

Multimodal pre-screening can predict BCI performance variability: A novel subject-specific experimental scheme

Seyyed Bahram Borgheai^{a,c}, Alyssa Hillary Zisk^b, John McLinden^a, James McIntyre^a,
Reza Sadjadi^d, Yalda Shahriari^{a,b,*}

^a Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, Kingston, RI, United States

^b Interdisciplinary Neuroscience Program, University of Rhode Island, Kingston, RI, United States

^c Neurology Department, Emory University, Atlanta, GA, United States

^d Neurology Department, Massachusetts General Hospital, Boston, MA, United States

ARTICLE INFO

Keywords:

Electroencephalography (EEG)
Functional near-infrared spectroscopy (fNIRS)
Multivariate linear regression (MLR) model
Predictive models
General linear model (GLM)

ABSTRACT

Background: Brain-computer interface (BCI) systems currently lack the required robustness for long-term daily use due to inter- and intra-subject performance variability. In this study, we propose a novel personalized scheme for a multimodal BCI system, primarily using functional near-infrared spectroscopy (fNIRS) and electroencephalography (EEG), to identify, predict, and compensate for factors affecting competence-related and interfering factors associated with performance.

Method: 11 (out of 13 recruited) participants, including five participants with motor deficits, completed four sessions on average. During the training sessions, the subjects performed a short pre-screening phase, followed by three variations of a novel visuo-mental (VM) protocol. Features extracted from the pre-screening phase were used to construct predictive platforms using stepwise multivariate linear regression (MLR) models. In the test sessions, we employed a task-correction phase where our predictive models were used to predict the ideal task variation to maximize performance, followed by an interference-correction phase. We then investigated the associations between predicted and actual performances and evaluated the outcome of correction strategies.

Result: The predictive models resulted in respective adjusted R-squared values of 0.942, 0.724, and 0.939 for the first, second, and third variation of the task, respectively. The statistical analyses showed significant associations between the performances predicted by predictive models and the actual performances for the first two task variations, with ρ values of 0.7289 (p-value = 0.011) and 0.6970 (p-value = 0.017), respectively. For 81.82 % of the subjects, the task/workload correction stage correctly determined which task variation provided the highest accuracy, with an average performance gain of 5.18 % when applying the correction strategies.

Conclusion: Our proposed method can lead to an integrated multimodal predictive framework to compensate for BCI performance variability, particularly, for people with severe motor deficits.

1. Introduction

Despite significant progress in the development of brain-computer interface (BCI) systems, long-term use of BCIs for the purpose of providing people with severe motor impairments a means of communication remains a challenge. Enthusiasm for BCI systems diminishes when the technology is translated from the laboratory to the patients' bedsides [1]. In addition to inter-subject variability, day-to-day intra-subject BCI performance variations have been reported, limiting the robustness and efficacy of these systems [2,3]. Furthermore, most

current BCI systems fail to provide reliable performance for completely locked-in patients who have lost all voluntary muscle control, including oculomotor movements [4]. One major reason for this inefficacy is the variability of disease-related conditions and environmental factors across end-users.

To meet long-term subjective demands in a real-world environment, BCI systems must be customized with respect to individual conditions and circumstances [2]. To achieve user satisfaction, all BCI users would need to be able to freely communicate and control their environments, regardless of the degree of their motor deficits and the progression of

* Corresponding author. Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, Kingston, RI, United States.

E-mail address: yalda_shahriari@uri.edu (Y. Shahriari).

<https://doi.org/10.1016/j.combiomed.2023.107658>

Received 23 March 2023; Received in revised form 20 October 2023; Accepted 31 October 2023

Available online 2 November 2023

0010-4825/© 2023 Elsevier Ltd. All rights reserved.

their disease [5,6]. Given the variation in individual characteristics, there is current interest in subject-specific systems, where the machine learns individual characteristics and adjusts the processing pipeline to optimize the system for different persons or sessions accordingly [7]. These subject-dependent BCI systems are ultimately the most realistic solution for convenient long-term BCI use at patients' bedsides [2,8].

2. Related work

2.1. Variability factors

To compensate for intra- and inter-subject variability, the first step is to identify the major causes of these variabilities. Two major types of confounds have been identified in the literature: Competence and interfering factors.

2.1.1. Competence factors

The first type of confound involves the subject's competence factors, i.e. confounds affecting their ability to perform the tasks. These include cognitive, physiological, and neurological confounds associated with the subject's abilities to perform a BCI task, including mental states and anatomical functions [9,10]. 15–30 % of users are reported to be unable to elicit task-specific responses robust enough to control a BCI [11]. Different users may also adopt various cognitive strategies to perform the same task, which can further lead to inter-subject performance variability [12]. Moreover, time-variant cognitive confounds such as fatigue, attention, and working memory load can result in inconsistent intra-subject BCI performance [13,14]. Thus, incorporating the mental state of a user [15,16] can serve as a valuable compensative component in BCI design to account for performance variabilities. In addition to direct brain-related factors, other physiological confounds have been reported to impact BCI users' performance competence. For instance, impaired eye gaze control can negatively impact visual evoked potential-based BCI performance, including the widely used P300 Speller, a communicative electroencephalography (EEG)-based BCI system [17]. Furthermore, vital signs, such as heart rate (HR), respiration rate, or blood oxygen saturation (SpO₂), have been reported to be associated with BCI performance [18,19]. While some studies have reported individual competency confounds associated with BCI performance, there is a need to develop an integrative approach that combines several measures to construct a unified predictive framework to compensate for performance variability.

Moreover, in practical, non-laboratory settings, BCI users are not exclusively focused on the instructed task. Instead, they may experience changes in mental workload. Background distractions can also occur in parallel with the main BCI task. As a result, the isolated laboratory settings common to P3Speller studies cannot practically be assumed [16, 20]. In their study, Ke et al. (2015) introduced mental operations, such as the n-back task and mental arithmetic, alongside the conventional flashing task in P3Speller. This approach served as an interfering factor to simulate the (negative) impact of additional workload in a realistic environment [21]. Furthermore, in addition to the external non-task related workload affecting the BCI task, studies have reported variations in task-related workload among BCI users. These studies have also investigated associated neuromarkers [16]. Variations in cognitive loads or mental overloads have been shown to impact BCI performance differently, either due to different cognitive capacities in subjects or different strategies used by participants to perform the tasks [22,23]. Different levels of task difficulty have also been reported to impact BCI performance [24,25]. However, despite previous exploratory investigations into the impact of workload, few attempts have been made in the BCI field to compensate for workload-associated variations in BCI performance.

Despite reports of parallel non-motor dysfunctions, little has been done to investigate how these cognitive declines can specifically impact the BCI performance of users with severe motor deficits. For example,

verbal fluency and executive functioning impairments have been demonstrated in people with ALS [26,27], but their possible associations with BCI performance have not yet been investigated in this cohort. In one case, Perez et al. (2020) reported a significantly lower number of responses in two verbal fluency tasks in an ALS cohort compared to the control group [28], but this study did not involve a BCI component. Another ALS study using functional near-infrared spectroscopy (fNIRS) and diffusion tensor imaging (DTI) reported disease-specific associations with homotopic resting state functional connectivity [29], but did not investigate any possible impact on their BCI performance. On the other hand, Shahriari et al. (2019) showed that temporal and spectral EEG features were correlated with performance variability in a longitudinal assessment of P3 Speller performance in a cohort of ALS patients but did not investigate connections with possible cognitive impairments. Although we reported significant differences in non-motor electrophysiological and hemodynamic features between the ALS and neurotypical groups during a visuo-mental BCI task in our previous works [17, 30], the association with performance was not evaluated.

2.1.2. Interfering factors

The second confound involves interfering factors, which are confounds affecting the quality of the acquired signal independent of the tasks [13,31,32]. Besides task-specific cognitive confounds, time-varying physiological interfering factors can also contribute to both inter- and intra-subject BCI performance variability. Interfering factors are subject-related physiological confounds that affect the quality of the recorded signals independently of the subject's brain activity. For instance, in most common BCI systems that rely on EEG, physiological artifacts such as eye movements and facial muscle tension can decrease the signal-to-noise ratio (SNR) of the data and consequently diminish task-driven EEG neuromarkers [33,34]. For example, Florich et al. (2015) reported the adverse effects of muscle artifacts on imagery BCI performance [33]. Electro-oculogram (EOG)-related artifacts have also been shown to impact P3 speller performance [35]. Similarly, in BCI systems relying on hemodynamic responses, including fNIRS, the acquired responses are often heavily contaminated by superficial physiological signals such as blood pressure, respiratory, and cardiac artifacts. This contamination occurs because the optical measurement path includes layers of highly vascularized skin, the skull, and hair [36,37]. These physiological interferences can be misinterpreted as task-driven neural responses, consequently increasing both false positive and negative rates [38,39]. Therefore, considering physiological interferences is an important factor in a real-world BCI design, as it may compensate for the variability of performance variability across subjects and sessions.

2.2. Current compensative solutions

One early solution to compensate for performance variability was to conduct training sessions prior to the main tasks/sessions, aimed at improving the subject's ability to perform the task [40]. However, the required training sessions and runs for each use of the BCI are tedious [41]. Additionally, repeating the same task may increase the risk of losing the response strength due to habituation effect [42]. Therefore, new BCI designs equipped with machine learning techniques were proposed to reduce the need for excessive training sessions, wherein a model is calibrated based on data collected at the beginning of each session [43,44]. For instance, transfer learning techniques have been proposed to leverage very few training samples from the target group [45,46] before proceeding to the actual experiment [47,48]. Blankertz et al. (2001) reported that resting-state alpha activity can be associated with performance in subsequent motor imagery (MI)-based BCI tasks. Similarly, Sannelli et al. (2019) found that resting-state EEG markers can be used to categorize sensorimotor rhythm (SMR)-based BCI users, enabling them to adapt to a better subject-dependent training strategy [49]. Subsequently, the constructed inter-subject performance

predictors, sometimes referred to as inter-subject associativity [41], could be incorporated into BCI design [45,50] to enhance BCI system efficacy [51,52]. Although markers extracted from calibration recordings at the beginning of a session have been shown to be associated with performance in subsequent BCI tasks within the same session, an integrative experimental scheme that uses those individual markers acquired prior to the main BCI tasks in a predictive framework to compensate for the upcoming BCI performance variability has rarely been introduced.

On the other hand, automated removal of interfering factors has also been investigated to improve BCI performance. Thompson et al. (2019) tried ten different automated EOG removal methods to compensate for P3 speller performance, though the tested methods in fact all reduced performance [35]. Integrated EEG + EOG designs have also been proposed to compensate for ocular artifacts in BCI systems [53,54]. Similarly, other types of hybrid BCI designs [55], such as electromyography (EMG) + EEG [56] or EEG + fNIRS [17,57] have been introduced to integrate the compensative nature of different recording modalities to improve performance. While the combination of different pairs of modalities has shown benefits in the BCI system, little work has been done to investigate the advantages of multimodal designs in which more than two modalities are integrated. Further, lacking the consideration of subject- and session-specific variabilities in performing BCI tasks makes these designs less robust for long-term daily use.

Here, we have proposed a customizable multimodal BCI design to address performance variability in both neurotypical BCI users and people with severe motor deficits. Our goal in this study is to design and validate an integrative subject-dependent closed-loop predictive scheme that identifies markers affecting subjects' performance, predicts possible variabilities in upcoming sessions, and subsequently employs correction strategies to compensate for the identified sources of variability, i.e. competence and interfering factors. Accordingly, we propose a novel experimental scheme in which a quick pre-test screen (pre-screen) is conducted before the main BCI protocol to construct a predictive platform that enhances the robustness of BCI systems. This is achieved by exploring the associations of BCI performance variation with the markers extracted from the pre-screen. To reduce variability in the main BCI task performance, we have provided new variations of a single trial fNIRS-EEG BCI paradigm by extending a novel visuo-mental (VM) protocol previously shown to be effective in facilitating communication in people with severe motor deficits [17,30]. Adopting such a paradigm offers several advantages, including: a) providing a feasible BCI communication platform where conventional BCIs fail, specifically for patients without fine eye-gaze control; b) identifying potential cognitive issues in people with motor deficits due to the dual-task nature of the paradigm; c) better simulating the interfering non-task related workloads which makes the setting more reflective of patients' daily needs; and d) providing the flexibility to adjust the task-related workload using different variations of the VM task. These steps enables the predictive platform to adapt to intra- and inter-subject variations in task performance capacity by augmenting the visuo-spatial P3Speller task with a controllable mental arithmetic component.

Our proposed predictive platform employs two correction strategies. First, a task correction strategy is employed, wherein the best task variation with the highest performance for a specific subject and session is selected. This selection relies on the features extracted from the pre-screen. Then, the platform employs a subject-dependent interference correction strategy to compensate for the impact of physiological interference. Unlike most BCI studies, which conduct preprocessing for correction of interfering confounds without considering the subjects' status or the type of BCI tasks, the correction strategies used here are applied with respect to the task-specific and subject-dependent terms appearing in the corresponding predictive models. The outcomes of the proposed experimental scheme from this study can advance the understanding of intra- and inter-subjective BCI performance variabilities and provide personalized correction strategies necessary for developing

long-term personalized BCI systems. These systems would rely on embedded adaptive correction strategies tailored to each user's status.

