2021

QUANTIFYING AND COMPENSATING FOR P300 VARIATIONS IN AMYOTROPHIC LATERAL SCLEROSIS

Alyssa Hillary Zisk
University of Rhode Island, alyssazisk@uri.edu

QUANTIFYING AND COMPENSATING FOR P300 VARIATIONS IN AMYOTROPHIC LATERAL SCLEROSIS

BY

ALYSSA HILLARY ZISK

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

INTERDISCIPLINARY NEUROSCIENCE

UNIVERSITY OF RHODE ISLAND

2021

DOCTOR OF PHILOSOPHY IN INTERIDISCIPLINARY NEUROSCIECE

DISSERTATION

OF

ALYSSA HILLARY ZISK

APPROVED:

Dissertation Committee:

Major Professor Yalda Shahriari

Co-Major Professor Kunal Mankodiya

Frank Menniti

Manbir Sodhi

Brenton DeBoef
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2021

ABSTRACT

P300-based brain–computer interface (BCI) systems enable people with neuromuscular disabilities, including amyotrophic lateral sclerosis (ALS), to communicate and to control their environments using brain activity. However, BCI systems have not yet fulfilled their promise as reliable communication systems for all who need them. Despite continued work on improving BCIs for end users, people with ALS can experience both reduced performance overall compared to neurotypical users and significant day to day variations in BCI performance and event-related potential (ERP) characteristics.

The P300 response, which the P300 speller is based on, is also known to exhibit trial-to-trial latency variability. The importance of latency jitter is established in cognitive studies, and its relevance to BCIs is of growing interest. Increased latency jitter is associated with decreased BCI performance, and preliminary comparisons indicated that jitter may be increased in ALS, similar to increased jitter found in a variety of neurological conditions.

Therefore, we quantify latency jitter and its correlates in people with ALS, longitudinally investigate within-session variability in event-related potentials (ERPs), session-average ERPs, and their relationships, and develop and evaluate a correction method to compensate for latency jitter in BCI use. To this end, we use longitudinal EEG data collected from 6 participants with ALS, and, when applicable, from neurotypical control participants, using a P300 BCI. Data recorded in each session had session-average ERP amplitudes and latencies extracted. Stepwise linear discriminant analysis was used both to evaluate BCI performance and to support the use of classifier-

based latency estimation (CBLE) to estimate whole-epoch latency shifts for single trials in all aims.

To quantify latency jitter and its correlates in people with ALS, latency jitter was calculated with CBLE. Then, ERP components and latency jitter were compared between participants with ALS and neurotypical control participants using Wilcoxon rank-sum tests. Correlations between latency jitter and each of the clinical measures, ERP features, and performance measures were investigated using Spearman and repeated measures correlations. We found that latency jitter was significantly increased in participants with ALS and significantly negatively correlated with BCI performance in both ALS and control participants. We also found significant correlations between ERP amplitudes and latency jitter in neurotypical participants and reduced ERP amplitudes in participants with ALS. However, there was no significant correlation between latency jitter and clinical measures.

Based on these results, we proposed a data augmentation and jitter correction (A/C) scheme with parameters determined individually using latency shifts calculated with CBLE. Performance metrics including character selection accuracy and binary accuracy, precision, recall, and F-score were calculated using both the proposed classification scheme and a reference classifier that did not implement data augmentation or correction. Performance was compared between the two classification methods using paired *t*-tests and investigated longitudinally using correlation analyses. Correlations between performance improvements and clinical measures were also investigated. The proposed classification scheme significantly improved character selection accuracy, required for usability, as well as recall and F-scores. However, precision was reduced, and binary

accuracy was not significantly affected. Overall, BCI performance deteriorated over time with both classification methods, and latency jitter calculated with CBLE increased over time. Improvements in selection accuracies using the proposed A/C approach were greater for participants with more significant physical impairments.

Also following the results from the first aim, we extracted single-trial N100, P200, N200, and P300 amplitudes and latencies in each session using Woody-type filters on spatially filtered data. That is, spatial principal component analysis was conducted on the responses to stimuli containing the intended characters in each session, and appropriate spatial factors were selected from the results of this analysis. Then, session-average time series for these spatial factors were used as templates. The cross-covariance of the templates with the single-trial time series were calculated. The maximum value of these cross-covariances and the latency shifts to achieve this maximum were then used as single-trial amplitudes and latencies. Within-session variability in N100, P200, N200, and P300 latencies were compared between participants with ALS and neurotypical participants using Wilcoxon rank-sum tests, and P200, N200, and P300 jitter were all found to be increased in ALS. In addition, linear models were used to investigate which ERP feature latencies contributed to the shifts detected with CBLE, determining that single-trial N100, P200, N200, and P300 latencies were all significant contributors in data recorded from neurotypical participants. However, the relationships between ERP feature latencies and CBLE were disrupted in ALS, with single-trial N100 latencies no longer a significant contributor to latency shifts calculated with CBLE and reduced but still significant contributions from single-trial P200, N200, and P300 latencies. There were, however, some contributions

to jitter from single-trial ERP amplitudes, with increased latency shifts detected with both CBLE and Woody filters on trials with reduced ERP amplitudes. Considering these results, we conclude that CBLE reflects both latency jitter and other factors which affect BCI performance. Despite the increase in latency jitter calculated with CBLE over time in participants with ALS, there was not a significant increase in N100, P200, N200, or P300 jitter calculated with Woody filters over time.

Overall, the research presented in this dissertation advances knowledge on latency variability in the use of P300 BCIs, both for neurotypical participants and for people with ALS. The importance of latency jitter in P300 BCIs is elucidated, both whole-epoch jitter calculated with CBLE and latency variation in specific ERP features are shown to be increased in people with ALS, a theoretical limitation of CBLE is investigated, and a compensation strategy is proposed to address increased latency jitter in people with ALS using P300 BCIs.

ACKNOWLEDGMENTS

First, I would like to thank my major professors, Drs. Yalda Shahriari and Kunal Mankodiya, for their support throughout my time in the Interdisciplinary Neuroscience Program, helping me to become a better scientist. Their support enabled this dissertation.

I would also like to thank my committee members. Dr. Menniti has supported me in honing my presentation skills and in understanding where I want my career to go. Dr. Sodhi broadened my horizons during my undergraduate career with the study program in India, and he has continued to support me since then. Dr. Peckham's course on no-boundary thinking has served as a guide to how I want to approach my research – with as little regard for disciplinary divisions as I can reasonably manage.

While they are not on my committee, Dr. Dalton and Dr. Weyandt also deserve mention. Dr. Dalton supported my interest in communication supports for people with disabilities when I took her course on Augmentative and Alternative Communication – which P300-based brain computer interfaces, the topic of my dissertation, form a part of. The review of P300-based brain computer interfaces I did under Dr. Weyandt's guidance supported my knowledge on the topic throughout my dissertation research.

