

Figure 4. Distribution of percentage of target letters copied and number of targets copied per minute under each condition for control participants (n=24). Each box contains the 25th to 75th percentiles of values for one condition, with the internal line marking the median. Whiskers indicate adjacent values (the highest and lowest values within 1.5 times the interquartile range), and circles represent outlier observations outside the adjacent value range. IP = Inquiry Preview.

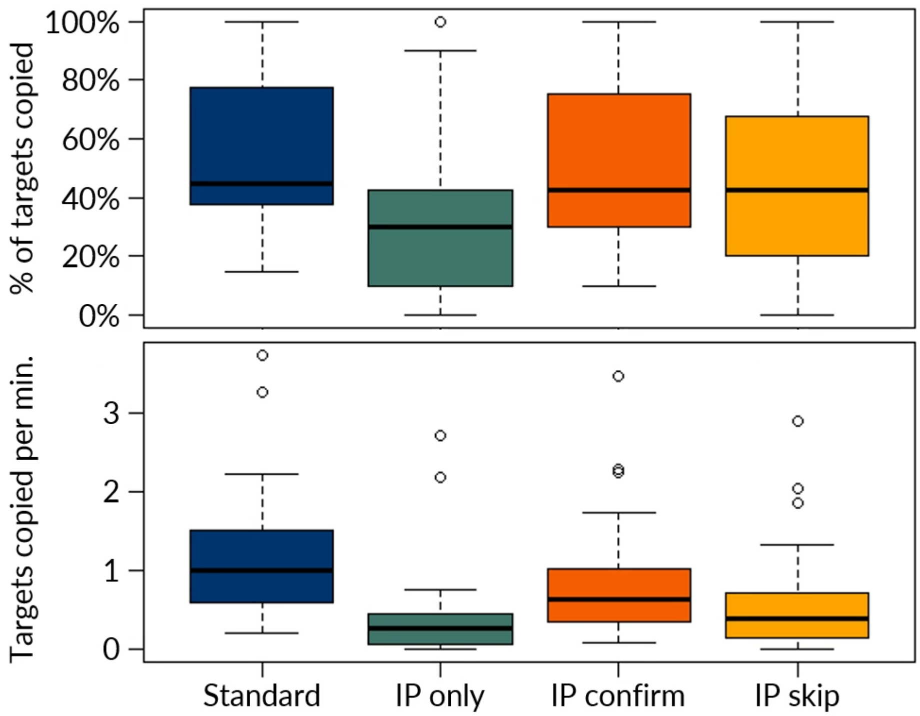


Table 2. Results of population-averaged Poisson model estimations of average marginal effects of condition and condition order on the percentage of target letters copied and target letters copied per minute for control participants (n=24). Each Inquiry Preview condition is compared to standard RSVP, and the second, third, and fourth conditions attempted in each data collection session are compared to the first condition. IP = Inquiry Preview; CI = confidence interval.

	% of targets copied		**Targets copied per minute**	
Condition	**Average marginal effect [95% CI]**	**p value**	**Average marginal effect [95% CI]**	**p value**
IP only	-22.7% [-41.3%, -4.1%]	0.019	-2.89 [-4.22, -1.57]	0.000
IP confirm	-8.8% [-24.7%, 7.2%]	0.269	-1.35 [-2.72, 0.01]	0.052
IP skip	-12.3% [-27.9%, 3.3%]	0.117	-2.24 [-3.46, -1.03]	0.001
Condition order				
2^{nd}	-9.4% [-23.2%, 4.4%]	0.174	-0.62 [-1.41, 0.17]	0.117
3^{rd}	-8.5% [-22.8%, 5.7%]	0.226	-0.53 [-1.43, 0.37]	0.234
4^{th}	-19.2% [-34.4%, -3.9%]	0.016	-1.11 [-2.06, -0.17]	0.023

Classifier performance

The effects of condition and condition order on classifier performance (number of correct selections out of total number of selections and number of correct selections per minute) were similar to those observed for the typing performance variables: overall performance across participants was worse under the IP conditions and for each condition attempted after the first, with larger effects for the IP only condition and the last condition of the day. See supplemental table S3 for details.

Switch activations per selection

The mean number of switch activations per selection for the two conditions involving switch input was 4.0 (SD 1.47) for IP confirm and 0.4 (SD 0.18) for IP skip.

User experience and preferences

Participant ratings of workload and satisfaction are summarized in supplemental table S4. Median ratings were similar across all four conditions, with little difference between standard RSVP and the IP conditions. Participants rated physical effort fairly low (median 1 to 2 on a 7-point scale), mental effort fairly high (median 5 to 6), and overall workload moderately difficult (median 4). Of note, physical effort ratings for IP confirm and IP skip were very similar, despite IP confirm requiring approximately ten times as many switch activations per selection as IP skip. Median satisfaction ratings of 3 to 5 on a 7-point scale indicated that participants were generally “somewhat satisfied” to “somewhat unsatisfied” with typing speed, typing accuracy, and the overall experience using the system. There was considerable variability across participants on all UX questionnaire items for all conditions, with ranges of 6 to 7 points for each.