3. Subjects, materials, and methods

3.1. Participants

13 participants (5 female) were recruited for this study, with an average age of 57.85 ± 6.00 years old. The participant group consisted of six diagnosed with amyotrophic lateral sclerosis (ALS) (age: 57.71 ± 6.80 years old, 2 female) and seven healthy (H) subjects with no known motor deficit (age: 58.5 ± 5.55 , 3 females). All participants acknowledged no history of non-motor deficits including visual, mental, or substance-related disorders, that would otherwise impact their ability to perform the experimental tasks. Demographic and clinical information, including age, sex, disease duration, disability score, medication, and education level, is listed in Table 1. All participants except for one healthy subject (H04) had at least some post-secondary education. The Revised ALS Functional Rating Scale (ALSFRS-R) scores, a validated screen for the dysfunctional progression of the ALS disease, averaged 26.17 ± 15.36 (min 0, and max 42) on a 48-point scale. The highest score [48] reflects normal function in activities of daily living (ADL), and the lowest score (0) represents a complete loss of function [58]. Patient disease durations were 6.3 ± 3.4 years on average. One patient (ALS-05) had both gastrostomies and tracheostomies. One of the healthy (H-07) and one of the ALS (ALS-06) participants withdrew from the study after finishing their familiarization session. Others completed on average 4.0 ± 0.8 sessions of recording, including at least one familiarization and one test session. All procedures were in accordance with the study protocol approved by the Institutional Review Board (IRB) of the University of Rhode Island (URI). All participants provided informed consent or assent prior to the experiment and were financially compensated. All participants in the ALS group were tested in either their homes or care centers, while the healthy cohort participated in the experiments at the NeuralPC lab at URI.

To construct our predictive models, we integrated the training data of all subjects without considering the motor deficits in our participants. The rationale behind this approach was to avoid imposing any a priori disease-related hypothesis about the impact of motor deficits on building our predictive models. Instead, we allowed the model itself to reveal disease-related confounds that could potentially reflect the possible impact of motor deficit on the performance variability. In other words, if there are any neurophysiological and cognitive differences between the two groups, these differences should be reflected in the features we extract and use in the model. The ideal predictive models should incorporate all the important features and mirror these differences. Our ideal model should provide predictive power for both groups.

3.2. The proposed experimental scheme overview

Fig. 1 illustrates the proposed experimental scheme, which centers around conducting a pre-screen recording prior to the main BCI task based. This pre-screen recording enables us to predict performance (variability) in the upcoming BCI run and, based on the prediction, implement appropriate performance correction strategies. Each subject participated in 4 (3–5) sessions of recording on average, including at least one familiarization session, two training sessions (except for ALS01), and one test session. For BCI tasks, our design focused on our proposed VM paradigm, which has been previously demonstrated as a feasible and effective paradigm for communication purposes, especially for people with severe motor deficits [17]. To achieve this, we designed three different variations of the VM task, aiming to address individual differences in task performance and to account for potential session-to-session cognitive variations that might impact the participants' competence in BCI task performance. As a result, we constructed three predictive models, each based on the features extracted from the

Table 1
Participant's demographic information.

	ALSFRS-R	Duration (years)	Age (years)	Gender	Post- Secondary Education	Medication	#Sessions
H01	N/A	0	44	M	Y	N	4
H02	N/A	0	60	M	Y	N	4
H03	N/A	0	62	F	Y	N	4
H04	N/A	0	57	M	N	N	4
H05	N/A	0	64	F	Y	N	4
H06	N/A	0	55	M	Y	N	4
ALS01	42	4	56	F	Y	Y	3
ALS02	32	12	58	M	Y	Y	4
ALS03	29	4	69	M	Y	N	4
ALS04	37	4	54	M	Y	Y	5
ALS05	0	5	56	F	Y	N	5

pre-screen runs in the training sessions. These models provided three different performance predictions for each of the three VM task variations. In the test session, we adopted two phases of correction. First, we applied task correction, and subsequently interference correction (correction related to physiological interfering factors), as explained in the following sections with more comprehensive details.

3.2.1. Training sessions

As illustrated in Fig. 1 (Top), each training session consisted of two sections: a pre-screen and the main BCI tasks, as explained below.

3.2.1.1. Pre-screening phase. Integrating our previous works [30,59,60], with other similar BCI studies [10,19,48], we incorporated three steps in our pre-screening phase: a cognitive questionnaire, a resting-state recording, and a standard P3Speller recording. The components of the pre-screen are as follows: 1) Cognitive questionnaire: To assess the internal cognitive status of the end-user, the first step is to identify the appropriate cognitive screens reported to be associated with cognitive impairments, particularly for people with motor deficits. For the cognitive screens, we adopted the Cognitive Behavioral Screening (CBS) test, a shorter cognitive battery developed specifically for ALS patients as one of the end-users for our proposed experimental scheme. The CBS was conducted at the beginning of each run to assess attention, concentration, tracking/monitoring, and linguistic initiation and retrieval (fluency test). In instances where our participants were unable to answer the CBS questions orally, they used their own eye-tracking systems to answer the questions. 2) Resting-state: All subjects were instructed to focus on a dot on the screen while in a relaxed state. In each run, 5 min of resting state fNIRS-EEG data were acquired, which provided a sufficient recording duration to obtain robust functional connectivity in resting-state brain networks [61]. 3) P3Speller: All subjects performed the standard P3Speller paradigm, where a 6×6 matrix of letters was used, and each row/column was randomly intensified with a picture of a celebrity face superimposed over the letters for 93.75 ms, followed by a 62.5 ms inter-stimulus-interval (ISI). An eye tracking device (TOBII) was used to monitor participants' eye gaze during the pre-screen P3Speller task.

3.2.1.2. Visuo-mental (VM) task variations. Extending our previous works, where we demonstrated the feasibility and effectiveness of using our proposed visuo-mental (VM) paradigm for communicative BCI communication, especially for people with severe motor deficits [17], we developed three variations of our previously introduced VM paradigm to include varying levels of mental task workload as follows (see Fig. 2 (B)): 1) VM task variation 1 (VM-v1): This variation extended the conventional oddball P3Speller paradigm, by displaying a 2×2 matrix of digits (1–9) over the intensified letter. Each subject was instructed to focus on a target character, resulting in 14 targets per run. Upon each target intensification, subjects were instructed to perform predefined mental arithmetic tasks. Specifically, they were asked to add pairs of numbers in the matrix either diagonally (at the first target flash) or

vertically (at the second target flash), and then double the larger result. The stimulation intensification time was set to 300 ms, and each intensification was followed by a 5-s inter-stimulus interval (ISI). 2) VM task variation 2 (VM-v2): At each target intensification, instead of displaying a 2×2 matrix, a 2-digit number within the range of 11–29 appeared over the letters. Subjects were required to start with a given 3-digit number and continuously subtract the numbers that appeared over the specified target. They subtracted the number presented as the current target intensification from the result of the previous target intensification. 3) VM task variation 3 (VM-v3): At each target intensification, the subject observed a 2×2 matrix containing four single-digit numbers. After the first flash, the task was to add up the upper row, compare the sum with the number seven, and multiply the larger value by two. After the second flash, the subjects were instructed to add up the lower row, compare it with the number seven, and multiply the larger value by two. To avoid order effects, we counterbalanced the order of the task variations within each session.

3.2.2. Test session

Fig. 1 (Bottom) illustrates the experimental flow in the test sessions. In the test session, the pre-screen was conducted following the same procedure as previously described. Then, the predictive model was fed features extracted from the pre-screen to predict which variation of the task would yield the highest performance for each specific subject. Based on this prediction, we selected the optimum VM task variation (VM-Vmax) for the test session of each subject. Afterward, the subject performed the chosen VM task variation, and the customized correction strategies were applied based on the interfering terms that appeared in the predictive model of the selected variation. For instance, if a blink related feature appeared in the predictive model for a task variation, we applied blink correction to the corresponding data.

3.3. Data acquisition

EEG data were recorded simultaneously using a g.USBamp amplifier (g.tec Medical Tech.) and digitized at 256 Hz. fNIRS data were recorded using a NIRScout (NIRx Inc.) with two NIR lights (760 nm and 850 nm wavelengths) and digitized at 7.81 Hz. Fig. 2 (A) shows a schematic head model of the fNIRS-EEG sensors' placement. As depicted in this figure, eight emitters and seven (long-channel) detectors acquired 14 fNIRS channels covering the pre/frontal region and two channels on temporo-parietal areas to capture higher cognitive functions associated with mathematical operation paradigms. To cancel hemodynamic-related systemic noise, we used an 8x short-channel detector measurement bundle with each detector mounted around each emitter, all connected to a single detector attached to the back of the cap. The distances between source and detector were 30 and 7.5 mm for long-distance (LD) and short-distance (SD) channels, respectively. Following the Modified Combinatorial Nomenclature (MCN) montage, emitters were placed at Fz, F3, F4, AF3, AF4, Fp1, Fp2, CP5, and CP6 and detectors at F1, F2, AFz, Fp1, Fp2, P5, and P6. EEG was recorded from 16 channels: Fz*, F5*,

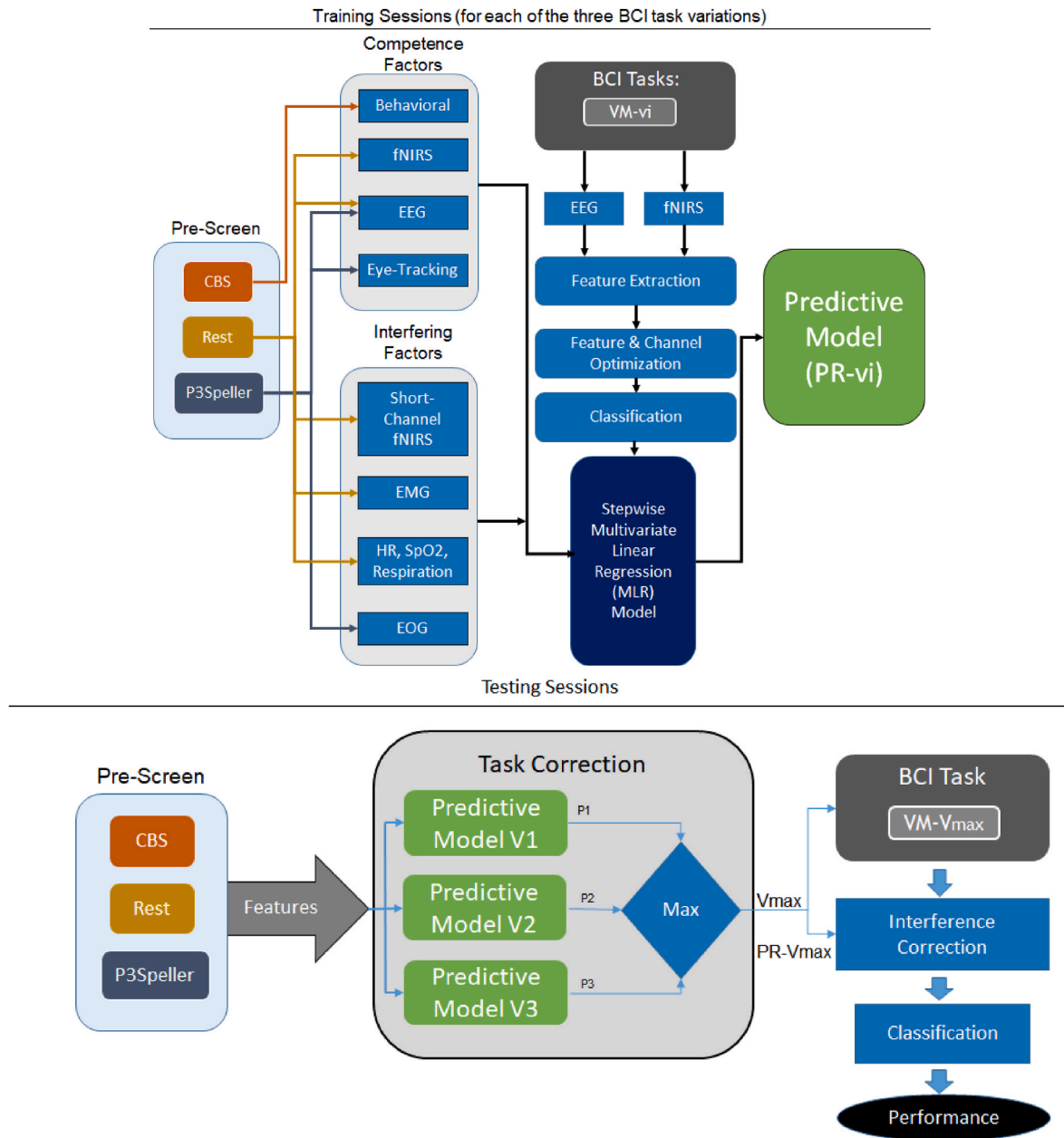


Fig. 1. The Proposed Experimental Scheme. The top part demonstrates the task flow in the training sessions. Firstly, a pre-screening phase consisting of cognitive-behavioral screening (CBS), a resting run, and P3Speller is conducted prior to the main BCI tasks. Next, three variations of the VM task are performed, and a predictive model is constructed for each of the variations based on features extracted from the pre-screening phase along with the performances of each variation. The bottom part demonstrates the experimental flow for the test session. The features extracted from the pre-screening phase are fed to the previously constructed predictive models to provide the prediction for performances (P1, 2, and 3). The task variation with the highest predicted performance, i.e. VM-Vmax, is then selected to be run (task correction). Subsequently, an appropriate interference correction strategy is applied concerning the terms appearing in the predictive model for the selected variation (PR-Vmax).