Truly, I would like to thank my entire academic community: all my lab mates, past and present, my fellow graduate students in the INP, my fellow graduate students from mathematics, my professors, and the scholars I may never have taken a class from but who nevertheless influenced my work. I look forward to being able to see you again.

Finally, I would like to thank my family and friends. Both your support and your distractions were needed – the former so I could finish, and the latter so I could enjoy life on the way.

PREFACE

This dissertation is written in manuscript format. The first chapter serves as an introduction to the dissertation as a whole, providing an overview of the main topics and a justification for the research. The first chapter additionally notes the three primary aims for the research described in the dissertation. The second through fourth chapters are the manuscripts. The first manuscript, P300 latency jitter and its correlates in people with amyotrophic lateral sclerosis, was published in the February 2021 issue of *Clinical Neurophysiology*: volume 132, issue 2, pages 632-642. This manuscript primarily addresses the first research aim. The second manuscript, Improving Longitudinal P300-BCI Performance for People with ALS Using a Data Augmentation and Jitter Correction Approach, was submitted to the *Journal of Neural Engineering* on March 4, 2021 and primarily addresses the third research aim. The third manuscript, A Longitudinal Study Latency Jitter and Disrupted Interrelationships in ALS Using a Woody Filter Approach, is in preparation for submission to *Clinical Neurophysiology*. This manuscript primarily addresses the second research aim.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

FIGURE PAGE

FIGURE PAGE

FIGURE PAGE

CHAPTER 1: INTRODUCTION AND OVERVIEW

1.1 MOTIVATION

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative condition which can affect both the upper and lower motor neurons, along with their frontotemporal connections, leading to the loss of voluntary motor control (Turner & Swash, 2015). Rates of progression and survival times vary based on a variety of factors including age at diagnosis and site of onset (Pupillo, Messina, Logroscino, Beghi, & SLALOM Group, 2014), but ALS is eventually fatal. Assistive technologies are a frequent part of care to improve quality of life, including mobility supports such as wheelchairs, life supports such as ventilators, and tools for communication and environmental control including high-tech augmentative and alternative communication (AAC) devices (Eicher et al., 2019). For people with ALS, AAC access can maintain social participation, improve self-determination, and increase quality of life (Eicher et al., 2019), and so most people with ALS accept AAC options when offered and continue to use them for as long as they are able to (Ball, Beukelman, & Pattee, 2004).

However, as ALS continues to progress, conventional assistive technologies become harder, and then impossible, to use effectively. Brain-computer interfaces (BCIs), devices which rely on a computer interpreting neural signals rather than intentional movements by the user, are a potential option for communication and control at this stage (Liberati et al., 2015). While there is a record of successful home use of BCIs (Hill, Kovacs, & Shin, 2014; Holz, Botrel, Kaufmann, & Kübler, 2015; Sellers, Vaughan, & Wolpaw, 2010; Shahriari, Yalda et al., 2019; Speier, Chandravadia,

Roberts, Pendekanti, & Pouratian, 2017; Wolpaw et al., 2018), there are also barriers to use, such as a lack of information about BCIs and their everyday applications (Linse, Aust, Joos, & Hermann, 2018) and day-to-day variations in BCI performance (Shahriari et al., 2019).

1.2 AUGMENTATIVE AND ALTERNATIVE COMMUNICATION

Augmentative and alternative communication (AAC) addresses the needs of people with speech and communication disabilities using a variety of techniques and tools. These tools include communication boards, speech generating devices, manual signs, and other electronic and nonelectronic supports. AAC can both support existing speech and serve as an alternative to speech (American Speech-Language-Hearing Association [ASHA] n.d.a). At its most inclusive, AAC covers all the ways people communicate without, or in addition to, verbal speech, including methods used by people with typical speech (ASHA n.d.b). However, AAC research is typically focused on people with disabilities affecting speech, and it may be suggested for people with ALS after they score a 90% or lower intelligibility of speech or a 100 words-per-minute or lower speaking rate on the Sentence Intelligibility Test (Ball et al., 2004; Yorkston & Beukelman, 1996). AAC systems can be based on text or symbols, and their technological requirements can be zero (e.g., gesture) or vary from low (e.g., symbol cards, or pen and paper) to high (e.g., mobile applications and dedicated AAC devices).

High-tech AAC options utilizing adapted keyboards or boards, specialized touch screens, single- or multiple-switch scanning, head tracking, and BCIs can improve communication and quality of life for people with ALS (Cipresso et al., 2012; Linse et

al., 2018), and people with ALS who are offered AAC options typically both initiate and continue use of them to the extent that they can (Ball et al., 2004). However, there are significant barriers to use (Cipresso et al., 2012) including limited information about AAC (Liberati et al., 2015), limited availability of devices (Linse et al., 2018), and issues with the technology itself (da Silva-Sauer, Valero-Aguayo, de la Torre-Luque, Ron-Angevin, & Varona-Moya, 2016; Shahriari et al., 2019) – the same barriers that are common for AAC use in general (Lund & Light, 2007).

1.3 ELECTROENCEPHALOGRAPHY (EEG)

Electroencephalography, one of the most common non-invasive methods of neural recording, involves measuring electrical potentials from electrodes on the scalp. Each electrode detects electrical potentials from thousands to millions of neurons through volume conduction, recording synchronized neural activities from a scalp area on the order of $10cm^2$ (Shahriari et al., 2020). Despite the fact that EEG recordings are affected by noise from both physiological and environmental sources, including muscular activity and other electrical equipment in the area, it can be used for a variety of applications, including diagnostics, neuroimaging, uncovering neural correlates of psychological constructs, and controlling BCIs (Biasiucci, Franceschiello, & Murray, 2019).

In an EEG recording, the electrodes can be attached to the head using spiky contacts in the case of dry electrodes, with conductive paste or gel for some medical recordings, or with a cap holding electrodes that may have gel added after donning the cap. The arrangement of electrodes in a montage can be described using the 10-20, 10-10, or 10-

5 international systems, with these numbers referring to inter-electrode intervals of 5%, 10%, or 20% of the span covered by electrodes. Letters denote placement along the anterior-posterior axis, and numbers denote placement along the left-right axis, with zero (z) representing the midline, positive even numbers representing the right side of the head, and positive odd numbers representing the left side of the head. Larger numbers mean that electrodes are further to the left or the right. (Shahriari et al., 2020). Figure 1.1 shows the montage used throughout this dissertation, with frontal channel Fz*, central channel Cz, parietal channels Pz, P3, and P4, parietal-occipital channels PO7 and PO8, and occipital channel Oz. Fz* is noted with a star as Fz itself was occupied by another sensor and the nearest available location, FAF2, was used instead.