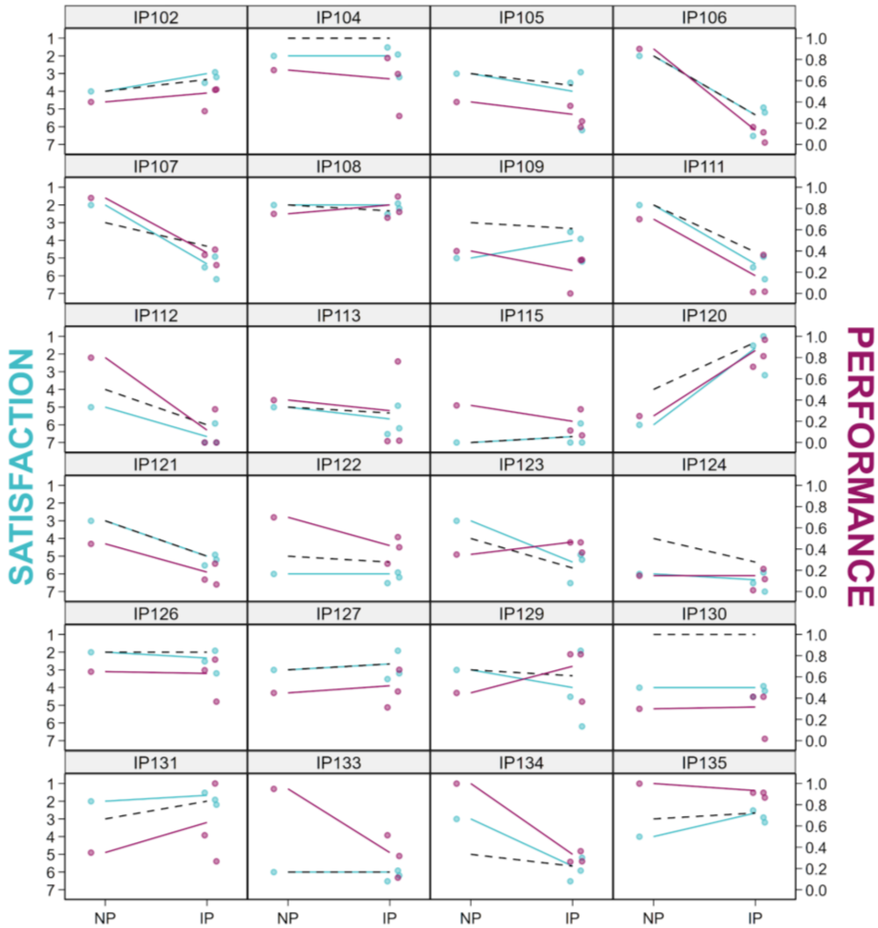
Differences in satisfaction and typing performance for the standard RSVP and IP conditions for individual participants are illustrated in figure 5. As an example of how to interpret this figure, consider the plot for IP104 in figure 5a. This participant correctly copied 70% of target letters for the standard RSVP condition (left purple marker and left endpoint of purple line) and a mean of 61.7% of target letters for the three IP conditions (right endpoint of purple line; performance in individual IP conditions is indicated by right purple markers). They rated their satisfaction with typing accuracy as a 2 (considerably satisfied) for standard RSVP (left teal marker and left endpoint of teal line), and the mean of their accuracy satisfaction ratings for the IP conditions was also 2 (right endpoint of teal line; ratings for individual IP conditions are indicated by right teal markers). They rated their overall satisfaction with the system as a 1 for all conditions, with standard RSVP represented by the left endpoint of the dashed black line and the mean for the three IP conditions represented by the right endpoint. The flat slope of the teal and dashed lines indicates that there was little difference in this participant’s satisfaction ratings for the standard

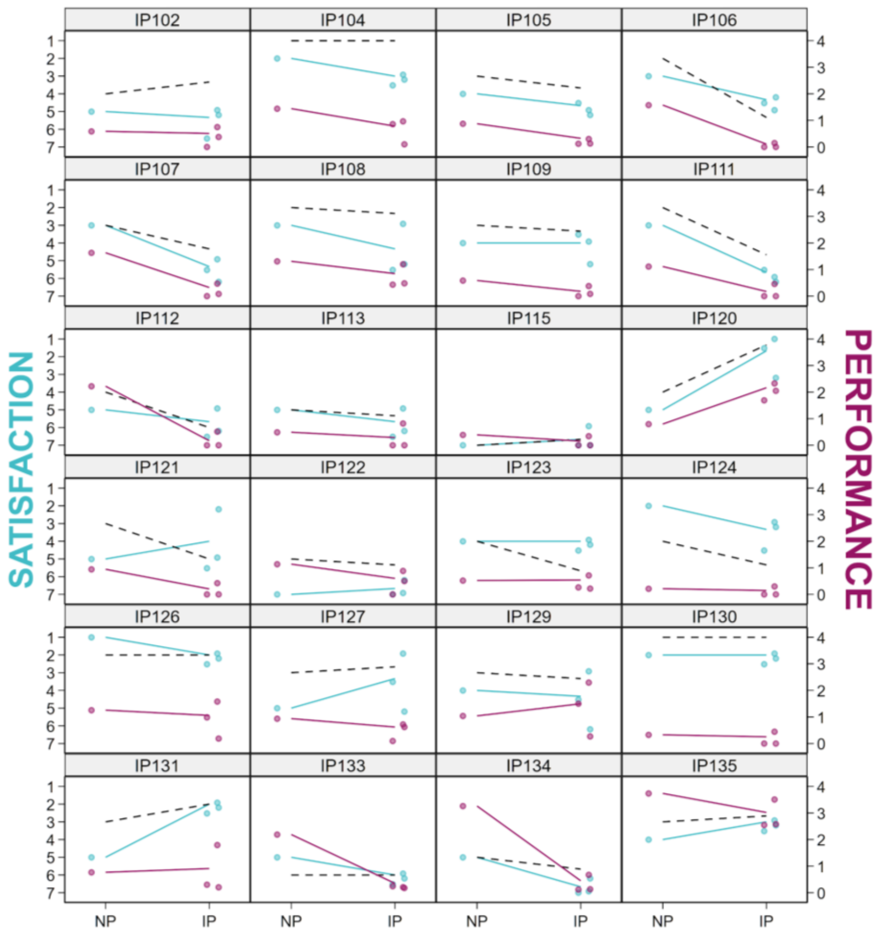
RSVP and IP conditions, while the negative slope of the purple line reveals that their percentage of correct letters copied for standard RSVP was higher than their average for the three IP conditions.