F6*, Fp1*, Fp2*, Cz, P3, Pz, P4, T7, T8, P7, P8, PO7, PO8, and Oz to include all commonly used channels in conventional P300 paradigms as well as pre/frontal and temporal channels to capture possible responses to mental aspects of the VM tasks (star (*) denotes the nearest electrode placement to fNIRS occupied channels according to the 128-channel montage). Extending our previous works [17,30], this montage was intended to capture both aspects of the visuo-mental task variation as well as resting-state activity and activity related to the P3S in the pre-screen. All experimental protocols, data acquisition, and stimulus presentation labels were controlled using BCI2000 and NIRStar software.

Additionally, simultaneous to neuroimaging modality recordings (i.

e. EEG and fNIRS), we measured multiple vital and potential interfering signals as shown in Fig. 2 (B). For this purpose, HR and SpO2 were recorded using a g.SpO2sensor (g.tec Medical Tech.) with pulse frequency 0–300 bpm attached to g.USBamp. For respiration rate, we used the Respiration Effort Sensor (g.tec Medical Tech.) attached to a chest-belt to measure changes of the circumference of the torso related to breathing activity. Furthermore, to capture physiological interfering factors contaminating EEG signal, we recorded four facial electromyogram (EMG) channels and one electro-oculogram (EOG) channel simultaneous to EEG and fNIRS signals using the g.USBamp amplifier and digitized at 256 Hz. In the P3S run, to assess the level of eye-gaze control in the subjects, we bundled the Tobii Pro Nano eye-tracking

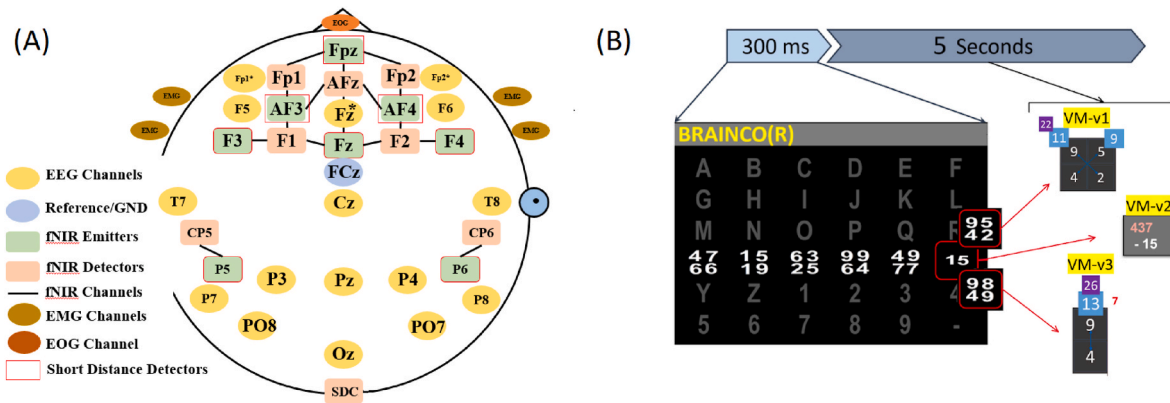


Fig. 2. (A) Schematic head model showing placements of 16 EEG, four EMG, and one EOG electrodes, as well as fNIRS optodes, including eight sources, eight long-distance (LD), and eight short-distance (SD) detectors. One of the long-distance detectors was placed on the back of the head to collect all short-distance channel data, here, labelled as short-distance collector (SDC). (B) The visuo-mental (VM) task with 300 ms intensification time and 5-s inter-stimulus interval (ISI) with three variations, VM-v1, VM-v2, and VM-v3.

device with our EEG recording system through BCI2000. In this run, due to the interference between fNIRS emitters and the eye-tracking system, we did not record any hemodynamic data along with EEG.

3.4. Data analysis

3.4.1. Pre-screening feature extraction

The pre-screen features were extracted in two phases: competency features (36 in total) and interfering features (16 in total), as described below.

3.4.1.1. Competency features. Cognitive and Behavioral Screen (CBS): Individual scores in each category, including 1) Attention, 2) Concentration, 3) Tracking/monitoring, and 4) Fluency, i.e. linguistic initiation and retrieval (fluency test), and 5) Sum of these individual scores, resulting in a total of 5 features for each session/subject.

Resting-State: Replicating our previous work [59], 30 s of the 5-min segments were cut from the beginning and end of data to avoid any potential non-relaxing state of the subject in the beginning and the boredom in the end, resulting in 4-min of data. Eye-blinks in the frontal EEG channels were, first, extracted using cues of EOG signal peaks, using MATLAB's *findpeaks* function, with prominence over 80% and within the range of its width, and then interpolated. Similar to another study [62], the prominence threshold was selected to guarantee that only peaks with high prominence could be selected, as eye-blinks are typically the strongest artifacts on frontal EEG channels. Then, EEG power spectra were extracted from the Delta (1–3 Hz), Theta (4–7 Hz), Alpha (8–12 Hz), and Beta (13–30 Hz) frequency bands using a set of 30 complex Morlet wavelets ranging from 1 to 30 Hz with 3–10 cycles and a 1 s time-window were used for time-frequency decomposition. To reduce feature space dimensionality, spectral power in each band was averaged over four regions of interest (ROIs): Frontal (F), including Fz, F5 and F6, Centro-Parietal (CP), including Cz, P3, Pz, and P4, Temporo-Parietal (TP), including T7, T8, P7, and P8, and Occipital (O), including PO7, PO8, and Oz. In total, we extracted 16 EEG features (4 (power-band) * 4 (ROI)) denoted as band-ROI (e.g., Alpha-F referring to frontal Alpha) for each session/subject.

To extract fNIRS features, the nirs-toolbox was used to deploy an autoregressive integrative (ARI) algorithm to remove statistical outliers, including both spike and shift variations of motion artifacts in fNIRS data [63]. Oxygenated hemoglobin (HbO2) and deoxygenated hemoglobin (HbR) concentration changes were extracted from raw optical intensity data using the modified Beer-Lambert law [64,65]. Then, we applied short-separation filtering to the hemodynamic responses [66] through nirs-toolbox. SD channel data were projected out of the LD

channels to bolster the brain signals [37,63]. Then, the outcome was bandpass-filtered using a 0.01–0.09 Hz passband commonly used in resting-state fNIRS studies [59,67] to remove higher frequency physiological artifacts, including respiratory artifacts (0.2–0.3 Hz) [68], cardiac signals (0.8–1.3 Hz), and Mayer waves (~0.1 Hz). In contrast to EEG, spectral analysis of fNIRS data has been shown to reflect little non-motor functional dynamics of the brain [59]. Thus, in this study, we adopted correlation-based connectivity analysis relying on time samples for our fNIRS data. To reduce dimensionality, we first averaged the hemodynamic responses across channels in four ROIs: left frontal (LF), right frontal (RF), left temporal (LT), and right temporal (RT), and then calculated Pearson's correlation coefficient for all pairs of ROIs, constituting five pair connectivity features (i.e., RF-LF, RF-RT, RF-LT, LF-RT, LF-LT). We did not include inter-temporal connectivity (RT-LT) in our connectivity feature lists, as it was rarely reported to be cognitively interesting, particularly in people with motor deficits. Given that the HbO2 signal has been shown to better characterize resting-state blood flow dynamics than the HbR signal, and that significant connectivity results in similar studies are primarily HbO2-based [69,70] we focused on HbO2 features. Additionally, the mean of (resting) HR, SpO2, and respiration rate were extracted from the data recorded during the resting-state task.

P3Speller: Replicating our previous work [60], features related to performance were extracted from EEG data. To calculate the performances, stepwise linear discriminant analysis (SWLDA) classifiers were used to derive features related to P3Speller performance. Through forward and backward stepwise regression using the *fitdiscr* and *stepwisefit* functions in MATLAB, the best predictors ($p < 0.1$) were selected and the least significant predictors ($p > 0.15$) were removed. This procedure was repeated for up to 60 steps, or until no additional terms satisfied the entry/removal criteria [71]. However, compared to neurotypical users, people with motor deficits, particularly those with ALS, are known to exhibit trial-by-trial latency variability in their P300 responses, commonly referred to as latency jitter [60,72]. Accordingly, we included jitter in our P3Speller feature list as it has been reported to be negatively correlated with BCI Performance [72,73]. To extract jitter-related features, we adopted the approach proposed by Thompson and colleagues [73] known as classifier-based latency estimation (CBLE). In CBLE, a classifier is trained as usual, and then time-shifted epochs are fed to this classifier. For instance, if the 0–800 ms post-stimulus epoch is used for classification, then time-shifted epochs from the –100 ms–700 ms epoch through the 100 ms–900 ms epoch could be classified. Then, for each target flash, the time shift corresponding to the highest classifier score (the probability that the flash was a target flash) is extracted as the latency shift for that particular flash. This approach allows us to estimate

the latency jitter for each target flash during the P3Speller task. The variance of these latency shifts within a session, denoted as vCBLE, reliably measures latency jitter [73]. In sum, the following features were extracted from P3Speller runs: 1) vCBLET: A measure of latency jitter using the variance of latency shifts for target epochs calculated with CBLE; 2) QDJitter: Another measure of latency jitter calculated as the difference between third and first quartile (Q3 and Q1) of latency shifts for target epochs using CBLE; 3) P3S-Acc: Stimulation (target) accuracy (representing correctly classified epochs); 4) Default Precision: Flash/-target precision (the number of correctly classified targets divided by the number of all returned results); and 5) 1stFlash-Acc: The character selection accuracy using the first trial epochs (both row and column flashes) per character. To assess subjects' gaze-control, along with P3Speller, we used an eye-tracking system and accordingly extracted two features: 1) Eye-Acc: The accuracy of target selection based on the eye-gaze data, and 2) Eye-msd: The mean of standard deviation of eye-gaze data during each character's intensifications.

3.4.1.2. Interfering factors. EEG Interfering Factors: As mentioned before, we recorded two types of cue signals to identify the EEG interfering factors using EOG and EMG. For EMG, we selected the resting-state runs to extract the relevant interfering features to target non-task-related facial muscle tensions. In contrast, for EOG, we wanted to extract variables to quantify the representative blinking behavior of the subject during the VM task. Eye blink metrics from the resting state data were considered non-representative of behavior during the VM task since subjects were instructed to stare at a point in the former task, which would reduce the number of the blinks. Two features were extracted from the EOG signal recorded during P3 Speller task, including the total number of blinks (#blinks), and the mean (excessive) blink artifact peak amplitude. For EMG, first, the EMG signal recorded during the resting-state task was segmented into two bandpassed frequency ranges of low frequency (EMG-LF, ≤ 20 Hz) and high frequency (EMG-HF, $20 < f < 100$ Hz). Then, simple square integral (SSI) [74] and variance of the entire time course of both EMG-LF and EMG-HF were extracted as EMG features.

fNIRS Interfering Factors: Two major cues to represent fNIRS interfering factors were extracted from the resting-state signal (segmented into a 4-min window as mentioned in the resting-state feature extraction section): 1) Peak-to-peak (LD-p2p): the sum of the difference between the maximum and minimum of the HbO2 signal in all LD channels, and 2) Mean of the short-separation (SS) contribution coefficients: Two methods were used to calculate the SS contribution to the hemodynamic responses in the LD channels, considering both HbO2 and HbR, as both are important in the SS context. In the first method, the contribution was calculated through regression as below:

$$\alpha_i = \langle SD_i, LD_i \rangle / \langle SD_i, SD_i \rangle \quad (1)$$

where $\langle . \rangle$ operator is the dot product, α_i is a temporal correlation factor between each (LD) channel i and associated SD, LD_i is the hemodynamic response of the long-distance channel (i), and SD_i is the short-distance (SD) channel attributed to the emitter of corresponding LD channel (i) [39]. We calculated the mean of α_i over all LD channels for both HbO2 and HbR signals. In the second method, we used a combination of principal component analysis (PCA) and general linear model (GLM) regression analysis to extract the contribution coefficients [75]. First, up to two of the first principal components (PCs) of the SD channels were fed to GLM regression. The component selection was based on the elbow criterion. In cases where a short channel was not useable due to poor connection in the calibration phase, the SD signal was replaced with the average of other SD channels. For constructing the GLM model we used the regress function in MATLAB:

$$LD_i = X * \beta_i + \varepsilon \quad (2)$$

where β_i is the contribution of short-channel PCs in the LD $_i$ channels

data, X is the matrix of two short-distance PCs with a constant column 1 for the intercept term, and ε is the residual error of the regression. We calculated the mean of β_i s over all LD channels for both HbO2 and HbR signals.