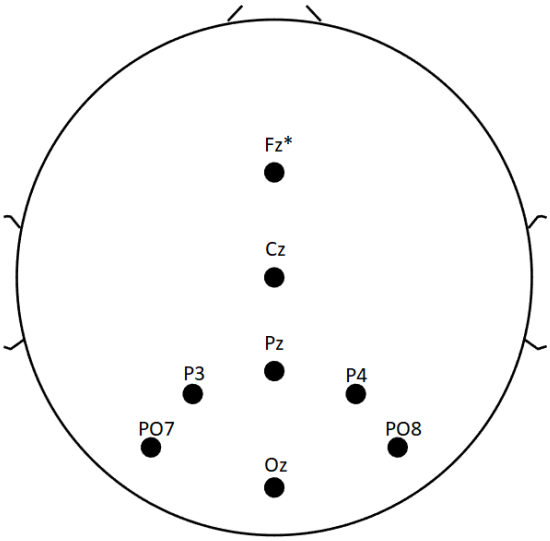

Figure 1.1. The electrode montage used in EEG recordings for the current studies.

The signals recorded from EEG can then be described in terms of oscillatory waves, which are considered in terms of frequency bands, or in terms of transient responses, such as event-related potentials (ERPs), which appear in response to certain events or

stimuli (Shahriari et al., 2020). Transitory responses, such as ERPs, are typically studied by averaging signals over several trials to increase the signal-to-noise ratio, but the responses can be detected in single trials (Biasiucci et al., 2019). However single-trial responses and their variations have been studied in cognitive contexts for some time (e.g., Kutas, McCarthy, & Donchin, 1977). Sample EEG data recorded over 20 seconds from the 8 channels used in this dissertation and bandpass filtered to allow frequencies between 0.5 and 30 Hz are shown in Figure 1.2.

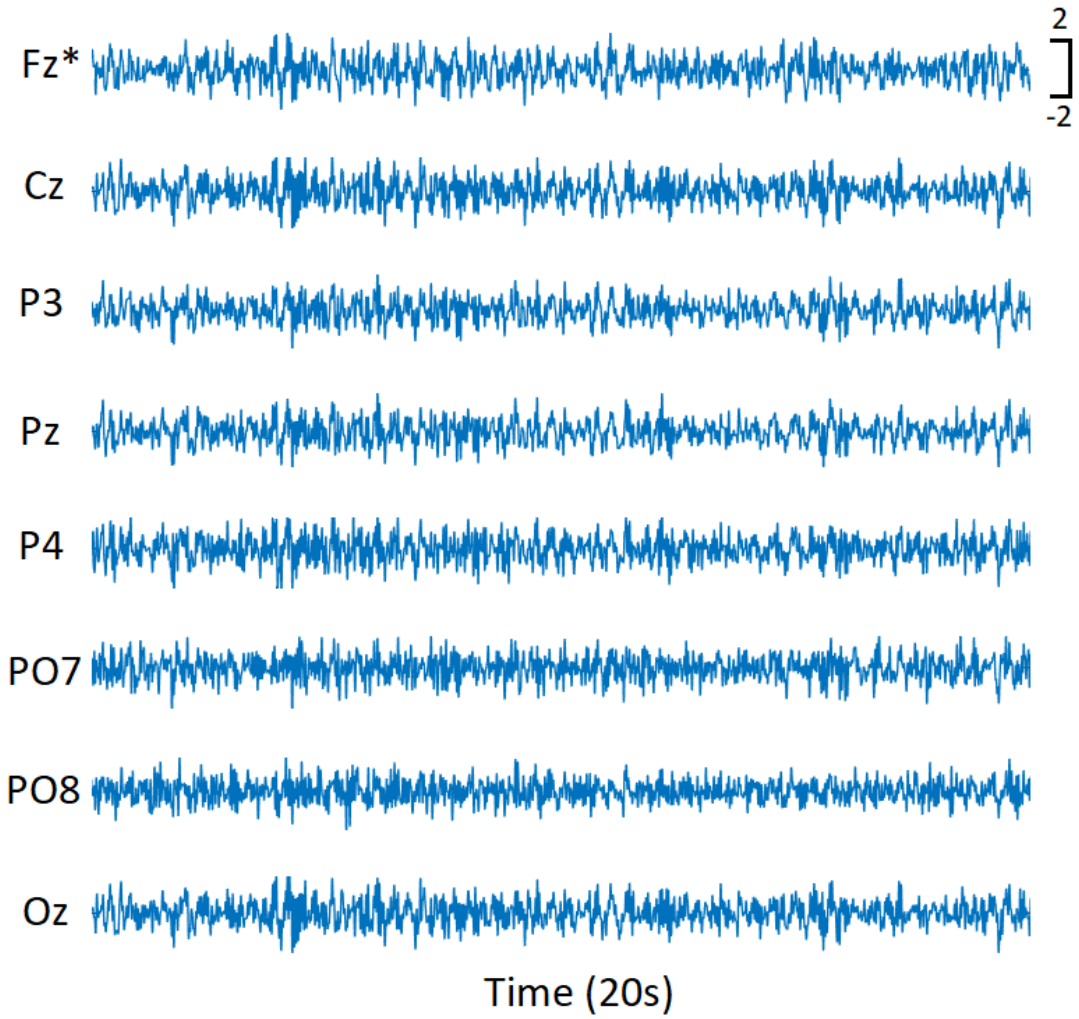

Figure 1.2. Sample EEG traces from each of the 8 channels used in the current studies, over 20 seconds.

1.4 P300-BASED BCIS

In 1988, Farwell and Donchin proposed a BCI in which the P300, an evoked electrical response to an unusual and attended event, was used to select characters, allowing users to spell words (Farwell & Donchin, 1988). In the original P300 speller, characters were arranged in a 6x6 matrix, with rows and columns intensified by randomly flashing them, while participants counted intensifications of their intended character and the computer detected the P300 response, named for the fact that it typically appears about 300 ms after the stimulus that evoked it. Like most ERP-based BCIs, P300 systems are relatively quick to learn to use (Lazarou, Nikolopoulos, Petrantonakis, Kompatsiaris, & Tsolaki, 2018), but they function as switch systems, which are generally slower than direct selection methods in daily use. The original 6x6 P300 speller matrix is shown in Figure 1.3, and an example of the target and non-target ERPs elicited by the P300 speller is shown in Figure 1.4