For many participants, typing performance and satisfaction ratings were similar across the three IP conditions, as indicated by the clustered markers on the right side of their plots, though there were some exceptions (e.g. IP104, IP129, and IP131). Ratings of satisfaction with typing accuracy appeared closely related to the actual percentage of targets correctly copied for many participants, though some reported low satisfaction across conditions regardless of performance (e.g. IP112, IP122, and IP133). This relationship was less pronounced for ratings of satisfaction with typing speed and correct characters per minute, with many participants reporting favorable ratings for typing speed even when performance was slow (e.g. IP104, IP124, and IP130). Overall, these figures highlight the wide variety in participant performance and preferences. For some, standard RSVP had better performance and satisfaction ratings than the IP conditions (lines with negative slopes in figure 5). For others the opposite was true (lines with positive slopes), and still others demonstrated little difference across conditions (lines with flat slopes).

Across all participants, typing accuracy, as measured by percent of target letters copied, demonstrated moderate correlations with ratings of satisfaction with system accuracy (r = 0.60) and overall satisfaction (r = 0.49). Typing speed, as measured by targets copied per minute, demonstrated weak correlations with ratings of satisfaction with typing speed (r = 0.32) and overall satisfaction (r = 0.38)

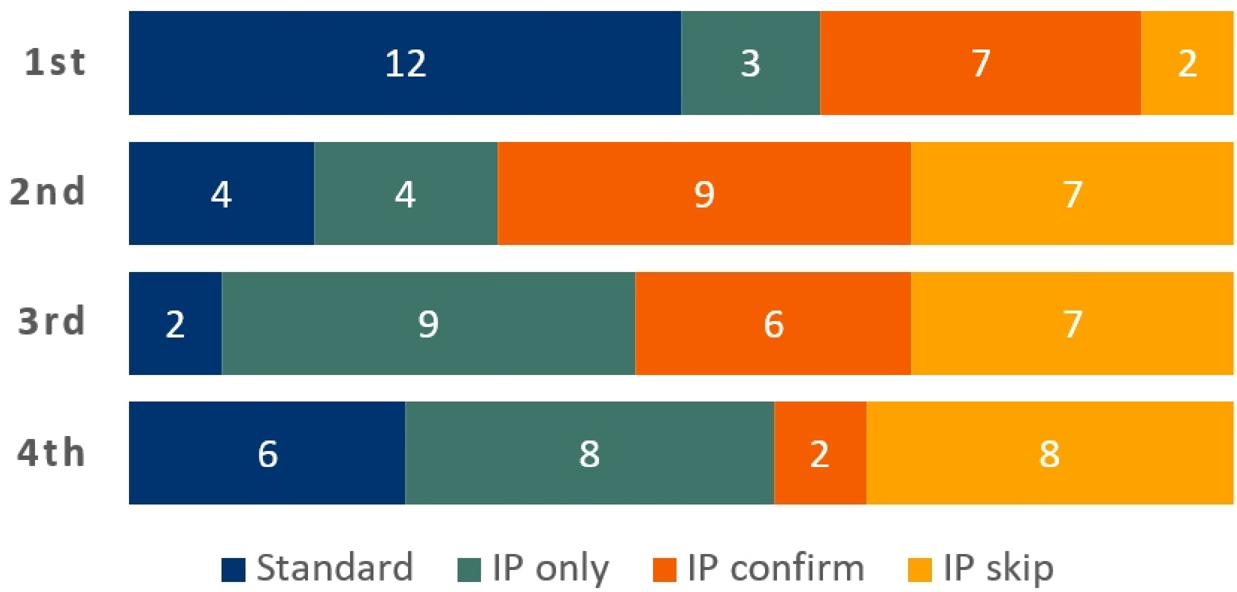
Figure 5a. Satisfaction with system accuracy and actual typing accuracy performance (percent of target letters copied) for each condition for individual participants. **Figure 5b.** Satisfaction with system speed and actual typing speed performance (targets copied per minute) for each condition for individual participants. Teal markers represent participant ratings of their satisfaction with typing accuracy (figure 5a) or speed (figure 5b) for a given condition on a 7-point Likert scale (left y-axis; 1 = extremely satisfied, 7 = extremely unsatisfied). Purple markers represent the percentage of target letters copied in a given condition (figure 5a) or target letters copied per minute (figure 5b). For each participant, markers on the left correspond to satisfaction and performance measures for the standard RSVP condition, while markers on the right correspond to the three IP conditions. Lines are included to emphasize differences in satisfaction and performance between the standard RSVP and the mean of the IP conditions. Dashed lines represent differences in participant ratings of their satisfaction with overall system performance for the standard RSVP and IP conditions.