3.4.2. Performance calculations

For each VM task variation, we calculated the performance similar to the method used in our previous work [17] with several extensions. First, we focused on hybrid (fNIRS-EEG) linear discriminant analysis (LDA) classification by first extracting seven EEG and 17 fNIRS features, as described in our previous work [17], from each target and non-target epoch for each channel. To balance the number of features from both modalities, we selected the optimum five EEG and two fNIRS channels to retain 35 ($=7*5$) EEG and 34 ($=17*2$) fNIRS features. For EEG, the optimum channels were selected using Pearson correlation between target/non-target epochs and the class labels. For fNIRS channel optimization, we used the GLM-statistical parametric mapping (SPM) method as explained in our previous work for each fold [17]. We performed 3-fold cross-validation for classification, since we required the continuous signal to apply the GLM regression method—with seven total characters to-be-spelled. Using a 70 % ratio for the training set (five characters in each fold for training and the remaining for the test), only three training sets could be constituted: the first, middle, and last five characters. For each training fold, we repeated the GLM-SPM process and extracted fold-specific optimum channels. Then, the EEG and fNIRS features were concatenated and 50 % of the highest correlated features (Pearson correlation) were selected as optimum features. To build up a fair comparison between the task variation performances and also to exclude lower than chance level results, we picked the highest accuracies with the corresponding maximum area under the curve (AUC) value as the task's performance metric to be used further in constructing the proposed predictive model.

3.4.3. Predictive model

As illustrated in Fig. 1 (Top), after extracting the pre-screen features and performances in all training sessions, we constructed the predictive model for each task variation resulting in a total of three predictive models for each of the three VM task variations. The predictive model of each task variation (VM-v1,2, and 3) was constructed based on the training sessions of all subjects in two stages, including the predictive model (PR) built based on both competency and interfering features, hereinafter known as PR-CI. The predictive model relating the pre-screen features to each task variation performance was constructed through the following multivariate linear regression (MLR) model:

$$P_i = \beta_{i0} + \beta_{i1} * X + \varepsilon \quad (3)$$

where P_i is the performance i th VM task variation, β_{i0} is the intercept term, β_{i1} is the regression coefficient for task variation i , ε is the residual error, and X is the matrix containing the pre-screen features, including competence and the interfering factors. We used stepwise MLR algorithm through the stepwiselm function in MATLAB. Stepwise multivariate regression uses a systematic method to add and remove terms from a linear model based on their statistical significance in explaining the response variable (Fig. 3). It starts with a constant (intercept-only) model and then automatically adds to or trims the model with respect to a criterion, which we set to be the p-value (<0.05 for inclusion and >0.1 for exclusion) for the F-test of the change in the sum of squared error (SSE) that results from adding or removing a term. Terms from the set of variable terms not in the model with the smallest p-value are added iteratively subject to an entrance tolerance p-value threshold. Similarly, if any of the available terms in the model have p-values greater than an exit tolerance, they will be removed with the largest p-value and return to the above step. We did not consider any limit for the number of steps.

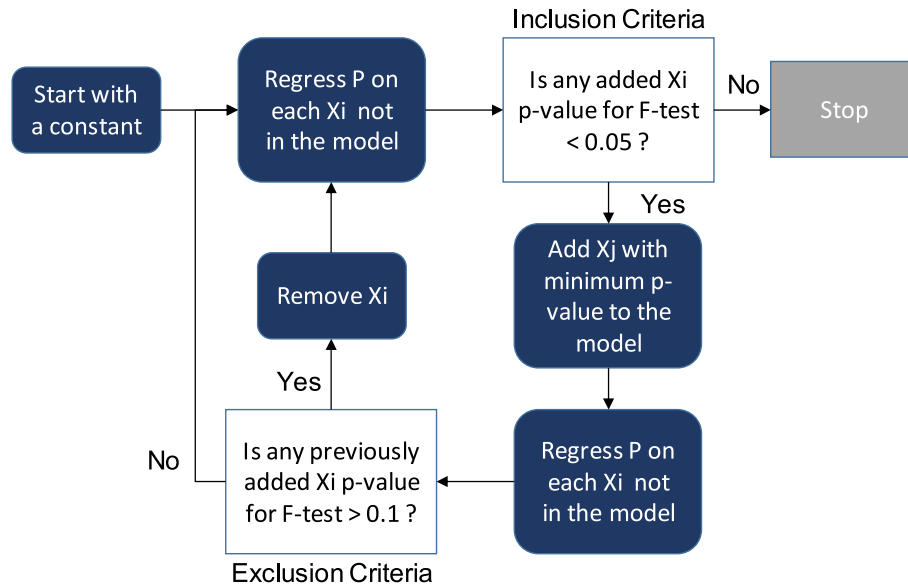


Fig. 3. The flowchart of recursive feature selection through stepwise multivariate linear regression (MLR). P is the performance of VM task and each Xi refers to a pre-screen feature. Inclusion and exclusion criteria determine if a feature would be added or removed from the model.

3.4.4. Correction strategies

With respect to the two major sources affecting BCI performance, i.e. competence and interference, two correction strategies were taken accordingly: task/workload correction and interference correction.

3.4.4.1. Task correction. As explained before and illustrated in Fig. 1 (Bottom), the variations in the task have been designed to compensate for variabilities in the cognitive- and neural-related confounds affecting the competence of the subjects to perform a certain BCI task. Therefore, to predict which variation of the VM task leads to the highest performance for a specific subject in a specific test session, we used the three models previously trained with features collected from the pre-screen to predict task performance. The variation with the highest predicted performance was selected as the optimum task variation for a specific subject for that specific subject's test session.

3.4.4.2. Interference correction. To compensate for the impact of physiological interfering factors, we applied different correction methods to the signal acquired during the selected task variation, i.e. the outcome of the task correction phase. The strategies listed below were adopted based on the presence of the relevant terms in the predictive model corresponding to the selected optimum task variation. For example, if the predicted model of the selected task variation included QDJitter appeared in its model terms, latency jitter would be identified as an important factor affecting performance. Consequently, a jitter correction strategy was applied to the corresponding data in this specific instance.

Correction for EEG Interfering Factors: Jitter Correction: to compensate for jitter or variance of event-related potential (ERP) latencies, we extended the window range for extraction of EEG-ERP features by adding the QDJitter value to both ends of the corresponding ERP window. First, we extracted QDJitter from the P3Speller run during the pre-screening phase, by computing the difference between the third and first quartiles (Q3 and Q1) of latency shifts. These latency shifts were calculated for target epochs using CBLE [73], as explained in more detail in the pre-screening feature extraction section (Section 3.4.1). Then, we extended the time windows used to extract ERPs by the QDJitter value at both ends for the performance calculation of the primary BCI tasks. Consistent with our previous works [17,30], we initially employed the following time window: 250–400 ms post-stimulus for calculating the P300 component, 350 and 560 ms post-stimulus for the

N400, and 650–800 ms post-stimulus for the P600 component. Thus, if a participant QDJitter value were 20 ms, then the time windows would be extended to 230–420 ms post-stimulus for the P300, 330–580 ms post-stimulus for the N400, and 630 to 820 post-stimulus for the P600.

Blink and EMG Removal: We used Independent Component Analysis (ICA) for blink and EMG removal using EEGLAB fastica algorithm. The components with the highest correlation with the EOG and EMG signals were removed and the denoised EEG was reconstructed based on the remaining components. An example of an EEG signal, before and after blink removal, is illustrated in the supplementary material (Fig. S5).

Correction for fNIRS Interfering Factors: Similar to what was explained in the resting-state analysis section (Section 5.3.1.1), we applied short-separation filtering to the hemodynamic responses using the ShortDistanceFilter function in nirs_toolbox. SD channel data were projected out of the LD channels to bolster the brain signals. An example of corrected HbO2 t-scores, before and after short-separation filtering, is illustrated in the supplementary material (Fig. S6).

3.4.5. Statistical analysis

For the evaluation of the predictive model, we calculated the root mean squared error (RMSE), R-squared, Adjusted R-squared, and the statistical significance (p-values) of each selected term in the model in addition to the F-statistic evaluating if the overall model provides a better fit to the data in comparison to a model that contains no independent variables. Additionally, the standard error (SE), t-score (tStat), and p-value of the coefficients appearing in the models were calculated for statistical evaluation. We calculated the variance inflation factor (VIF) metric by deriving the diagonal elements of the inverse of the correlation matrix of each model's independent variables [76] to assess the impact of multicollinearity on the models [77,78].

To evaluate the task correction outcomes, we performed all the task variations within the test session and then compared the predicted performances with the actual performances in each variation of the task. Accordingly, we assessed if the variation selected by the predictive models corresponded with the highest performance in the actual runs in the test session. Additionally, for each variation, we ran a non-parametric Spearman correlation analysis to investigate the associations between predicted performances and the actual performances across all subjects.

To investigate any significant difference in the predictors' values

between subjects without and with motor deficit (i.e., ALS), the between-group intra-subject average measures, i.e. the average of each predictor's value over all sessions of each subject in each group, were compared using Mann-Whitney U tests (Wilcoxon rank-sum test), appropriate for non-normal distributions and small, potentially uneven samples.

4. Results

4.1. Predictive models

Table 2 displays the predictive models constructed for each task variation (VM-v1, VM-v2, and VM-v3) using the training sessions of all subjects in two stages. In the first stage, we built the predictive model (PR) based solely on competency features, referred to as PR-C. Subsequently, in the second stage, we extended the model to include both competency and interfering features, namely PR-CI. The rationale behind separating these steps was that the interfering factors may, hypothetically, equally impact the tasks performed within the same session as they are present in all of the task variations. Therefore, we constructed models in two steps to investigate the influence of interfering factors on the predictive power of our models. To streamline the presentation of results, we reported the final predictive models, i.e. PR-CI, in the main body of the text. The model that does not include interfering factors (PR-C models) was included in the supplementary materials.

4.1.1. Task variation 1 (VM-v1)

Table 3 illustrates the terms and estimated coefficients (β) as well as their corresponding standard error (SE), t-score (tStat) and p-value for the predictive model, i.e., PR-CI, of the first task variation. In the PR-CI model, in addition to attention, RF-RT, and Eye-msd, Delta-TP from EEG resting-state, P3S-Acc from P3S Speller remained in the model. Additionally, the number of blinks (#blink) feature from EOG and LD-p2p feature related to fNIRS interfering features together with the interaction term between LD-p2p and RF-RT appeared in the model. The model's RMSE was 0.0154, R-squared and Adjusted R-squared values were respectively 0.967 and 0.942 for PR-CI model, and the model's F-

score was 39.8 (p-value $\ll 0.01$). VIF values were all less than 5 which raised no multicollinearity concern.

Fig. 4 (and Table S2) shows the predicted accuracies using PR-CI model versus the actual performance in the VM-v1 test session across all subjects. The spearman correlation analysis showed significant associations between the performance predicted by the PR-CI model and the actual performance with a rho of 0.7289 (p-value = 0.011) in the VM-v1 task.

4.1.2. Task variation 2 (VM-v2)

Table 4 shows the terms and estimated coefficients as well as their corresponding SE, t-score (tStat), and p-value of the proposed PR-CI predictive model of the second task variation. In this model, fluency from CBS test, LF-LT from fNIRS resting-state connectivity, QDJitter and P3S-Acc from P3S Speller remained in the model. As mentioned, none of the interference features appeared in the model. The RMSE was 0.0343, and R-squared and Adjusted R-squared values were 0.785 and 0.724 for the PR-CI model, respectively. The F-score as opposed to the constant model was 12.8 (p-value $\ll 0.01$). VIF values for QDJitter and P3S-Acc were marginally greater than five, which should be considered cautiously due to multicollinearity concerns. However, as multicollinearity does not in principle affect prediction accuracy, we postponed exclusion of either of those factors until checking the ultimate predictive power of the model.

Fig. 5 (and Table S4) shows the predicted accuracies using the proposed PR-CI model versus the actual performances in the VM-v2 test session across all subjects. The correlation analysis showed significant associations between the performances predicted by the model and actual performances with a rho of 0.6970 (p-value = 0.017) in the VM-v2 task. With the significant association of the predictive performances with actual ones, none of the predictors were excluded from the model due to the multicollinearity concern.

4.1.3. Task variation 3 (VM-v3)

Table 5 demonstrates the terms and estimated coefficients as well as their corresponding standard error (SE), t-score (tStat) and p-value of the proposed predictive model of the third task variation. From the competence features, concentration from the CBS battery, HR from the vital recordings, LF-RF from fNIRS connectivity measurements, QDJitter from the P3S speller features, and Eye-msd from the eye-tracking features, remained in the model. In addition to the competence terms, the LD-p2p term from the fNIRS interfering features and an interaction term between HR and Eye-msd were added to the model. The RMSE was 0.0147 for the proposed PR-CI model, while the R-squared and Adjusted R-squared values were, respectively, 0.964 and 0.939, and the F-score was 38.6 (p-value $\ll 0.01$). VIF values were all less than five which raised no multicollinearity concern.

Fig. 6 (and Table S6) shows the predicted accuracies using the proposed PR-CI model versus the actual performance in the VM-v3 test session across all subjects. For this variation, the associations between the performances predicted by the PR-CI model and the actual performances were not statistically significant with a rho of 0.2883 (p-value = 0.390).