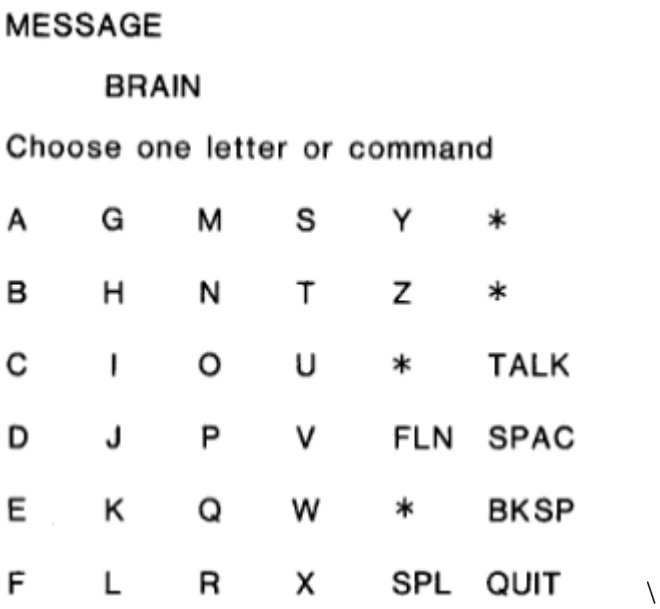

Figure 1.3. The original P300 speller matrix (Farwell & Donchin, 1988)

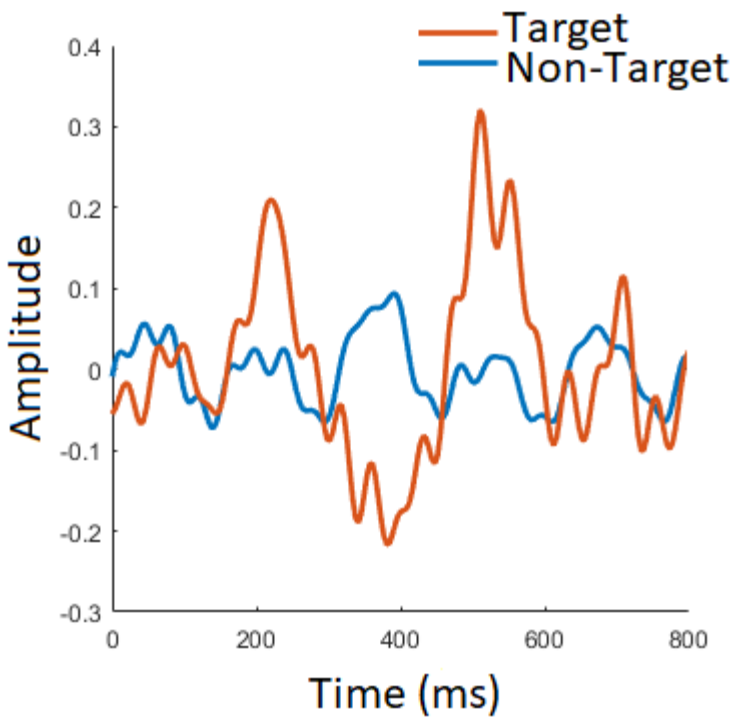

Figure 1.4. Target (P300) and non-target (no P300) ERPs.

The majority of BCI studies do not include participants with neuromuscular disabilities such as ALS, as indicated by the fewer than 10% of BCI publications that mention the term "patients". (Allison, Kübler, & Jin, 2020). Nevertheless, the P300 BCI is one of the most studied BCI paradigms, both generally (Allison et al., 2020), and for people with ALS (Kellmeyer, Grosse-Wentrup, Schulze-Bonhage, Ziemann, & Ball, 2018). As such, there are still many studies of P300 BCIs which use data collected from participants with ALS (Borgheai et al., 2020; Carabalona et al., 2012; Clements et al., 2016; Halder, S., Käthner, & Kübler, 2016; Halder, Sebastian et al., 2016; Holz et al., 2015; Hou, Li, Liu, & Wang, 2017; Kübler & Birbaumer, 2008; Kübler et al., 2009; Kübler et al., 2014; Mainsah et al., 2015; Mak et al., 2012; McCane et al., 2014; McCane et al., 2015; Mowla, Huggins, & Thompson, 2017; Pasqualotto et al., 2015; Piccione et

al., 2006; Riccio et al., 2018; Ryan, David B. et al., 2018; Schettini, Riccio et al., 2015; Sellers & Donchin, 2006; Sellers et al., 2010; Speier et al., 2017; Spüler et al., 2012; Spüler, Walter, Rosenstiel, & Bogdan, 2013; Thompson, Warschausky, & Huggins, 2012; Townsend et al., 2010).

These studies have investigated a variety of research questions relevant to the use of P300-based BCIs for people with ALS, including factors associated with BCI performance such as trial-to-trial latency variability, or latency jitter (Thompson et al., 2012; Zisk et al., 2020), the expansion of C9ORF72, a gene associated with hereditary ALS (Geronimo et al., 2017), cognitive factors (Geronimo et al., 2016; Geronimo & Simmons, 2017; Riccio et al., 2013; Riccio et al., 2018), and EEG features (Mak et al., 2012; Shahriari, Y. et al., 2013; Shahriari et al., 2019). They have also aimed to improve BCI performance and user experience. The inclusion of word prediction (Ron-Angevin et al., 2015) and language models (Speier et al., 2017) can speed up communication and make typing more convenient. Other arrangements of characters or intensification patterns have also been studied to improve BCI performance, such as region-based spellers (Ikegami et al., 2014; Severens et al., 2014), a lateral single-character speller (Pires et al., 2012), a checkerboard paradigm in which no two adjacent items are intensified at the same time (Townsend et al., 2010), flashes with colors other than black, grey, and white (Ikegami et al., 2014; Ryan, D. B., Townsend, Gates, Colwell, & Sellers, 2017; Ryan et al., 2018), or the use of face images to cover characters rather than changing the color of the characters (Fernández-Rodríguez, Velasco-Álvarez, Medina-Juliá, & Ron-Angevin, 2019; Kaufmann, Schulz, Grünzinger, & Kübler, 2011; Kaufmann et al., 2013). When comparisons are made, participants typically prefer the

checkerboard or region-based arrangements to the row-column arrangements, as well as preferring face or colored stimuli for intensification (Ikegami et al., 2014; Ryan et al., 2018; Townsend et al., 2010). Recently, alternatives to the typical task of counting intensifications of the intended character, such as mental arithmetic, and the inclusion of other recording modalities such as fNIRS, have also been proposed with promising results (Borgheai, S. B., Abtahi, Mankodiya, McLinden, & Shahriari, 2019; Borgheai et al., 2020).