IP102
IP104
IP105
IP106
IP107
IP108
IP109
IP111
IP112
IP113
IP115
IP120
IP121
IP122
IP123
IP124
IP126
IP127
IP129
IP130
IP131
IP133
IP134
IP135
SATISFACTION
PERFORMANCE
NP
IP
1
2
3
4
5
6
7
1.0
0.8
0.6
0.4
0.2
0.0



Participants' rankings of the four conditions in order of preference are summarized in figure 6. Standard RSVP and IP confirm were the two most popular conditions, with 12 and 7 first-place rankings, respectively. However, other participants rated these same conditions in last place, and even the less-popular conditions, IP only and IP skip, were the top choice of two or more participants. Participants did not always prefer the best-performing condition, as shown in supplemental figure S1. A selection of quotes from participants' narrative feedback about why they liked or disliked each condition are presented in supplemental table S5, and reveal a wide range of opinions on RSVP Keyboard, Inquiry Preview, and switch input. Different participants had very different reactions to the same condition; for example, with regard to IP only, one participant "liked having the letters beforehand," but another "did not like the box format or seeing the letter ahead of time.” Some participants specifically mentioned that they preferred using the switch to confirm that they had seen their letter instead of skipping to a new preview, while others felt the opposite.

Figure 6. Number of participants ranking each condition as their 1st, 2nd, 3rd, and 4th most preferred typing interface. IP = Inquiry Preview.



Sensitivity analysis

We reanalyzed all four outcomes (typing accuracy and speed, classifier accuracy and speed) in the performance data omitting responses from the four participants whose EEG quality appeared poor. The same population-averaged clustered Poisson models were used as above, including clustered jackknife estimation of standard errors, and for each outcome the average marginal effects for each contrast (each IP condition vs no IP and each condition order vs 1st in sequence) under this sensitivity analysis (n=20) were compared to the primary results that included performance data from all participants (n=24). We found no important differences: no effects changed sign or differed in magnitude by more than approximately 40% in either direction (usually less than 30%, and always so for the condition effects), and the standard errors were larger by approximately the amount expected from the smaller sample size. Effect estimates were sometimes smaller and sometimes larger in the sensitivity analyses, with no apparent systematicity, and our conclusions about the impacts of IP condition and condition order on performance were entirely unchanged. These sensitivity results suggest that we experienced no outlier problems or lack of model fit due to the inclusion of data derived from sessions characterized by poor EEG quality. See supplemental tables S6 and S7 for direct comparisons of sensitivity analysis results to primary results.

Experiment 2: End user case study

James completed all four planned study visits. Four words from three typing sessions were repeated due to signal noise. The bar chart in figure 7 displays his typing performance results (percent of targets copied and targets copied per minute) for the first and second conditions within each session; boxes representing the median and interquartile range for the control group are included for comparison. James averaged 2.5 switch activations per selection for the IP confirm condition, and 0.3 activations per selection for the IP skip condition. His mean UX questionnaire responses (across two sessions for each condition) and condition rankings in order of preference are presented in table 3. He gave his highest UX ratings to the IP confirm condition, selecting the most favorable rating for every workload and satisfaction item for both IP confirm sessions. He gave generally favorable ratings (2.5 or below on the 7-point scale, with lower scores indicating easier workload or greater satisfaction) to the standard RSVP and IP skip conditions, and slightly less favorable ratings (particularly for frustration and typing speed satisfaction) to the IP only condition. He ranked the conditions as follows, from most to least preferred: standard RSVP, IP skip, IP confirm, IP only. James's comments about each condition are included in supplemental table S5. He preferred the standard RSVP condition in part because “it’s what I’m used to” from previous interaction with the RSVP Keyboard as a research team member and study participant. He reported that both IP confirm and IP skip made him feel “more in control.”

Figure 7. Percentage of target letters copied and targets copied per minute for each typing session conducted with James. Data are presented in order of collection, with two typing sessions in each of the four visits. Bars represent James’s results, and the box on each bar indicates the median and interquartile range for control participants for the same condition, as shown in figure 4. IP = Inquiry Preview.