4.2. Correction

4.2.1. Task correction

Table 6 demonstrates how the proposed predictive model could predict which variation leads to the highest performance for each subject. For each variation, the table lists the actual performance in the test session compared with the predicted performance using the proposed PR-CI model. The variation with the highest accuracy for each subject was selected and then compared to the variation with the highest predicted accuracy based on the PR-CI model. Overall, the PR-CI model achieved 81.8 % correct variation selection, outperforming the PR-C model, which achieved 72.7 % correct variation selection (Tables 6

Table 2

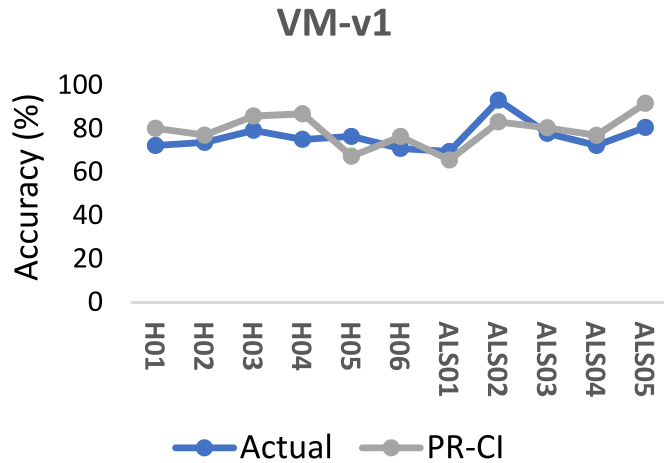
The ultimate predictive models constructed for each variation (VM-v1, VM-v2, and VM-v3). The models are categorized into two groups: first, considering only competence features (PR-C), and then with the inclusion of interfering features (PR-CI). Each formula in the table represents an expansion of equation [3] as described in the method section. The 'P' in the equation represents the performance, which serves as the response variable on the left side of the equation. The first coefficient represents the intercept, while the other coefficients represent the weights assigned to the selected pre-screen features derived through the process of stepwise multivariate linear (MLR) model construction. As described in the feature extraction section, the features used in the models are as follows: Attention, Fluency, and Concentration which are extracted from cognitive-behavioral screen (CBS); HR is heart rate; Delta-TP which refers to the EEG spectral power in the delta band in temporal-parietal region; LF-LT or RF-RT representing fNIRS connectivity in annotated ROIs; P3S_Acc and QDJitter, which are extracted from P3S Speller run; #blinks, Eye_msds, and LD_p2p are extracted interfering features.

VM-v1	PR-C	$P = 0.92 + 0.05 * \text{Attention} - 0.24 * \text{LF_RF} - 0.24 * \text{RF_RT} + 0.24 * \text{Eye_msd}$
	PR-CI	$P = 0.62 + 0.06 * \text{Attention} - 0.01 * \text{Delta_TP} - 0.79 * \text{RF_RT} + 0.51 * \text{P3S_Acc} + 0.26 * \text{Eye_msd} - 0.01 * \text{\#Blink} - 0.01 * \text{LD_p2p} + 0.01 * \text{RF_RT} * \text{LD_p2p}$
VM-v2	PR-C	$P = 2.38 + 0.06 * \text{Fluency} - 0.17 * \text{LF_LT} - 0.01 * \text{QDJitter} - 1.54 * \text{P3S_Acc}$
	PR-CI	$P = 2.38 + 0.06 * \text{Fluency} - 0.17 * \text{LF_LT} - 0.01 * \text{QDJitter} - 1.54 * \text{P3S_Acc}$
VM-v3	PR-C	$P = 0.43 + 0.04 * \text{Concentration} + 0.01 * \text{HR} - 0.71 * \text{LF_RF} - 0.01 * \text{QDJitter} - 0.67 * \text{Eye_msd}$
	PR-CI	$P = 0.88 + 0.03 * \text{Concentration} + 0.01 * \text{HR} - 0.72 * \text{LF_RF} + 0.01 * \text{QDJitter} - 2.01 * \text{Eye_msd} + 0.02 * \text{HR} * \text{Eye_msd}$

Table 3

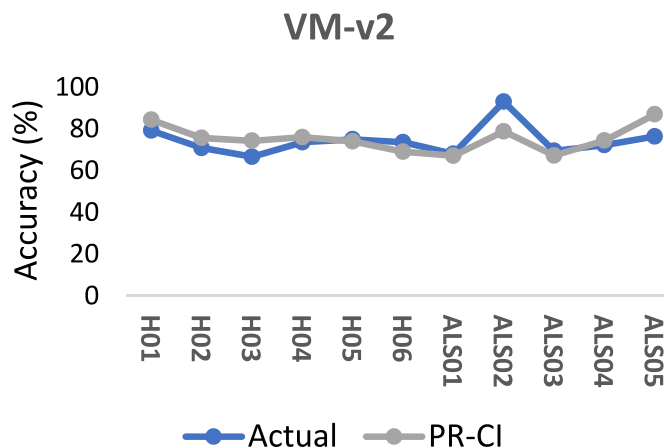
Predictive model constructed from both competency and Interfering features (PR-CI) model for the first variation of VM task (VM-v1).

	Predictive Model's Terms and Statistical Terms							
	Intercept	Attention	Delta-TP	RF-RT	P3S-Acc	Eye-msd	#Blink	LD-p2p
β	0.62	0.06	$\ll -0.01$	-0.79	0.51	0.26	< -0.01	< -0.01
SE	0.20	0.01	$\ll 0.01$	0.15	0.18	0.05	$\ll 0.01$	$\ll 0.01$
tStat	3.01	11.09	-3.50	-5.34	2.76	4.75	-4.84	-3.17
p-value	0.01	$\ll 0.01$	< 0.01	< 0.01	0.02	< 0.01	< 0.01	< 0.01
VIF	NA	1.98	1.35	1.37	1.86	2.45	1.14	2.24

**Fig. 4.** Comparison of actual accuracy (%) of the 1st VM task variation (VM-v1) and the respected predicted accuracy (%) based on PR-CI predictive model in the test session across all subjects (See also Tables S1 and S2, and Fig. S1 for comparison with PR-C model predictions).**Table 4**

Predictive model constructed from both cognitive and interfering features PR-CI model specification for variation 2 of VM task (VM-v2).

	Predictive Model's Terms and Statistical Parameters				
	Intercept	Fluency	LF-LT	QDJitter	P3S-Acc
β	2.38	0.06	-0.17	-0.01	-1.54
SE	0.71	< 0.01	0.06	< 0.01	0.67
tStat	3.35	6.38	-3.11	-3.24	-2.29
p-value	< 0.01	$\ll 0.01$	< 0.01	< 0.01	0.04
VIF		1.22	1.11	5.82	5.62

**Fig. 5.** Comparison of actual accuracy (%) of the 2nd VM task variation (VM-v2) and the respected predicted accuracy (%) based on PR-CI predictive model in the test session across all subjects (See also Tables S3 and S4, and Fig. S2 for comparison with PR-C model).

and S7). Both the F1- and Kappa-scores were also improved in the PR-CI model as compared to the PR-C model (Table 7).

4.2.2. Interference correction

With respect to the terms appearing in the PR-CI model of each variation, we took the corresponding correction strategy. For example, for task variation 1 (VM-v1), as #blink and LD-p2p appeared in its associated PR-CI model, we applied blink removal and short-separation filtering, respectively. For task variation 2 (VM-v2), as QDJitter and LF-LT appeared in its model, we applied jitter correction as well as short-separation filtering, respectively. For task variation 3 (VM-v3), as QDJitter and LD-p2p appeared in the corresponding PR-CI model, we applied jitter-correction and short-separation filtering, respectively.

Fig. 7 (and Table S8) shows the performance (accuracy) changes after applying both task and interference correction strategies. As demonstrated, except for two subjects (H06 and ALS-02), the correction strategies led to a higher performance for all the subjects. Interference correction did not result in any change in performance in the two identified subjects. On average, the correction strategies led to 5.16 % gain in the performance accuracies over all subjects.

5. Discussion

In this work, as an extension of our previously developed fNIRS-EEG spelling BCI [17,30], we proposed a novel multimodal experimental scheme. This predictive scheme involves conducting a quick pre-screening phase prior to the main BCI protocol to produce a feature set. This feature set is then used to predict the appropriate task and artifact correction approaches for the specific subject and session, using a predictive platform trained on data from previous sessions. To actualize and evaluate this scheme, we designed three variations of our previously proposed VM task to account for possible intra- and inter-subject variation arising from differences in the competence and strategies adopted to perform BCI task-related activities. Then, in the task-correction phase, our predictive models predicted which of the variations would lead to the highest performance for each subject on that specific session. This phase was followed by an interference correction phase in which correction strategies were adopted with respect to the terms appearing in the predictive model with the highest performance.

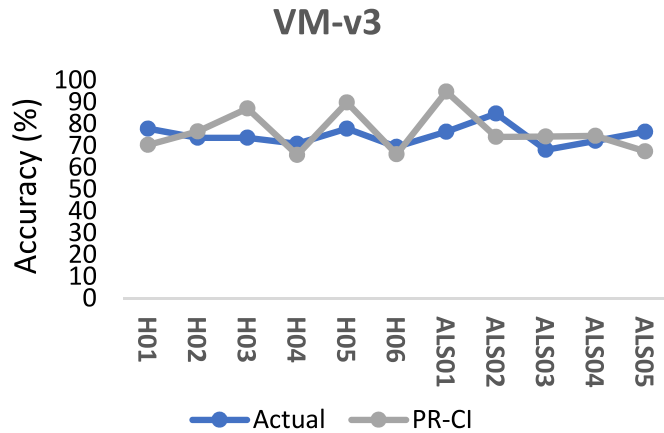
5.1. Impact of the constructed models for people with ALS

To construct our predictive model, we integrated the training data of all subjects without consideration of the motor deficits in our subjects. The rationale behind this approach was twofold; the first reason was to avoid the imposition of any a priori disease-related hypothesis about the impact of motor deficits on the predictive power of our models and allow the model to reveal disease-related confounds which can potentially reflect the possible impact of motor deficit on performance. The second reason for constructing the models without reference to the degree of motor deficit was that the statistical analysis did not reveal any significant between-group difference in either the cognitive competence (CBS score) or the eye-gaze control (eye tracking) features, the major competencies needed for performing VM protocol. It is notable that

Table 5

Predictive model constructed from competency and interfering features PR-CI model for third variation of VM task (VM-v3).

	Predictive Model's Terms and Statistical Parameters							
	Intercept	Concentration	HR	LF-RF	QDJitter	Eye-msd	HR* Eye-msd	LD-p2p
β	0.88	0.03	<0.01	-0.72	<0.01	-2.01	0.02	<0.01
SE	0.19	0.01	<0.01	0.06	<0.01	0.53	<0.01	<0.01
tStat	4.67	3.40	1.87	-11.94	4.37	-3.81	2.57	3.36
p-value	<0.01	<0.01	0.09	<0.01	<0.01	<0.01	0.03	<0.01
VIF	NA	1.10	2.54	1.21	2.37	2.16	NA	1.62

**Fig. 6.** Comparison of actual accuracy (%) of the 3rd variation of the VM task (VM-v3) and the respective predicted accuracy (%) based on both predictive models (PR-C and PR-CI) in the test session across all subjects (See also Tables S5 and S6, and Fig. S3 for comparison with PR-C model).**Table 6**

Comparison of correctness of the task selection by the predictive models.

	Actual	PR-CI	Correctness
H01	v2	v2	1
H02	v1,3	v1	1
H03	v1	v3	0
H04	v1	v1	1
H05	v3	v3	1
H06	v2	v1	0
ALS01	v3	v3	1
ALS02	v1,2	v1	1
ALS03	v1	v1	1
ALS04	v1,2,3	v1	1
ALS05	v1	v1	1
Accuracy (%)			81.8 %

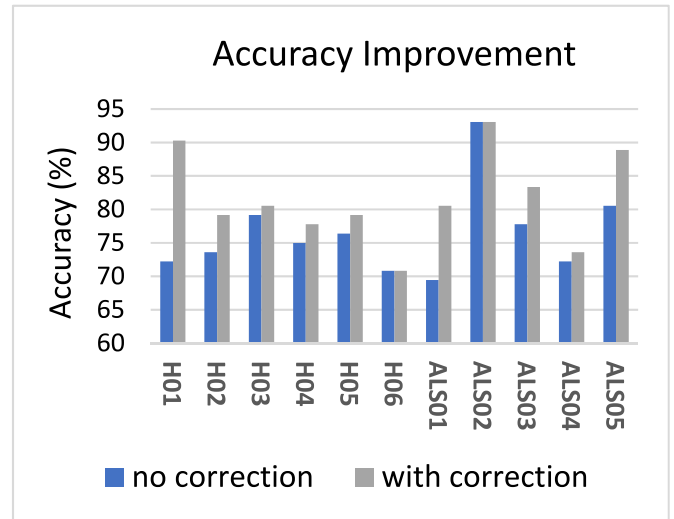
Table 7

The performance metrics for PR-C and PR-CI models, including accuracy, F1-, and Kappa-score.

	Predictive Model's Performance		
	Accuracy (%)	F1 (%)	Kappa
PR-C	71.7	72.5	0.515
PR-CI	81.8	81.3	0.660

performing the VM protocol requires mental arithmetic and visuo-spatial tasks which involve the user's cognitive competence and eye-gaze control. We also observed no statistically significant association between the ALSFRS-R score of the ALS subjects and their CBS score or eye-tracking features, meaning that, at least in our sample population, the presence of motor deficits had no significant association with task-related competencies.