Evidence suggests that P300 BCI performance is fairly stable over time (Holz et al., 2015; Sellers et al., 2010; Shahriari et al., 2019; Silvoni et al., 2009; Silvoni et al., 2013; Wolpaw et al., 2018). However, as people with ALS can develop ocular issues as their disease progresses, other stimulus modalities such as auditory (Halder et al., 2016; Kleih et al., 2015; Kübler et al., 2009; Onishi et al., 2017; Simon et al., 2015) and tactile (Guger et al., 2017; Severens et al., 2014) P300 BCIs have been considered. Despite some success with non-visual P300 BCIs for people with ALS in a completely locked-in state (Guger et al., 2017), visual EEG-based P300 BCIs such as the BCI studied in this dissertation are generally not effective at that point (Kübler & Birbaumer, 2008; Murguialday et al., 2011). Given the importance of effectiveness, reliability, and speed for people with ALS considering BCI use (Zickler et al., 2011), work to improve the technical aspects of P300 BCIs continues.

1.5 LATENCY JITTER

While the P300 occurs approximately 300 ms after an attended, unusual stimulus, it can have significant trial-to-trial latency variability, or jitter (Aricò et al., 2014; Fjell,

Rosquist, & Walhovd, 2009; Jaśkowski & Verleger, 2000; Yu, 2016). A variety of methods have been proposed to extract these single trial latencies, whether to measure jitter, correct for it, or do both. Single-trial latencies and their associated latency variations have been studied in cognitive contexts for some time, where single-trial P300 latencies are associated with stimulus evaluation times (Kelly & O'Connell, 2013; Verleger, 1997) and single-trial reaction times (Saville et al., 2011), particularly when the focus is on accuracy rather than speed (Kutas et al., 1977). However, this relationship is disrupted in neurotypical participants with comparatively higher P300 latency jitter (Saville et al., 2011) and when the task prioritizes speed (Kutas et al., 1977; Verleger, 1997). Latency variability can be studied as one form of neural variability (Magnuson, Iarocci, Doesburg, & Moreno, 2020), which is required for learning but increased in a variety of neurological conditions (Dinstein, Heeger, & Behrmann, 2015). For example, latency jitter is increased in people with ALS (Zisk et al., 2020), attention deficit hyperactivity disorder (Saville et al., 2015), schizophrenia (Ford, White, Lim, & Pfefferbaum, 1994), depression, (Patterson, Michalewski, & Starr, 1988), traumatic brain injuries (Unsal & Segalowitz, 1995), disorders of consciousness (Schettini, Risetti et al., 2015), and dementia (Patterson et al., 1988).

In the context of brain computer interfaces, latency jitter is important to consider both because increased latency jitter is associated with decreased performance (Aricò et al., 2014; Huggins, Alcaide-Aguirre, & Hill, 2016; Mowla, Gonzalez-Morales, Rico-Martinez, Ulichnie, & Thompson, 2020; Schettini et al., 2015; Thompson et al., 2012; Zisk et al., 2020; Zisk et al., 2021) and because methods which aim to compensate for

latency jitter have provided some improvements in BCI performance (Mowla et al., 2017; Togashi & Washizawa, 2013; Zisk et al., 2021).

Given the importance of latency jitter in cognitive studies and its growing recognition in the context of BCIs, a variety of methods exist to quantify this phenomenon (Fabiani, Gratton, Karis, & Donchin, 1987; Jaśkowski & Verleger, 2000; Ouyang, Hildebrandt, Sommer, & Zhou, 2017; Smulders, Kenemans, & Kok, 1994). In general, latency jitter is calculated by extracting single-trial latencies or latency shifts for responses of interest for each trial in a data set, and then using a measure of variability on the calculated latencies or latency shifts. To calculate these single-trial latencies, Woody proposed an adaptive filter using iterative cross-correlations of time-shifted single-trial responses with the averaged response (Woody, 1967). Kutas and colleagues applied this Woody filter to the P300 response to provide evidence that P300 latencies are correlated with stimulus evaluation time (Kutas et al., 1977). While a study by Verleger and colleagues found that increased decision complexity led to reduced P300 amplitudes (Verleger, Baur, Metzner, & Śmigasiewicz, 2014), Yu used principal component analysis (PCA) to extract spatial patterns for use with the Woody filter and reported that this apparent reduction in amplitude is instead due to an increase in jitter (Yu, 2016). Other groups utilized filtering and peak-picking methods (Ouyang et al., 2017), ranging from simple but harsh lowpass filters (Smulders et al., 1994, Jaśkowski and Verleger, 2000) to more complex methods including wavelet-based filtering (Aricò et al., 2014; Chennu, Craston, Wyble, & Bowman, 2009). Both peak-picking and cross-correlational approaches can be applied to spatially filtered data (Coles, Gratton, Kramer, & Miller, 1986; Fabiani, Karis, & Donchin, 1986; Fabiani et al., 1987; Saville

et al., 2011; Saville et al., 2015; Yu, 2016), and cross-correlational approaches similar to the Woody filter applied to time series combining data from multiple channels are a comparatively reliable way to assess both single-trial amplitudes and latencies (Fabiani et al., 1987). Woody filtering approaches have also been used to extract single-trial latency information for several ERP features simultaneously, using appropriate segmentation of time series (Michalewski, Prasher, & Starr, 1986; Patterson et al., 1988).

As another method of quantifying single-trial latency shifts and latency jitter, Thompson and colleagues introduced classifier-based latency estimation (CBLE) as an additional way to quantify P300 jitter and thereby predict BCI performance (Thompson et al., 2012). In their proposed classifier-based latency estimation method, a classifier is applied to epochs starting at multiple time points surrounding target stimulus presentation, and the time shift corresponding to the highest classifier score is selected as the latency shift for the corresponding epoch. The authors suggested their proposed method can be utilized with any linear classification algorithm and effectively predicts BCI performance (Thompson et al., 2012). They additionally used their classifier-based latency estimation method on simulated data to evaluate the accuracy of their proposed technique on data with known jitter and further confirmed the presence of P300 jitter in data from people with ALS (Thompson et al., 2019).

1.6 THIS DISSERTATION

This dissertation is focused on latency jitter, with three primary research aims:

Research Aim 1: Quantify latency jitter and its correlates in people with ALS. Using classifier-based latency estimation, this aim will compare latency variability, or

jitter between people with ALS and neurotypical controls and examine within-group correlations of latency jitter with session average ERP features in both groups and clinical features in people with ALS.

Research Aim 2: Longitudinally investigate single-trial ERP feature variations, session-average ERP feature variations, latency jitter, and their relationships. This aim will extract single-trial ERP features from longitudinal BCI recordings. This aim will then investigate these features longitudinally and examine relationships between variabilities in single-trial features, session-average features, and latency jitter.

Research Aim 3: Develop and evaluate a correction method to compensate for latency jitter. As BCI performance does not seem to show negative trends over time, but does show day-to-day variations and correlation with jitter, I hypothesize that correcting for jitter will improve performance and reduce day-to-day performance variability for some BCI users with ALS.

Aims 2 and 3 build on aim 1, but in different directions: aim 2 is primarily exploratory, while aim 3 is primarily about improving BCI performance.