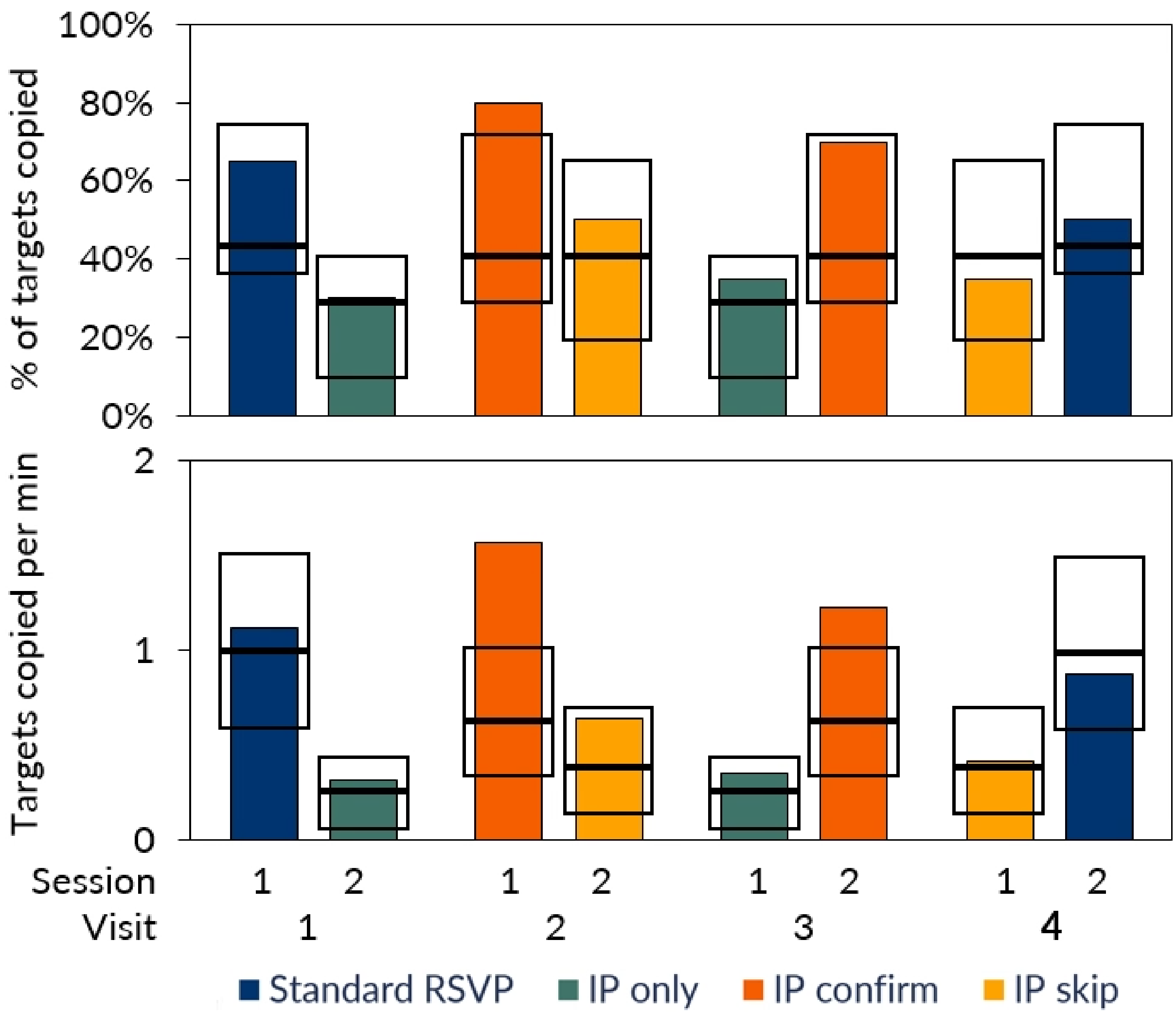


Table 3. Mean user experience ratings (averaged across two sessions for each condition), condition ranking in order of preference, and mean performance outcomes for each condition in Experiment 2. Rating scale ranged from 1 (best: extremely low workload or extremely satisfied) to 7 (worst: extremely high workload or extremely unsatisfied). For preference ranking, 1 represents the participant's most preferred condition. IP = Inquiry Preview.

	Standard RSVP	**IP only**	**IP confirm**	**IP skip**
Workload				
Physical effort	1	2	1	2
Mental effort	1.5	1	1	1
Time pressure	1	1	1	1
Frustration	2	4.5	1	2.5
Overall workload	**1.5**	**1**	**1**	**1**
Satisfaction				
Typing accuracy	1.5	2.5	1	2.5
Typing speed	1.5	3.5	1	1
Overall satisfaction	**2**	**2.5**	**1**	**1**
Preference ranking	1	4	3	2
Performance				
% target letters copied	57.5%	32.5%	75.0%	42.5%
Targets copied per minute	1.00	0.34	1.40	0.53
Switch activations per selection	N/A	N/A	2.5	0.3

4. Discussion

As non-implantable, wearable BCIs become a reality, design must be driven by potential end users, and pilot experiments must be presented as preliminary research and development steps. In this study, potential cBCI end users provided guidance on the choice of switch functionality and the Inquiry Preview interface layout as part of a user-centered design process. We demonstrated proof of concept for the IP interface and for fusing switch input evidence with

EEG and LM evidence in RSVP Keyboard typing tasks, with improved typing performance and user experience for some (but not all) study participants.