However, the terms that ultimately appeared in the models could

**Fig. 7.** Accuracy (%) changes after applying task- and interference-correction across all subjects (See also Table S8 and Fig. S4).

potentially elucidate how the variation of performance could manifest in a larger patient sample size. The terms retained in the proposed model after variable selection can potentially reflect important information about the association between ALS pathological measures and BCI classification accuracy. For example, verbal fluency, the term that appeared in the predictive model of the second task variation, is a confound reported to be impaired in ALS patients and associated with executive dysfunction [26,28]. Latency jitter, appearing in the models of the 2nd and 3rd task variations, has also been reported to be significantly increased in individuals with ALS and negatively correlated with BCI performance (P3Speller) in both ALS and control participants [60,73]. fNIRS connectivity features, which appeared in all of our models, were identified as confounding factors in other studies, showing between group differences in comparisons between ALS cohorts and healthy controls [59,79]. Kopitzki et al. (2016) have reported significant association between interhemispheric resting-state functional connectivity and ALS pathology [29]. Additionally, fronto-temporal atrophy associated with impaired executive function has been reported in several ALS studies [27,80,81]. While we could not observe any significant between-group performance differences in our relatively small sample size, all previously mentioned features appeared in our models, including the eye-gaze features shown to be impaired in the late stages of ALS [82]. Such findings indicate that it may be possible to predict the most appropriate BCI task for a particular user even in the broader population of individuals with ALS with reference to the range of motor deficits present in these individuals. In future works, including more participants with motor deficits and with more variance in their cognitive and clinical scores could help achieve a more robust between-group categorization in BCI performance variability.

5.2. Scalability of the method

Considering all subjects' inclusion during model training has another methodological advantage in that it could provide us with a unified and comprehensive model in which the training sessions from other subjects could potentially predict the performance of another subject with no training session. This reduces the need for excessive training sessions for each participant and potentially implies that the sources of performance variability could be common among different subjects. This is aligned with recent BCI designs equipped with generic machine learning techniques proposed with the same goal of reducing the need for excessive training sessions [43,44]. For instance, Kang et al. (2014) proposed to replace conventional subject-by-subject training with a multi-subject EEG classification scheme to capture subject-to-subject information transfer by exploiting a linear combination of common spatial patterns (CSP) [45]. Lu et al. (2010) similarly proposed an aggregation of regularized CSP algorithms to give an integrated solution compensating for sample-based covariance matrix estimation shortfalls in a small-sample setting. Transfer learning methods have also been proposed to overcome the need for excessive training runs through utilizing machine learning-based techniques to transfer data from previous sessions or subjects to a new session or subject [46,83]. For example, Kraudelant et al. (2008) have used clustering methods to extract similar CSP filters across sessions and subjects to move towards a zero-training scheme for trained BCI users [83]. Fahimi (2019) suggested an inter-subject transfer learning framework using a deep convolutional neural network (CNN) where the network first learns a general model based on the data from a subject pool, then transfers the trained information to a new subject [46]. Interestingly, our proposed scheme showed the potential to be used in an inter-subject or subject-to-subject transfer learning approach. Regarding this matter, although two of our subjects had no or insufficient training sessions (specifically on the third task variation), the constructed models could still correctly predict how well those subjects perform that task variation in the test session without previous training. This shows the potential of our proposed scheme to be extended to a broader population by constructing a reference model to obviate the need for excessive training sessions for all subjects.

5.3. Efficacy of pre-screening

The cornerstone of our proposed scheme was constructing models to predict BCI performance using a quick pre-screen recording collected at the beginning of each session. Our designed pre-screen recording consists of three parts: the CBS battery, a resting-state recording, and the standard P3Speller. The CBS provided important features in building our proposed predictive model. This is aligned with other studies that have reported similar associations between cognitive battery scores and BCI performance. For example, the ability to better concentrate on performing a task has been demonstrated to have a positive correlation with motor imagery (MI)-based BCI accuracy [31]. Motivational factors have also been reported to be associated with BCI performance in ALS patients [84]. Furthermore, the retention of resting-state features in our predictive models is aligned with studies that have also suggested that integrating resting-state features into BCI design may advance our understanding of task-specific neural dynamics and provide methodological advantages [48,85]. For instance, Blankertz et al. (2018) have reported that higher resting-state Alpha activity can be associated with larger power decreases when performing MI BCI tasks. While other studies have reported individual confounds associated with BCI performance, we propose an integrative predictive framework to aggregate both competence-related and interfering factors to construct a single unified multimodal predictive tool to compensate for performance variability.

5.4. Model variability across different task variations

The different terms appearing in the models of different task variations might be attributed to the differences in the strategies taken by participants and in the competence of subjects to handle different task difficulty levels or workloads. This statement is supported by several previous studies. For example, individual performance in ERP-based BCIs has been reported to be sensitive to variations of cognitive loads [22]. Gu et al. (2020) have reported different effects of adding mental workload to motor imagery-based BCIs when compared with BCIs relying on motor execution [86]. Mental overload has also been reported to negatively impact BCI performance as a result of both subject-related characteristics and task-related components [87], either due to the learning interactivity level in the task (intrinsic) or the method of stimulus presentation (extrinsic) [23]. Differences in the tasks may lead to the allocation of different resources to handle the workload depending on the components of the task [88] and variability across subjects in capacity-limits in working memory [23] or short-memory [89]. The appearance of eye-blink features among interfering factors in the models is aligned with other studies observing that eye-blink interval is associated with mental workload or alertness [90]. Our integrative model is aligned with the Ryu et al. (2005) non-BCI study in which the authors showed a combination of Alpha power, eye-blink and heart rate variability can categorize the workload of four different difficulty levels or workload of a mental arithmetic task [88]. The main advantage of our work is that the terms retained in our models were selected automatically from a broad physiological and interference feature set based on the statistical evaluation of the training data within the context of a machine learning platform, while they selected the terms a priori independent of any training data.

5.5. Task correction

Our subject-dependent task correction is a novel methodological approach to adapt to different effects of mental workload across different subjects and tasks. Providing task variations compensated for the baseline performance variation (VM-v1) in approximately 36% of our participants. Although multiple studies have reported physiological factors associated with BCI workload or difficulty level [24,86], few correction strategies have been proposed to adapt for the corresponding variation in performance. Our proposed method notably accomplishes this using a predictive platform during a calibration and pre-screen phase rather than the main BCI task. In a study with the aim to simulate a practical P3Speller setting where a subject performs the spelling in presence of other mental processes, Ke et al. (2016) and Chen et al. (2017) suggested interleaving two mental tasks (n-back and mental arithmetic) into the P3Speller task and showed that adding mental task features into the training set may enhance speller performance and compensate for the augmented workload [16,91]. Our proposed method has the following advantages compared to their suggested methodology: 1) Their augmented mental dimension is intentionally designed to diminish the oddball response in order to simulate a realistic multi-tasking setting and is present throughout the entire task. In contrast, our VM paradigm is designed such that the mental task is performed only during target intensifications and is additive to the oddball effect. 2) While their suggested paradigm relies on visual competence, our proposed VM protocol eschews this reliance on gaze control by including simultaneous fNIRS recordings. fNIRS has been shown to be synergistic for conventional EEG-only BCIs, particularly for individuals in the late-stage locked-in state, as this modality relies less on eye-gaze competence [17]. 3) Our predictive platform and task correction approach attempts to adopt the best task variation for each subject, while their proposed method lacks subject-specific or session-specific considerations. 4) As previously mentioned, our reliance on features extracted from the pre-screening phase rather than features from the main BCI task can alleviate the tedium of the task or prevent the

within-session habituation effect resulting from repetition of the main BCI task during calibration.

In summary, the proposed predictive scheme was fruitful in the task/workload correction stage by correctly predicting which of the variations of the VM task leads to the highest accuracy for 81.82 % of the subjects. Even for the two subjects where the optimal task was not correctly predicted by the model, the interference correction approaches suggested by the predictive framework could be used to compensate for the error imposed by the identified interfering factors and ultimately led to a higher than baseline performance. In particular, the predicted performances for VM-v1 and VM-v2 could be significantly associated with the actual performances observed in each task, meaning that this framework could potentially be used to estimate inter-subject variance within each task variation. Additionally, both the F1- and Kappa-scores were improved in the PR-CI model as compared to the PR-C model. The kappa score for the PR-CI model indicated substantial agreement between predictions and ground truth for ideal variation selection (0.61–0.80), while the PR-C model showed moderate agreement (0.41–0.60). Kappa scores in both models could have been improved further, if, in our dataset, we had more samples from each variation, as the subjects' optimal task.

5.6. Interference correction

Among interference features, the 'number of blinks' feature from the EOG feature set appeared in some of our models, while no EMG features were retained in any of the models. One explanation could be that none of the participants exhibited facial muscle tension that could cause differences in features extracted from targets as opposed to non-targets. Applying short-separation filtering as a step in our interference correction approach demonstrated promising improvements aligned with other fNIRS studies that similarly show improvements when using short-distance channel information in the preprocessing stage [92,93]. The subject-dependent adoption of correction methods can provide a more efficient processing pipeline as opposed to excessive pre/processing methods blind to the dynamics of the task-related demands, interfering factors, and intra- and inter-subject variations.

5.7. Conclusion

Overall, with methodological emphasis on the pre-screen stage, our proposed experimental scheme showed efficacy in predicting, identifying, and compensating for performance variability factors in a BCI task. In particular, it showed fruitfulness for people with ALS, with the potential of being extended to a broader population with other motor deficits. As, to the best of our knowledge, no reliable communicative BCI design currently exists for people in a completely locked-in state, our proposed variations of a novel visuo-mental task within our predictive-corrective framework constitute one more step towards a practical and ideal BCI platform. While few similar studies have sparsely reported individual confounds associated with BCI performance variations, we proposed an integrative predictive framework to aggregate both competence-related and interfering factors to construct a single unified multimodal predictive tool which then compensates for performance variability. In addition to the integrative approach, the subject-dependent adoption of correction strategies was a novel aspect of this work. Finally, the simple predictive models constructed from a comprehensive pre-screening feature could provide a dynamic ground to customize experimental designs based on variability across subjects and sessions.

5.8. Future work and limitations

However, to extend this work to capture intra-subject longitudinal variation, we must record from the same participant for longer periods to validate our constructed models, although our current results indicate

that information about the sources of intra-subject variability could potentially be transferred from information about sources of inter-subject variability. Furthermore, the scalability of our study was limited by its small sample size due to the rare nature of the disease and the difficulty of conducting a longitudinal activity-based paradigm with partly/completely locked-in ALS patients. For future work, replicating our results with larger sample sizes would facilitate the generalizability of the reported outcome. Additionally, it may be informative to repeat this experiment using a broader range of tasks designed to involve greater degrees of cognitive complexity. For instance, as lengthening the stimulation time may make numbers available for longer time, it might provide a controlling factor for adjusting the difficulty of the task and correspondingly compensating performance variability. This approach could provide the benefits of capturing greater variability in the strategies used by participants to complete the tasks and explore different facets of participant BCI task competence, subsequently offering more degrees of freedom to the constructed models. In future recordings, the findings in this work can be used to customize the experimental design based on each subject's data. For instance, in the BCI-task runs, we can reduce the recording of fNIRS and EEG channels in the test sessions based on the optimum channels extracted in their training runs. Additionally, we can reduce the number of channels or ROI based on the terms that appeared in the predictive model. This reduction of channels can ease the recording set up which can particularly benefit people with severe motor deficits using these systems on a daily basis. Moreover, using alternative classification methods, such as Support Vector Machine (SVM), Neural Networks (e.g., CNN) and deep learning techniques (e.g., DNN) may enhance individual performances for each subject and session. It is essential to consider that the relative difference between performances and their variability might remain consistent. As a result, comparatively, the reported results should not be significantly affected. Furthermore, while our correction stage can be considered a hyperparameter optimization stage [94] since it adjusts the calculation of certain parameters/features, there is room for further improvement by incorporating additional hyperparameters. Future works could explore the use of additional hyperparameters, particularly in the stepwise feature entrance and removal process. To investigate the stability of our proposed method and to address concerns about its dependency on the recorded data, it would be helpful to replicate the methodology with a different dataset. In this regard, we have examined available similar BCI datasets, such as the largest SCP dataset [95] and BCI Competition IV-2a&b [96,97]. However, to the best of our knowledge, none of these datasets could serve the evaluative purpose. Among the available BCI datasets, none of them a) had a pre-screen run before the main BCI task, b) were tested on people with severe motor deficits, or c) recorded an interfering modality for applying interference correction. Additionally, since these datasets neither were multimodal nor included the mental non-motor aspect in the tasks, even applying our method could not conclusively address the stability problem. The comparability of these datasets to our own is thus too compromised to be an effective test of our method at this time. Future recording with the same experimental scheme as proposed in this work will help to draw more robust conclusions about the stability of the reported results. Extending the proposed scheme to a real-time platform would be the next step to validate its applicability on a daily basis.

Declaration of competing interest

The authors do not have any declaration of interest.

Acknowledgements

This work was supported by the National Science Foundation (NSF-1913492, NSF-2006012) and the Institutional Development Award (IDeA) Network for Biomedical Research Excellence (P20GM103430). We would like to thank the participants who took part in this study,

without whom this study would not have been possible. We would also like to thank the ALS Association Rhode Island Chapter and the National Center for Adaptive Neurotechnologies (NCAN) for their continuous support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.compbimed.2023.107658>.