Throughout this dissertation, latency shifts and jitter are quantified using CBLE, and Woody filters are additionally used to extract single-trial features in the third manuscript (Aim 2). Data was collected longitudinally from six participants with ALS, and in 2-3 sessions each from neurotypical participants, while participants used a P300 speller. All participants have at least some post-secondary education. All neurotypical control participants have normal or corrected to normal vision, as do participants with ALS other than ALS-1, who is in the late stages of locked-in syndrome with significant

ocular impairments. Additional details about participants with ALS are provided in each manuscript as applicable.

1.6.1 Manuscript 1

In the first manuscript, three sessions from each participant with ALS were used, and nine neurotypical control participants also participated in data recordings. The aim was to quantify latency jitter and its correlates in people with ALS, as well as to determine whether jitter was increased in ALS. ERP amplitudes and latencies were extracted. Classifier-based latency estimation (CBLE) was used to calculate latency jitter. ERP components and latency jitter were compared between groups using Wilcoxon rank-sum tests. Correlations between latency jitter and each of the clinical measures, ERP features, and performance measures were investigated using Spearman and repeated measures correlations. We found that latency jitter, calculated with CBLE, was significantly increased in participants with ALS compared to neurotypical control participants. Latency jitter correlated with BCI performance in both groups, but not with clinical measures for participants with ALS.

1.6.2 Manuscript 2

In the second manuscript, longitudinal recordings were used from participants with ALS, and there were no neurotypical control participants. The aim was to improve BCI performance for people with ALS. We proposed an augmentation and correction (A/C) classification scheme including data augmentation and correction for jitter, both relying on time-shifted responses with individualized parameters determined based on latency

jitter. The proposed A/C classification scheme significantly improved character selection accuracy, required for usability, as well as recall and F-scores. However, precision was reduced, and binary accuracy was not significantly affected. Overall, BCI performance deteriorated over time with both classification methods. Selection accuracies were more improved by the proposed A/C approach for participants with more significant physical impairments. Both data augmentation and latency jitter compensation can potentially improve BCI performance for people with ALS.

1.6.3 Manuscript 3

For the third manuscript, longitudinal recordings were used from participants with ALS, and sixteen neurotypical participants each underwent 2-3 data recording sessions. This study investigated latency jitter calculated using two different methods to understand which single-trial features are represented in classifier-based latency estimation (CBLE), compare latency jitter for specific ERP features between people with ALS and neurotypical controls, and longitudinally investigate latency. Both single-trial and session-average ERP amplitudes and latencies were extracted. Both a Woody filtering approach and CBLE were used to calculate latency jitter. ERP components and latency jitter were compared between groups using Wilcoxon rank-sum tests. Relationships between measures were investigated within and between sessions using linear regression models, Spearman correlations, and repeated measures correlations. Latency variations in the four ERP features considered, the N100, P200, N200, and P300, all contributed to whole-epoch latency variations calculated with CBLE in neurotypical participants. However, these contributions were disrupted in participants

with ALS, who also had increased P200, N200, P300, and whole-epoch latency jitter. Whole-epoch latency jitter increased over time in people with ALS, but N100, P200, N200, and P300 jitter did not. Neither whole epoch nor ERP feature latency jitter correlated with clinical scores in participants with ALS.

1.6.4 General Conclusions

Combined, the three manuscripts show that latency jitter is increased in people with ALS for multiple ERP features including the P300, though jitter is not correlated with clinical scores. Correction for latency jitter relying on whole-epoch shifts calculated with CBLE can effectively improve BCI performance for people with ALS. The manuscripts additionally show that CBLE effectively reflects latency variations in the N100, P200, N200, and P300, though this is disrupted somewhat in people with ALS.

REFERENCES

Allison, B. Z., Kübler, A., & Jin, J. (2020). 30 years of P300 brain–computer interfaces. *Psychophysiology*, *57*(7), e13569.

Aricò, P., Aloise, F., Schettini, F., Salinari, S., Mattia, D., & Cincotti, F. (2014). Influence of P300 latency jitter on event related potential-based brain–computer interface performance. *Journal of Neural Engineering*, *11*(3), 035008.

American Speech-Language-Hearing Association. (n.d.a.) Augmentative and alternative communication. https://www.asha.org/Practice-Portal/Professional-Issues/Augmentative-and-Alternative-Communication Accessed March 26, 2021.

American Speech-Language-Hearing Association. (n.d.b) Augmentative and alternative communication (AAC). https://www.asha.org/public/speech/disorders/AAC Accessed March 26, 2021.

Ball, L. J., Beukelman, D. R., & Pattee, G. L. (2004). Acceptance of augmentative and alternative communication technology by persons with amyotrophic lateral sclerosis. *Augmentative and Alternative Communication*, *20*(2), 113-122.

Biasiucci, A., Franceschiello, B., & Murray, M. M. (2019). Electroencephalography. *Current Biology, 29*(3), R80-R85.

Borgheai, S. B., Deligani, R. J., McLinden, J., Zisk, A. H., Hosni, S. I., Abtahi, M., . . . Shahriari, Y. (2019). Multimodal exploration of non-motor neural functions in ALS patients using simultaneous EEG-fNIRS recording. *Journal of Neural Engineering,*

Borgheai, S. B., McLinden, J., Zisk, A. H., Hosni, S. I., Deligani, R. J., Abtahi, M., . . . Shahriari, Y. (2020). Enhancing communication for people in late-stage ALS using an fNIRS-based BCI system. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28*(5), 1198-1207.

Carabalona, R., Grossi, F., Tessadri, A., Castiglioni, P., Caracciolo, A., & de Munari, I. (2012). Light on! real world evaluation of a P300-based brain–computer interface (BCI) for environment control in a smart home. *Ergonomics, 55*(5), 552-563.

Chennu, S., Craston, P., Wyble, B., & Bowman, H. (2009). Attention increases the temporal precision of conscious perception: Verifying the neural-ST2 model. *PLoS Computational Biology, 5*(11), e1000576.

Cipresso, P., Carelli, L., Solca, F., Meazzi, D., Meriggi, P., Poletti, B., . . . Riva, G. (2012). The use of P300-based BCIs in amyotrophic lateral sclerosis: From augmentative and alternative communication to cognitive assessment. *Brain and Behavior, 2*(4), 479-498.

Clements, J. M., Sellers, E. W., Ryan, D. B., Caves, K., Collins, L. M., & Throckmorton, C. S. (2016). Applying dynamic data collection to improve dry electrode system performance for a P300-based brain–computer interface. *Journal of Neural Engineering, 13*(6), 066018.