IP was designed to incorporate switch input as an additional control signal in a multimodal BCI system to take advantage of the residual motor function of some users. For example, some individuals may have enough strength and range of motion to activate a switch, but may struggle with the precise timing required for automatic scanning, or fatigue quickly such that they cannot use switch activation frequently for long periods of time. Depending on how it is configured, IP can offer single-switch input capability that is less time-dependent, more tolerant of unintended or missed activations, and requires fewer activations per character selection than automatic scanning. For certain users, adding switch input to a hybrid cBCI with the IP interface may support faster, more accurate, and more reliable communication access than either traditional switch scanning or EEG alone. More investigation would be needed to explore how the performance of a hybrid cBCI with IP compares to switch scanning for people with SSPI.

Among control participants, both typing performance and classifier performance were better under the standard RSVP condition compared to any of the IP conditions, with the worst performance observed under the IP only condition. However, there was considerable inter-subject variability, and some participants, including the end user participant in Experiment 2, demonstrated improved performance with one or more of the IP conditions compared to standard RSVP. This reinforces the results of several previous studies indicating that poor performance with one type of BCI does not necessarily predict poor performance with a different BCI.(14–16) The field continues to move away from the concept of “BCI illiteracy” and toward a landscape in which there are many options to suit the needs and preferences of individual users.(31)

Condition order appeared to affect typing and classifier performance, with large detrimental effects noted for the last condition of the day for control participants. Fatigue, boredom, and waning attention may affect RSVP Keyboard performance regardless of the interface and input method(s) used.(32)

UX feedback revealed wide variation in participants' perceptions of their workload and performance, as well as their preferences for one condition over another. Responses on nearly all Likert scale items related to workload and satisfaction spanned a range of 6 to 7 points on the 7-point scale. No condition emerged as a clear favorite among participants. Half of the control participants ranked standard RSVP as their favorite condition, but 25% ranked it as their least favorite. IP only and IP skip were tied for the worst-rated condition overall, but each was ranked first by at least two participants. Preference rankings did not necessarily align with typing performance results; for example, end user participant James identified standard RSVP as his favorite condition, despite demonstrating faster and more accurate performance with IP confirm, which he ranked as his 3rd most preferred condition. This discrepancy again highlights the need for multiple options in BCI systems and configurations, as well as the importance of collecting and reporting UX feedback in BCI studies. Self-rated user system satisfaction was moderately correlated with percent of target letters copied but weakly correlated with typing speed. This may reflect variability in user preferences in balancing speed vs accuracy.(33)

Participants had mixed opinions on using switch input with RSVP Keyboard. Some liked it, reporting that pressing the "button" made the task feel more "active," made the user feel "more in control," or "made the system move along faster." Participants who disliked the IP only condition (without switch input) reported that it felt slow and that they "didn't feel in control." One survey of potential BCI end users found that many report a preference for active strategies for cBCI control, such as imagined movement,(19) so the option of using an active strategy based on actual movement may be appealing to those with some residual motor function. The copy-spelling task also tended to take longer with IP only than with the other versions of the system (as shown in figure 4), in part because the preview was shown for the maximum duration before each inquiry instead of being cut short with switch input on some inquiries. Other participants in the current study preferred the conditions without switch input, with one stating that “the button felt weird when you are already using your thoughts for the letters.” On average, typing under the IP confirm condition required approximately ten times as many switch activations as typing under the IP skip condition. Most control participants, as well as the end user participant, reported that both switch-input conditions required low physical effort. However, some BCI users with minimal voluntary movement may find switch activation more effortful and fatiguing, and might prefer IP skip because of its relatively low physical demands.

An additional outcome of this study was the successful implementation of user-centered design activities via videoconference. Videoconferencing and screen sharing worked well for co-design and interface prototyping, allowing consultants and researchers to collaborate and to create and modify prototypes in real time. These techniques may be a valuable option for reducing barriers to the participation of people with SSPI in user-centered design, as well as reducing the costs to researchers. Our consultants all used AAC or alternative access methods in their daily lives, either currently or in the past, and had used earlier versions of the RSVP Keyboard as study participants. Their previous BCI use was helpful for their participation in the co-design process as they could imagine how the new features would add value, improve performance, or reduce frustration compared to their previous experiences. Co-design consultants expressed different preferences for new RSVP Keyboard features, as reflected in table 1. It is possible that involving a larger group of consultants may have led to stronger consensus, but that was not essential to our goals for this user-centered design process. We learned that this small group of experienced RSVP Keyboard users broadly agreed that all of the proposed features were potentially useful, and we were able to apply ideas generated during co-design to the final iteration of the feature that we believed was the most innovative and had the most potential to improve system performance. In this case, user-centered design did not lead to the identification of a single most popular feature to target for development, but it did reveal several features that end users felt would add value. One of our team's major goals is to increase the customizability of cBCI systems for individual users, and these features will be included in future system development to expand the options available.

Within this complex, interdisciplinary field, potential and actual end users with SSPI play many different roles. They guide us in design decisions, as is evident in the choice of the new IP feature. They share their values and opinions.(34) They participate in iterative development so that the systems remain relevant and usable. They test systems and provide feedback. In this study, although much of the usability testing for Inquiry Preview was conducted with control participants without disabilities, we involved end users with SSPI in several key phases of design, development, and evaluation. This combined approach, involving participants both with and without severe disabilities, may be appropriate in many cases in BCI research, particularly given the difficulty of finding and recruiting participants with SSPI and the need to accommodate their schedules, locations, and health needs. Inclusion of individuals with SSPI is a matter of justice,(35) and decisions about when it is best to add their values, opinions, suggestions and feedback should be guided by the end users themselves.