References

- [1] R. Chavarriaga, M. Fried-Oken, S. Kleih, F. Lotte, R. Scherer, Heading for new shores! Overcoming pitfalls in BCI design [Internet], *Brain-Computer. Interf.* (2017 Apr 3). ;4(1–2):60–73. Available from: <https://www.tandfonline.com/doi/abs/10.1080/2326263X.2016.1263916>.
- [2] Y. Shahrari, T.M. Vaughan, L.M. McCane, B.Z. Allison, J.R. Wolpaw, D. J. Krusienski, An exploration of BCI performance variations in people with amyotrophic lateral sclerosis using longitudinal EEG data [Internet], *J. Neural. Eng.* (2019 Sep 17). ;16(5). Available from: <https://pubmed.ncbi.nlm.nih.gov/31108477/>.
- [3] L. Holper, N. Kobashi, D. Kiper, F. Scholkmann, M. Wolf, K. Eng, Trial-to-trial variability differentiates motor imagery during observation between low versus high responders: a functional near-infrared spectroscopy study, *Behav. Brain Res.* 229 (1) (2012) 29–40.
- [4] P. Kellmeyer, M. Grosse-Wentrup, A. Schulze-Bonhage, U. Ziemann, T. Ball, Electrophysiological correlates of neurodegeneration in motor and non-motor brain regions in amyotrophic lateral sclerosis - implications for brain-computer interfacing, *J. Neural. Eng.* 15 (4) (2018).
- [5] M. Naito, Y. Michioka, K. Ozawa, Y.I. Ito, M. Kiguchi, T. Kanazawa, A communication means for totally locked-in ALS patients based on changes in cerebral blood volume measured with near-infrared light, *IEICE Trans. Info Syst.* E90-D (7) (2007) 1028–1037.
- [6] S.G.S. Shah, I. Robinson, S. Alshawi, Developing medical device technologies from users' perspectives: a theoretical framework for involving users in the development process [Internet], in: *International Journal of Technology Assessment in Health Care*, vol. 25, Int J Technol Assess Health Care, 2009 [cited 2020 Dec 1]. pp. 514–21. Available from: <https://pubmed.ncbi.nlm.nih.gov/19845981/>.
- [7] F. Lotte, F. Larue, C. Mühl, Flaws in current human training protocols for spontaneous Brain-Computer Interfaces: lessons learned from instructional design [Internet], *Front. Hum. Neurosci.* (2013 Sep 17). ;7(SEP):568. Available from: <http://journal.frontiersin.org/article/10.3389/fnhum.2013.00568/abstract>.
- [8] M. Rodríguez-Ugarte, E. Iáñez, M. Ortiz, J.M. Azorín, Personalized offline and pseudo-online BCI models to detect pedaling intent [Internet], *Front. Neuroinf.* (2017 Jul 11). ;11:45. Available from: <http://journal.frontiersin.org/article/10.3389/fninf.2017.00045/full>.
- [9] S. Smith, E. Duff, A. Groves, T.E. Nichols, S. Jbabdi, L.T. Westlye, et al., Structural variability in the human brain reflects fine-grained functional architecture at the population level [Internet], *J. Neurosci.* (2019 Jul 31). ;39(31):6136–49. Available from: <https://www.jneurosci.org/content/39/31/6136>.
- [10] V. Wens, M. Bourguignon, S. Goldman, B. Marty, M. Op De Beeck, C. Clumeck, et al., Inter- and intra-subject variability of neuromagnetic resting state networks [Internet], *Brain Topogr.* (2014 Apr 29). ;27(5):620–34. Available from: <https://link.springer.com/article/10.1007/s10548-014-0364-8>.
- [11] C. Vidaurre, B. Blankertz, Towards a cure for BCI illiteracy [Internet], *Brain Topogr.* (2010 Jun 28). ;23(2):194–8. Available from: <https://link.springer.com/article/10.1007/s10548-009-0121-6>.
- [12] M.L. Seghier, C.J. Price, Interpreting and utilising intersubject variability in brain function, *Trends Cognit. Sci.* 22 (6) (2018 Jun 1) 517–530.
- [13] C. Sannelli, C. Vidaurre, K.R. Müller, B. Blankertz, A large scale screening study with a SMR-based BCI: categorization of BCI users and differences in their SMR activity [Internet], *PLoS One* (2019 Jan 1). ;14(1):e0207351. Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0207351>.
- [14] M. Ahn, S.C. Jun, Performance variation in motor imagery brain-computer interface: a brief review, *J. Neurosci. Methods* 243 (2015 Mar 30) 103–110.
- [15] I. Käthner, S.C. Wriessnegger, G.R. Müller-Putz, A. Kübler, S. Halder, Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain-computer interface [Internet], *Biol. Psychol.* (2014). ;102(1):118–29. Available from: <https://pubmed.ncbi.nlm.nih.gov/25088378/>.
- [16] Y. Ke, P. Wang, Y. Chen, B. Gu, H. Qi, P. Zhou, et al., Training and testing ERP-BCIs under different mental workload conditions [Internet], *J. Neural. Eng.* (2016 Dec 11). ;13(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/26655346/>.
- [17] S.B. Borgheai, J. McLinden, A.H. Zisk, S.I. Hosni, R.J. Deligani, M. Abtahi, et al., Enhancing communication for people in late-stage ALS using an fNIRS-based BCI system, *IEEE Trans. Neural Syst. Rehabil. Eng.* 28 (5) (2020).
- [18] L.F. Nicolas-Alonso, J. Gomez-Gil, Brain computer interfaces, a review [Internet], *Sensors* (2012 Feb). ;12(2):121. Available from: <https://pmc/articles/PMC3304110/>.
- [19] T. Kaufmann, C. Vögele, S. Sütterlin, S. Lukito, A. Kübler, Effects of resting heart rate variability on performance in the P300 brain-computer interface [Internet], *Int. J. Psychophysiol.* (2012). Mar ;83(3):336–41. Available from: <https://pubmed.ncbi.nlm.nih.gov/22172335/>.
- [20] Curran EA, Stokes MJ. Learning to Control Brain Activity: A Review of the Production and Control of EEG Components for Driving Brain-Computer Interface (BCI) Systems. Available from: <http://www.elsevier.com/locate/b&c>.
- [21] Y. Ke, P. Wang, Y. Chen, B. Gu, H. Qi, P. Zhou, et al., Training and testing ERP-BCIs under different mental workload conditions, *J. Neural. Eng.* 13 (1) (2015).
- [22] S.J. Luck, 1 A Broad Overview of the Event-Related Potential Technique. An Introduction to the Event-Related Potential Technique, 2014, pp. 1–34, 2014.
- [23] J. Sweller, Can We Measure Working Memory without Contamination from Knowledge Held in Long-Term Memory? *Behavioral and Brain Sciences* [Internet], 1998 Dec. ;21(6):845–6. Available from: <https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/abs/can-we-measure-working-memory-without-contamination-from-knowledge-held-in-longterm-memory/EA976424D1EAC664B705D8306A9F374A>.
- [24] Felton EA, Williams JC, Vanderheiden GC, Radwin RG. Mental workload during brain-computer interface training. <https://doi.org/10.1080/001401392012662526> [Internet]. 2012 May [cited 2022 Apr 21];55(5):526–37. Available from: <https://www.tandfonline.com/doi/abs/10.1080/001401392012662526>.
- [25] M. Causse, Z. Chua, V. Peysakhovich, N. Del Campo, N. Matton, Mental workload and neural efficiency quantified in the prefrontal cortex using fNIRS [Internet], *Sci. Rep.* (2017 Dec 1), www.nature.com/scientificreports. ;7(1):1–15..
- [26] S. Abrahams, P.N. Leigh, A. Harvey, G.N. Vythelingum, D. Gries, L.H. Goldstein, Verbal fluency and executive dysfunction in amyotrophic lateral sclerosis (ALS), *Neuropsychologia* 38 (6) (2000 Jun 1) 734–747.
- [27] F. Trojci, P. Sorrentino, G. Sorrentino, G. Tedeschi, Neurodegeneration of brain networks in the amyotrophic lateral sclerosis-frontotemporal lobar degeneration (ALS-FTLD) continuum: evidence from MRI and MEG studies, *CNS Spectr.* 23 (6) (2018) 378–387.
- [28] M. Perez, I. Amaya, E. Lazaro, M. García, O. Martínez, P. Caballero, et al., Intrusion errors during verbal fluency task in amyotrophic lateral sclerosis, *PLoS One* 15 (5) (2020) 1–11.
- [29] K. Kopitzki, A. Oldag, C.M. Sweeney-Reed, J. Machts, M. Veit, J. Kaufmann, et al., Interhemispheric connectivity in amyotrophic lateral sclerosis: a near-infrared spectroscopy and diffusion tensor imaging study [Internet], *Neuroimage Clin* (2016). ;12:666–72. Available from: <https://pubmed.ncbi.nlm.nih.gov/27761397/>.
- [30] S.B. Borgheai, R.J. Deligani, J. McLinden, A. Zisk, S.I. Hosni, M. Abtahi, et al., Multimodal exploration of non-motor neural functions in ALS patients using simultaneous EEG-fNIRS recording, *J. Neural. Eng.* 16 (6) (2019).
- [31] E.M. Hammer, S. Halder, B. Blankertz, C. Sannelli, T. Dickhaus, S. Kleih, et al., Psychological predictors of SMR-BCI performance, *Biol. Psychol.* 89 (1) (2012 Jan) 80–86.
- [32] A. Vasilyev, S. Liburkina, L. Yakovlev, O. Perepelkina, A. Kaplan, Assessing motor imagery in brain-computer interface training: psychological and neurophysiological correlates, *Neuropsychologia* 97 (2017 Mar 1) 56–65.
- [33] L. Frolich, I. Winkler, K.R. Müller, W. Samek, Investigating effects of different artefact types on motor imagery BCI [Internet], *Annu Int Conf IEEE Eng Med Biol Soc* (2015 Nov 4). ;2015:1942–5. Available from: <https://pubmed.ncbi.nlm.nih.gov/26736664/>.
- [34] X. Yong, M. Fatourehchi, R.K. Ward, G.E. Birch, Automatic artefact removal in a self-paced hybrid brain-computer interface system [Internet], *J. NeuroEng. Rehabil.* (2012 Jul 27). ;9(1):1–20. Available from: <https://jneuroengrehab.biomedcentral.com/articles/10.1186/1743-0003-9-50>.
- [35] D.E. Thompson, M.R. Mowla, K.J. Dhuyvetter, J.W. Tillman, J.E. Huggins, Automated artifact rejection algorithms harm P3 Speller brain-computer interface performance, *Brain-Computer. Interf.* 6 (4) (2019) 141–148.
- [36] J.W. Barker, A. Aarabi, T.J. Huppert, Autoregressive model based algorithm for correcting motion and serially correlated errors in fNIRS, *Biomed. Opt Express* 4 (8) (2013) 1366.
- [37] H. Santosa, X. Zhai, F. Fishburn, P.J. Sparto, T.J. Huppert, Quantitative comparison of correction techniques for removing systemic physiological signal in functional near-infrared spectroscopy studies, *Neurophotonics* 7 (3) (2020) 1–21.
- [38] I. Tachtsidis, F. Scholkmann, False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward, *Neurophotonics* 3 (3) (2016), 031405.
- [39] Gregg, Brain specificity of diffuse optical imaging: improvements from superficial signal regression and tomography, *Front. Neuroenergetics* 2 (July) (2010) 1–8.
- [40] J.R. Wolpaw, D.J. McFarland, G.W. Neat, C.A. Forneris, An EEG-based brain-computer interface for cursor control, *Electroencephalogr. Clin. Neurophysiol.* 78 (3) (1991 Mar 1) 252–259.
- [41] S. Saha, M. Baumert, Intra- and inter-subject variability in EEG-based sensorimotor brain computer interface: a review, *Front. Comput. Neurosci.* 13 (2020 Jan 1) 87.
- [42] R. Caraballona, The role of the interplay between stimulus type and timing in explaining BCI-illiteracy for visual P300-based Brain-Computer Interfaces, *Front. Neurosci.* 11 (JUN) (2017 Jun 30) 363.
- [43] H. Ramoser, J. Müller-Gerking, G. Pfurtscheller, Optimal spatial filtering of single trial EEG during imagined hand movement, *IEEE Trans. Rehabil. Eng.* 8 (4) (2000) 441–446.
- [44] Blankertz B, Curio G, Müller KR. Classifying Single Trial EEG: towards Brain Computer Interfacing.
- [45] H. Kang, S. Choi, Bayesian common spatial patterns for multi-subject EEG classification, *Neural Network.* 57 (2014 Sep 1) 39–50.
- [46] F. Fahimi, Z. Zhang, W.B. Goh, T.S. Lee, K.K. Ang, C. Guan, Inter-subject transfer learning with an end-to-end deep convolutional neural network for EEG-based BCI [Internet], *J. Neural. Eng.* (2019 Jan 23). ;16(2):026007. Available from: <https://iopscience.iop.org/article/10.1088/1741-2552/aaf3f6>.