Coles, M. G., Gratton, G., Kramer, A. F., & Miller, G. A. (1986). Principles of signal acquisition and analysis. In M. G. Coles, E. Donchin & S. W. Porges (Eds.), *Psychophysiology: Systems, processes and applications* (pp. 183-221). New York: Guilford Press.

da Silva-Sauer, L., Valero-Aguayo, L., de la Torre-Luque, A., Ron-Angevin, R., & Varona-Moya, S. (2016). Concentration on performance with P300-based BCI systems: A matter of interface features. *Applied Ergonomics, 52,* 325-332.

Dinstein, I., Heeger, D. J., & Behrmann, M. (2015). Neural variability: Friend or foe? *Trends in Cognitive Sciences, 19*(6), 322-328.

Eicher, C., Kiselev, J., Brukamp, K., Kiemel, D., Spittel, S., Maier, A., . . . Greuèl, M. (2019). Expectations and concerns emerging from experiences with assistive technology for ALS patients. Paper presented at the *International Conference on Human-Computer Interaction,* 57-68.

Erlbeck, H., Mochty, U., Kübler, A., & Real, R. G. L. (2017). Circadian course of the P300 ERP in patients with amyotrophic lateral sclerosis - implications for brain-computer interfaces (BCI). *BMC Neurology, 17*(1), 3. doi:10.1186/s12883-016-0782-1

Fabiani, M., Gratton, G., Karis, D., & Donchin, E. (1987). Definition, identification, and reliability of measurement of the P300 component of the event-related brain potential. *Advances in Psychophysiology, 2*(3), 1-78.

Fabiani, M., Karis, D., & Donchin, E. (1986). P300 and recall in an incidental memory paradigm. *Psychophysiology, 23*(3), 298-308.

Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology, 70*(6), 510-523.

Fernández-Rodríguez, Á, Velasco-Álvarez, F., Medina-Juliá, M. T., & Ron-Angevin, R. (2019). Evaluation of emotional and neutral pictures as flashing stimuli using a P300 brain–computer interface speller. *Journal of Neural Engineering, 16*(5), 056024.

Fjell, A. M., Rosquist, H., & Walhovd, K. B. (2009). Instability in the latency of P3a/P3b brain potentials and cognitive function in aging. *Neurobiology of Aging, 30*(12), 2065-2079.

Ford, J. M., White, P., Lim, K. O., & Pfefferbaum, A. (1994). Schizophrenics have fewer and smaller P300s: A single-trial analysis. *Biological Psychiatry, 35*(2), 96-103.

Geronimo, A., Simmons, Z., & Schiff, S. J. (2016). Performance predictors of brain–computer interfaces in patients with amyotrophic lateral sclerosis. *Journal of Neural Engineering, 13*(2), 026002.

Geronimo, A. M., & Simmons, Z. (2017). The P300 'face' speller is resistant to cognitive decline in ALS. *Brain-Computer Interfaces, 4*(4), 225-235.

Geronimo, A., Sheldon, K. E., Broach, J. R., Simmons, Z., & Schiff, S. J. (2017). Expansion of C9ORF72 in amyotrophic lateral sclerosis correlates with brain-computer interface performance. *Scientific Reports, 7*(1), 1-6.

Guger, C., Spataro, R., Allison, B. Z., Heilinger, A., Ortner, R., Cho, W., & La Bella, V. (2017). Complete locked-in and locked-in patients: Command following assessment and communication with vibro-tactile P300 and motor imagery brain-computer interface tools. *Frontiers in Neuroscience, 11*, 251.

Guy, V., Soriani, M., Bruno, M., Papadopoulo, T., Desnuelle, C., & Clerc, M. (2018). Brain computer interface with the P300 speller: Usability for disabled people with amyotrophic lateral sclerosis. *Annals of Physical and Rehabilitation Medicine, 61*(1), 5-11.

Halder, S., Käthner, I., & Kübler, A. (2016). Training leads to increased auditory brain–computer interface performance of end-users with motor impairments. *Clinical Neurophysiology, 127*(2), 1288-1296.

Halder, S., Pinegger, A., Käthner, I., Wriessnegger, S. C., Faller, J., Antunes, J. B. P., . . . Kübler, A. (2015). Brain-controlled applications using dynamic P300 speller matrices. *Artificial Intelligence in Medicine, 63*(1), 7-17.

Halder, S., Ruf, C. A., Furdea, A., Pasqualotto, E., De Massari, D., van der Heiden, L., . . . Kübler, A. (2013). Prediction of P300 BCI aptitude in severe motor impairment. *PloS One, 8*(10), e76148.

Halder, S., Takano, K., Ora, H., Onishi, A., Utsumi, K., & Kansaku, K. (2016). An evaluation of training with an auditory P300 brain-computer interface for the Japanese hiragana syllabary. *Frontiers in Neuroscience, 10*, 446.

Hill, K., Kovacs, T., & Shin, S. (2014). Reliability of brain-computer interface language sample transcription procedures. *Journal of Rehabilitation Research and Development, 51*(4), 579-590.

Holz, E. M., Botrel, L., Kaufmann, T., & Kübler, A. (2015). Long-term independent brain-computer interface home use improves quality of life of a patient in the locked-in state: A case study. *Archives of Physical Medicine and Rehabilitation, 96*(3), S16-S26.

Hou, J., Li, Y., Liu, H., & Wang, S. (2017). Improving the P300-based brain-computer interface with transfer learning. Paper presented at the *2017 8th International IEEE/EMBS Conference on Neural Engineering (NER)*, 485-488.

Huggins, J. E., Alcaide-Aguirre, R. E., & Hill, K. (2016). Effects of text generation on P300 brain-computer interface performance. *Brain-Computer Interfaces, 3*(2), 112-120.

Ikegami, S., Takano, K., Kondo, K., Saeki, N., & Kansaku, K. (2014). A region-based two-step P300-based brain–computer interface for patients with amyotrophic lateral sclerosis. *Clinical Neurophysiology, 125*(11), 2305-2312.

Iqbal, S., Rizvi, B. A., Shanir, P. M., Khan, Y. U., & Farooq, O. (2017). Detecting P300 potential for speller BCI. Paper presented at the *2017 International Conference on Communication and Signal Processing (ICCSP)*, 295.

Jarmolowska, J., Turconi, M. M., Busan, P., Mei, J., & Battaglini, P. P. (2013). A multimenu system based on the P300 component as a time saving procedure for communication with a brain-computer interface. *Frontiers in Neuroscience, 7*, 39.

Jaśkowski, P., & Verleger, R. (2000). An evaluation of methods for single-trial estimation of P3 latency. *Psychophysiology, 37*(2), 153-162.

Käthner, I., Halder, S., Hintermüller, C., Espinosa, A., Guger, C., Miralles, F., . . . Solà, M. (2017). A multifunctional brain-computer interface intended for home use: An evaluation with healthy participants and potential end users with dry and gel-based electrodes. *Frontiers in Neuroscience, 11*, 286.