Limitations

Although this study demonstrated proof of concept for IP and switch input for the RSVP Keyboard, generalizability to potential BCI user populations is limited due to its small overall sample size and inclusion of a single participant with SSPI. Additionally, the use of a classifier based on data from a standard RSVP calibration task during copy-spelling for all four conditions may have affected typing performance during the IP conditions, as advance knowledge of the stimuli to be presented might modify the brain response. Further investigation is needed on this topic.

The language model bugs identified during data analysis for Experiment 1 form another limitation. While a few minor bugs decreased its performance slightly, the most notable bug was the way in which the RSVP Keyboard evaluated possibilities. We suspect that the bugs present in Experiment 1 resulted in a higher variance in the accuracy of the LM predictions. While the impact on any particular user depends on the exact stimuli they encountered, we do not have reason to believe that these bugs impacted any particular experimental condition disproportionately.

Finally, the fixed switch error rate (5%) used in this study may not be appropriate for all users or all situations. Some users may be prone to accidental switch activations due to involuntary movements, or missed activations due to fatigue, attention, or other fluctuating user states, and would benefit from a higher error rate. Others may have consistently high switch input accuracy and may prefer a lower switch error rate to more quickly drive the BCI toward their intended target. In future work, we will consider 5% as the default rate and adjust it based on user needs and preferences.

5. Conclusion

Inquiry Preview offers a novel multimodal interface for the RSVP Keyboard cBCI, incorporating switch input as an additional control signal for users with residual motor function. Variability in typing performance and user experience across participants highlights the importance of flexible, customizable cBCI systems that can adapt to individual users' needs and preferences.

Future work will involve additional testing and refinement of Inquiry Preview and switch input with end users, and continued development of BciPy with additional options for control signals, interfaces, and user-specific configuration.

Data availability statement

Data generated and analyzed for this study are available by request from the corresponding author.

Acknowledgements

We thank our participants for their time and effort, David Smith and Shijia Liu for their work on the language model used in this project, and Karuna Miller for assistance in formatting this manuscript.

Conflict of interest

The authors report no conflicts of interest.

Funding

This work was supported by the National Institutes of Health under National Institute on Deafness and Other Communication Disorders grant R01DC009834 and by the National Science Foundation under grant IIS-1750193.

References

1. Beukelman DR, Light JC. Augmentative & alternative communication: supporting children

and adults with complex communication needs. Fifth edition. Baltimore: Paul H. Brookes Publishing Co., Inc; 2020. 704 p.

2. Higginbotham DJ, Shane H, Russell S, Caves K. Access to AAC: Present, past, and future. Augment Altern Commun. 2007 Jan 1;23(3):243–57.
3. Fager S, Fried-Oken M, Jakobs T, Beukelman DR. New and emerging access technologies for adults with complex communication needs and severe motor impairments: State of the science. Augment Altern Commun. 2019 Jan 2;35(1):13–25.
4. Rezeika A, Benda M, Stawicki P, Gembler F, Saboor A, Volosyak I. Brain–Computer Interface Spellers: A Review. Brain Sci. 2018 Mar 29;8(4):57.
5. Peters B, Eddy B, Galvin-McLaughlin D, Betz G, Oken B, Fried-Oken M. A systematic review of research on augmentative and alternative communication brain-computer interface systems for individuals with disabilities. Front Hum Neurosci [Internet]. 2022 [cited 2024 Jun 28];16. Available from: https://www.frontiersin.org/journals/human-neuroscience/articles/10.3389/fnhum.2022.952380/full
6. Li Z, Zhang S, Pan J. Advances in Hybrid Brain-Computer Interfaces: Principles, Design, and Applications. Comput Intell Neurosci. 2019 Oct 8;2019(1):3807670.
7. Fishman I. Electronic communication aids: selection and use. Boston: College-Hill; 1987. 148 p.
8. Lees S, Dayan N, Cecotti H, McCullagh P, Maguire L, Lotte F, et al. A review of rapid serial visual presentation-based brain–computer interfaces. J Neural Eng. 2018 Jan 24;15(2):021001.
9. Perdikis S, Leeb R, Williamson J, Ramsay A, Tavella M, Desideri L, et al. Clinical evaluation of BrainTree, a motor imagery hybrid BCI speller. J Neural Eng. 2014 Apr 16;11(3):036003.
10. Riccio A, Holz EM, Aricò P, Leotta F, Aloise F, Desideri L, et al. Hybrid P300-Based Brain-Computer Interface to Improve Usability for People With Severe Motor Disability: Electromyographic Signals for Error Correction During a Spelling Task. Arch Phys Med

Rehabil. 2015 Mar 1;96(3, Supplement):S54–61.