- [47] H Il Suk, S. Fazli, J. Mehnert, K.R. Müller, S.W. Lee, Predicting BCI subject performance using probabilistic spatio-temporal filters [Internet], *PLoS One* (2014 Feb 14). ;9(2):e87056. Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0087056>.
- [48] H. Morioka, A. Kanemura, Hirayama J. ichiro, M. Shikachi, T. Ogawa, S. Ikeda, et al., Learning a common dictionary for subject-transfer decoding with resting calibration, *Neuroimage* 111 (2015 May 1) 167–178.
- [49] C. Sannelli, C. Vidaurre, K.R. Müller, B. Blankertz, A large scale screening study with a SMR-based BCI: categorization of BCI users and differences in their SMR activity [Internet], *PLoS One* (2019 Jan 1). ;14(1):e0207351. Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0207351>.
- [50] S. Saha, M.S. Hossain, K. Ahmed, R. Mostafa, L. Hadjileontiadis, A. Khandoker, et al., Wavelet entropy-based inter-subject associative cortical source localization for sensorimotor BCI, *Front. Neuroinf.* 13 (2019 Jul 23) 47.
- [51] M. Wronkiewicz, E. Larson, A.K.C. Lee, Leveraging anatomical information to improve transfer learning in brain–computer interfaces [Internet], *J. Neural. Eng.* (2015 Jul 14). ;12(4):046027. Available from: <https://iopscience.iop.org/article/10.1088/1741-2560/12/4/046027>.
- [52] S. Perdikis, L. Tonin, S. Saeedi, C. Schneider, J. del R. Millán, The Cybathlon BCI race: successful longitudinal mutual learning with two tetraplegic users [Internet], *PLoS Biol.* (2018 May 10). ;16(5):e2003787. Available from: <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.2003787>.
- [53] L. Mingai, G. Shuoda, Z. Guoyu, S. Yanjun, Y. Jinfu, Removing ocular artifacts from mixed EEG signals with FastKICA and DWT, *J. Intell. Fuzzy Syst.* 28 (6) (2015) 2851–2861.
- [54] R. Ramli, H. Arof, F. Ibrahim, N. Mokhtar, M.Y.I. Idris, Using finite state machine and a hybrid of EEG signal and EOG artifacts for an asynchronous wheelchair navigation [Internet], *Expert Syst. Appl.* 42 (5) (2015) 2451–2463, <https://doi.org/10.1016/j.eswa.2014.10.052>.
- [55] K.S. Hong, M.J. Khan, Hybrid brain-computer interface techniques for improved classification accuracy and increased number of commands: a review, *Front. Neuroinf.* 11 (2017) JUL.
- [56] B. Chen, Y. Feng, Q. Wang, Combining vibrotactile feedback with volitional myoelectric control for robotic transistibial prostheses, *Front. Neuroinf.* 10 (2016) 1–14. AUG.
- [57] M.J. Khan, K.S. Hong, Hybrid EEG-fNIRS-based eight-command decoding for BCI: application to quadcopter control, *Front. Neuroinf.* 11 (2017) FEB.
- [58] J.M. Cedarbaum, N. Stambler, E. Malta, C. Fuller, D. Hilt, B. Thurmond, et al., The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function, *J. Neurol. Sci.* 169 (1–2) (1999) 13–21.
- [59] R.J. Deligani, S.I. Hosni, S.B. Borgheai, J. McLinden, A.H. Zisk, K. Mankodiya, et al., Electrical and hemodynamic neural functions in people with ALS: an EEG-fNIRS resting-state study, *IEEE Trans. Neural Syst. Rehabil. Eng.* 28 (12) (2020 Dec 1) 3129–3139.
- [60] A.H. Zisk, S.B. Borgheai, J. McLinden, S.M. Hosni, R.J. Deligani, Y. Shahriari, P300 latency jitter and its correlates in people with amyotrophic lateral sclerosis, *Clin. Neurophysiol.* 132 (2) (2021 Feb 1) 632–642.
- [61] S. Geng, X. Liu, B.B. Biswal, H. Niu, Effect of resting-state fNIRS scanning duration on functional brain connectivity and graph theory metrics of brain network [Internet], *Front. Neurosci.* (2017 Jul 20). ;11(JUL), 392. Available from: <https://pubmed.ncbi.nlm.nih.gov/28775676/>.
- [62] A. Wunderlich, K. Gramann, Eye movement-related brain potentials during assisted navigation in real-world environments, *Eur. J. Neurosci.* 54 (12) (2021 Dec 1) 8336–8354.
- [63] H. Santosa, X. Zhai, F. Fishburn, T. Huppert, The NIRS brain AnalyzIR toolbox, *Algorithms* 11 (5) (2018).
- [64] Asgher U, Ahmad R, Naseer N, Ayaz Y, Khan MJ, Kamal Amjad M. Assessment and Classification of Mental Workload in the Prefrontal Cortex (PFC) Using Fixed-Value Modified Beer-Lambert Law.
- [65] M. Hiraoka, M. Firbank, M. Essenpreis, M. Cope, S.R. Arridge, P. Van Der Zee, et al., A Monte Carlo investigation of optical pathlength in inhomogeneous tissue and its application to near-infrared spectroscopy [Internet], *Phys. Med. Biol.* (1993). ;38(12):1859–76. Available from: <https://pubmed.ncbi.nlm.nih.gov/8108489/>.
- [66] H. Santosa, X. Zhai, F. Fishburn, T. Huppert, The NIRS brain AnalyzIR toolbox, *Algorithms* 11 (5) (2018).
- [67] M. Fraschini, M. Lai, M. Demuru, M. Puligheddu, G. Floris, G. Borghero, et al., Functional Brain Connectivity Analysis in Amyotrophic Lateral Sclerosis: an EEG Source-Space Study [Internet], *Biomed Phys Eng Express*, 2018 Apr 25, <https://doi.org/10.1088/2057-1976/aa9c64> [cited 2020 Aug 12];4(3):37004.
- [68] S. Coyle, T. Ward, C. Markham, Cerebral blood flow changes related to motor imagery, using near infrared spectroscopy (NIRS), in: *World Congress on Medical Physics and Biomedical Engineering*, 2003.
- [69] Y. Li, D. Yu, Weak network efficiency in young children with Autism Spectrum Disorder: evidence from a functional near-infrared spectroscopy study, *Brain Cognit.* 108 (2016 Oct 1) 47–55.
- [70] H. Chen, G. Miao, S. Wang, J. Zheng, X. Zhang, J. Lin, et al., Disturbed functional connectivity and topological properties of the frontal lobe in minimally conscious state based on resting-state fNIRS, *Front. Neurosci.* 17 (2023 Feb 10), 1118395.
- [71] D.J. Krusienski, E.W. Sellers, D.J. McFarland, T.M. Vaughan, J.R. Wolpaw, Toward enhanced P300 speller performance, *J. Neurosci. Methods* (2008);167(1), 15–21.
- [72] Mowla MR, Huggins JE, Thompson DE. Enhancing P300-BCI performance using latency estimation. <https://doi.org/10.1080/2326263X20171338010> [Internet]. 2017 Jul 3 [cited 2022 Jun 27];4(3):137–45. Available from: <https://www.tandfonline.com/doi/abs/10.1080/2326263X.2017.1338010>.
- [73] D.E. Thompson, S. Warschausky, J.E. Huggins, Classifier-based latency estimation: a novel way to estimate and predict BCI accuracy [Internet], *J. Neural. Eng.* (2013 Feb). ;10(1), 016006. Available from: <https://pubmed.ncbi.nlm.nih.gov/23234797/>.
- [74] C. Spiewak, A comprehensive study on EMG feature extraction and classifiers, *Open Access J Biomed Eng Biosci* 1 (1) (2018 Feb 7).
- [75] J.A. Noah, X. Zhang, S. Dravida, C. DiCocco, T. Suzuki, R.N. Aslin, et al., Comparison of short-channel separation and spatial domain filtering for removal of non-neural components in functional near-infrared spectroscopy signals, *Neurophotonics* 8 (1) (2021) 1–22.
- [76] D.A. Belsley, Kuh Edwin, R.E. Welsch, *Regression Diagnostics : Identifying Influential Data and Sources of Collinearity*, 2004, p. 292.
- [77] S. Baldini, M.E. Morelli, A. Sartori, F. Pasquin, A. Dinoto, A. Bratina, et al., Microstates in Multiple Sclerosis: an Electrophysiological Signature of Altered Large-Scale Networks Functioning? *Brain Commun* [Internet], 2022, <https://doi.org/10.1093/braincomms/fcac255> [cited 2023 Jul 3];5.
- [78] X. Ma, L. Song, B. Hong, Y. Li, Y. Li, Relationships between EEG and thermal comfort of elderly adults in outdoor open spaces, *Build. Environ.* 235 (2023 May 1), 110212.
- [79] S.B. Borgheai, J. McLinden, K. Mankodiya, Y. Shahriari, Frontal functional network disruption associated with amyotrophic lateral sclerosis: an fNIRS-based minimum spanning tree analysis, *Front. Neurosci.* 14 (2020). December.
- [80] D. Dimond, A. Ishaque, S. Chenji, D. Mah, Z. Chen, P. Seres, et al., White matter structural network abnormalities underlie executive dysfunction in amyotrophic lateral sclerosis, *Hum. Brain Mapp.* 38 (3) (2017) 1249–1268.
- [81] S.C. Woolley, M.J. Strong, Frontotemporal dysfunction and dementia in amyotrophic lateral sclerosis [Internet], *Neurologic Clinics*. W.B. Saunders 33 (2015) [cited 2020 Aug 12]. pp. 787–805. Available from: <https://pubmed.ncbi.nlm.nih.gov/26515622/>.
- [82] U. Chaudhary, N. Birbaumer, A. Ramos-Murguialday, Brain–computer interfaces in the completely locked-in state and chronic stroke [Internet], in: *Progress in Brain Research*, Elsevier B.V., 2016 [cited 2020 Dec 1]. pp. 131–61. Available from: <https://pubmed.ncbi.nlm.nih.gov/27590968/>.
- [83] M. Krauledat, M. Tangermann, B. Blankertz, K.R. Müller, Towards zero training for brain-computer interfacing [Internet], *PLoS One* 3 (8) (2008) 2967, www.plosone.org. Available from.
- [84] F. Nijboer, N. Birbaumer, A. Kübler, The influence of psychological state and motivation on brain-computer interface performance in patients with amyotrophic lateral sclerosis - a longitudinal study, *Front. Neurosci.* 4 (2010) 55. JUL.
- [85] G. Northoff, N.W. Duncan, D.J. Hayes, The brain and its resting state activity—experimental and methodological implications, *Prog. Neurobiol.* 92 (4) (2010 Dec 1) 593–600.
- [86] B. Gu, L. Chen, Y. Ke, Y. Zhou, H. Yu, K. Wang, et al., The effects of varying levels of mental workload on motor imagery based brain-computer interface, *Int. J. Embed. Syst.* 12 (3) (2020) 315–323.
- [87] E. Galy, C. Mélan, Effects of cognitive appraisal and mental workload factors on performance in an arithmetic task [Internet], *Appl. Psychophysiol. Biofeedback* (2015 Dec 1). ;40(4):313–25. Available from: <https://pubmed.ncbi.nlm.nih.gov/26205469/>.
- [88] K. Ryu, R. Myung, Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic, *Int. J. Ind. Ergon.* 35 (11) (2005 Nov 1) 991–1009.
- [89] M.W. Miller, J.C. Rietschel, C.G. McDonald, B.D. Hatfield, A novel approach to the physiological measurement of mental workload [Internet], *Int. J. Psychophysiol.* (2011 Apr). ;80(1):75–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/21320552/>.
- [90] Stern JA, Skelly JJ. The Eye Blink and Workload Considerations: <https://doi.org/10.1177/154193128402801101> [Internet]. 2016 Aug 6 [cited 2022 Jun 28];2: 942–4. Available from: https://journals.sagepub.com/doi/abs/10.1177/154193128402801101?casa_token=WsABgrf8ccgAAAA%3ANnBRppN-yXCaXgCRsb5xfvbmZ_CzeymVNrKQsx9Ky3IOVnMlpp7bKuC4YkJzttBilD5vXilQ.
- [91] Y. Chen, Y. Ke, G. Meng, J. Jiang, H. Qi, X. Jiao, et al., Enhancing performance of P300-speller under mental workload by incorporating dual-task data during classifier training, *Comput. Methods Progr. Biomed.* 152 (2017 Dec 1) 35–43.
- [92] H. Santosa, X. Zhai, F. Fishburn, P.J. Sparto, Quantitative comparison of correction techniques for removing systemic physiological signal in functional near-infrared spectroscopy studies 7 (3) (2020) 1–21.
- [93] D. Wyser, M. Mattille, M. Wolf, O. Lambercy, F. Scholkmann, R. Gassert, Short-channel regression in functional near-infrared spectroscopy is more effective when considering heterogeneous scalp hemodynamics, *Neurophotonics* 7 (3) (2020).
- [94] S.M.I. Hosni, S.B. Borgheai, J. McLinden, S. Zhu, X. Huang, S. Ostadabbas, et al., A graph-based nonlinear dynamic characterization of motor imagery toward an enhanced hybrid BCI [Internet], *Neuroinformatics* (2022 Oct 1). ;20(4):1169–89. Available from: <https://link.springer.com/article/10.1007/s12021-022-09595-2>.
- [95] M. Kaya, M.K. Binli, E. Ozbay, H. Yanar, Y. Mishchenko, Data descriptor: a large electroencephalographic motor imagery dataset for electroencephalographic brain computer interfaces, *Sci. Data* 5 (2018) 1–16. August.
- [96] R.R. Chowdhury, Y. Muhammad, U. Adeel, Enhancing cross-subject motor imagery classification in EEG-based brain–computer interfaces by using multi-branch CNN, *Sensors* 23 (18) (2023) 7908.
- [97] X. Wang, X. Dai, Y. Liu, X. Chen, Q. Hu, R. Hu, et al., Motor imagery electroencephalogram classification algorithm based on joint features in the spatial and frequency domains and instance transfer, *Front. Hum. Neurosci.* 17 (2023) 1–11. May.