11. Orhan U, Erdogmus D, Roark B, Oken B, Purwar S, Hild KE, et al. Improved accuracy using recursive Bayesian estimation based language model fusion in ERP-based BCI typing systems. In: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society [Internet]. 2012 [cited 2024 Jun 24]. p. 2497–500. Available from: https://ieeexplore.ieee.org/abstract/document/6346471
12. Oken BS, Orhan U, Roark B, Erdogmus D, Fowler A, Mooney A, et al. Brain–Computer Interface With Language Model–Electroencephalography Fusion for Locked-In Syndrome. Neurorehabil Neural Repair. 2014 May 1;28(4):387–94.
13. Speier W, Arnold C, Lu J, Taira RK, Pouratian N. Natural language processing with dynamic classification improves P300 speller accuracy and bit rate. J Neural Eng. 2012;9(1):016004.
14. Townsend G, LaPallo BK, Boulay CB, Krusienski DJ, Frye GE, Hauser CK, et al. A novel P300-based brain–computer interface stimulus presentation paradigm: Moving beyond rows and columns. Clin Neurophysiol. 2010 Jul 1;121(7):1109–20.
15. Combaz A, Chatelle C, Robben A, Vanhoof G, Goeleven A, Thijs V, et al. A Comparison of Two Spelling Brain-Computer Interfaces Based on Visual P3 and SSVEP in Locked-In Syndrome. PLOS ONE. 2013 Sep 25;8(9):e73691.
16. Severens M, Van der Waal M, Farquhar J, Desain P. Comparing tactile and visual gaze-independent brain–computer interfaces in patients with amyotrophic lateral sclerosis and healthy users. Clin Neurophysiol. 2014 Nov 1;125(11):2297–304.
17. Memmott T, Koçanaoğulları A, Lawhead M, Klee D, Dudy S, Fried-Oken M, et al. BciPy: brain–computer interface software in Python. Brain-Comput Interfaces. 2021 Oct 2;8(4):137–53.
18. Kübler A, Holz EM, Riccio A, Zickler C, Kaufmann T, Kleih SC, et al. The User-Centered Design as Novel Perspective for Evaluating the Usability of BCI-Controlled Applications. PLOS ONE. 2014 Dec 3;9(12):e112392.

19. Branco MP, Pels EGM, Sars RH, Aarnoutse EJ, Ramsey NF, Vansteensel MJ, et al. Brain-Computer Interfaces for Communication: Preferences of Individuals With Locked-in Syndrome. Neurorehabil Neural Repair. 2021 Feb 3;35(3):267–79.
20. Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I. Language Models are Unsupervised Multitask Learners. 8(9):1.
21. Hausman DM, Woodward J. Independence, Invariance and the Causal Markov Condition. Br J Philos Sci. 1999 Dec;50(4):521–83.
22. Friedman JH. Regularized Discriminant Analysis. J Am Stat Assoc. 1989 Mar 1;84(405):165–75.
23. Orhan U, Hild KE, Erdogmus D, Roark B, Oken B, Fried-Oken M. RSVP Keyboard: An EEG Based Typing Interface. Proc IEEE Int Conf Acoust Speech Signal Process Spons Inst Electr Electron Eng Signal Process Soc ICASSP. 2012;10.1109/ICASSP.2012.6287966.
24. Knopman DS, Roberts RO, Geda YE, Pankratz VS, Christianson TJH, Petersen RC, et al. Validation of the Telephone Interview for Cognitive Status-modified in Subjects with Normal Cognition, Mild Cognitive Impairment, or Dementia. Neuroepidemiology. 2010 Jan 1;34(1):34–42.
25. Kaida K, Takahashi M, Åkerstedt T, Nakata A, Otsuka Y, Haratani T, et al. Validation of the Karolinska sleepiness scale against performance and EEG variables. Clin Neurophysiol. 2006 Jul 1;117(7):1574–81.
26. Gillberg M, Kecklund G, Åkerstedt T. Relations Between Performance and Subjective Ratings of Sleepiness During a Night Awake. Sleep. 1994 May 1;17(3):236–41.
27. Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, et al. MEG and EEG data analysis with MNE-Python. Front Neurosci [Internet]. 2013 Dec 26 [cited 2024 Jun 24];7. Available from: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2013.00267/full
28. Peters B, Mooney A, Oken B, Fried-Oken M. Soliciting BCI user experience feedback from

people with severe speech and physical impairments. Brain-Comput Interfaces. 2016 Jan 2;3(1):47–58.

29. R: The R Project for Statistical Computing [Internet]. 2024 [cited 2024 Jun 24]. Available from: https://www.r-project.org/
30. Koester HH, Simpson RC. Effectiveness and usability of Scanning Wizard software: a tool for enhancing switch scanning. Disabil Rehabil Assist Technol. 2019 Feb 17;14(2):161–71.
31. Thompson MC. Critiquing the Concept of BCI Illiteracy. Sci Eng Ethics. 2019 Aug 15;25(4):1217–33.
32. Oken B, Memmott T, Eddy B, Wiedrick J, Fried-Oken M. Vigilance state fluctuations and performance using brain-computer interface for communication. Brain Comput Interfaces Abingdon Engl. 2018;5(4):146–56.
33. Fried-Oken M, Kinsella M, Stevens I, Klein E. What stakeholders with neurodegenerative conditions value about speed and accuracy in development of BCI systems for communication. Brain-Comput Interfaces. 2024 Apr 2;11(1-2):21-32.
34. Friedman B, Hendry DG. Value Sensitive Design: Shaping Technology with Moral Imagination. MIT Press; 2019.
35. Costanza-Chock S. Design Justice: Community-Led Practices to Build the Worlds We Need. The MIT Press; 2020.