Käthner, I., Kübler, A., & Halder, S. (2015). Comparison of eye tracking, electrooculography and an auditory brain-computer interface for binary communication: A case study with a participant in the locked-in state. *Journal of Neuroengineering and Rehabilitation, 12*(1), 76.

Kaufmann, T., Schulz, S. M., Grünzinger, C., & Kübler, A. (2011). Flashing characters with famous faces improves ERP-based brain–computer interface performance. *Journal of Neural Engineering, 8*(5), 056016.

Kaufmann, T., Schulz, S. M., Köblitz, A., Renner, G., Wessig, C., & Kübler, A. (2013). Face stimuli effectively prevent brain–computer interface inefficiency in patients with neurodegenerative disease. *Clinical Neurophysiology, 124*(5), 893-900.

Kellmeyer, P., Grosse-Wentrup, M., Schulze-Bonhage, A., Ziemann, U., & Ball, T. (2018). Electrophysiological correlates of neurodegeneration in motor and non-motor brain regions in amyotrophic lateral sclerosis-implications for brain-computer interfacing. *Journal of Neural Engineering, 15*(4), 041003. doi:10.1088/1741-2552/aabfa5

Kelly, S. P., & O'Connell, R. G. (2013). Internal and external influences on the rate of sensory evidence accumulation in the human brain. *Journal of Neuroscience, 33*(50), 19434-19441.

Kleih, S. C., Herweg, A., Kaufmann, T., Staiger-Sälzer, P., Gerstner, N., & Kübler, A. (2015). The WIN-speller: A new intuitive auditory brain-computer interface spelling application. *Frontiers in Neuroscience, 9*, 346.

Kübler, A., & Birbaumer, N. (2008). Brain–computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients? *Clinical Neurophysiology, 119*(11), 2658-2666.

Kübler, A., Furdea, A., Halder, S., Hammer, E. M., Nijboer, F., & Kotchoubey, B. (2009). A brain–computer interface controlled auditory event-related potential (P300) spelling system for locked-in patients. *Annals of the New York Academy of Sciences, 1157*(1), 90-100.

Kübler, A., Holz, E. M., Riccio, A., Zickler, C., Kaufmann, T., Kleih, S. C., . . . Mattia, D. (2014). The user-centered design as novel perspective for evaluating the usability of BCI-controlled applications. *PLoS One, 9*(12), e112392.

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science, 197*(4305), 792-795.

Lazarou, I., Nikolopoulos, S., Petrantonakis, P. C., Kompatsiaris, I., & Tsolaki, M. (2018). EEG-based brain–computer interfaces for communication and rehabilitation of people with motor impairment: A novel approach of the 21st century. *Frontiers in Human Neuroscience, 12*, 14.

Liberati, G., Pizzimenti, A., Simione, L., Riccio, A., Schettini, F., Inghilleri, M., . . . Cincotti, F. (2015). Developing brain-computer interfaces from a user-centered perspective: Assessing the needs of persons with amyotrophic lateral sclerosis, caregivers, and professionals. *Applied Ergonomics, 50*, 139-146.

Linse, K., Aust, E., Joos, M., & Hermann, A. (2018). Communication matters–pitfalls and promise of hightech communication devices in palliative care of severely physically disabled patients with amyotrophic lateral sclerosis. *Frontiers in Neurology, 9*, 603.

Lund, S. K., & Light, J. (2007). Long-term outcomes for individuals who use augmentative and alternative communication: Part III–Contributing factors. *Augmentative and Alternative Communication, 23*(4), 323-335.

Magnuson, J. R., Iarocci, G., Doesburg, S. M., & Moreno, S. (2020). Increased Intra-Subject variability of reaction times and Single-Trial Event-Related potential components in children with autism spectrum disorder. *Autism Research, 13*(2), 221-229.

Mainsah, B. O., Collins, L. M., Colwell, K. A., Sellers, E. W., Ryan, D. B., Caves, K., & Throckmorton, C. S. (2015). Increasing BCI communication rates with dynamic stopping towards more practical use: An ALS study. *Journal of Neural Engineering, 12*(1), 016013.

Mainsah, B. O., Collins, L. M., & Throckmorton, C. S. (2016). Using the detectability index to predict P300 speller performance. *Journal of Neural Engineering, 13*(6), 066007.

Mak, J. N., McFarland, D. J., Vaughan, T. M., McCane, L. M., Tsui, P. Z., Zeitlin, D. J., . . . Wolpaw, J. R. (2012). EEG correlates of P300-based brain–computer interface (BCI) performance in people with amyotrophic lateral sclerosis. *Journal of Neural Engineering, 9*(2), 026014.

Manyakov, N. V., Chumerin, N., Combaz, A., & Van Hulle, M. M. (2011). Comparison of classification methods for P300 brain-computer interface on disabled subjects. *Computational Intelligence and Neuroscience, 2011*

McCane, L. M., Heckman, S. M., McFarland, D. J., Townsend, G., Mak, J. N., Sellers, E. W., . . . Vaughan, T. M. (2015). P300-based brain-computer interface (BCI) event-related potentials (ERPs): People with amyotrophic lateral sclerosis (ALS) vs. age-matched controls. *Clinical Neurophysiology, 126*(11), 2124-2131.

McCane, L. M., Sellers, E. W., Mcfarland, D. J., Mak, J. N., Carmack, C. S., Zeitlin, D., . . . Vaughan, T. M. (2014). Brain-computer interface (BCI) evaluation in people with amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 15*(3-4), 207-215.

Mowla, M. R., Gonzalez-Morales, J. D., Rico-Martinez, J., Ulichnie, D. A., & Thompson, D. E. (2020). A comparison of classification techniques to predict brain-computer interfaces accuracy using classifier-based latency estimation. *Brain Sciences, 10*(10), 734.

Mowla, M. R., Huggins, J. E., & Thompson, D. E. (2017). Enhancing P300-BCI performance using latency estimation. *Brain-Computer Interfaces, 4*(3), 137-145.

Mugler, E. M., Ruf, C. A., Halder, S., Bensch, M., & Kubler, A. (2010). Design and implementation of a P300-based brain-computer interface for controlling an internet browser. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 18*(6), 599-609.

Münßinger, J. I., Halder, S., Kleih, S. C., Furdea, A., Raco, V., Hösle, A., & Kubler, A. (2010). Brain painting: First evaluation of a new brain–computer interface application with ALS-patients and healthy volunteers. *Frontiers in Neuroscience, 4*, 182.

Nijboer, F., Birbaumer, N., & Kubler, A. (2010). The influence of psychological state and motivation on brain–computer interface performance in patients with amyotrophic lateral sclerosis–a longitudinal study. *Frontiers in Neuroscience, 4*, 55.

Nijboer, F., Sellers, E. W., Mellinger, J., Jordan, M. A., Matuz, T., Furdea, A., . . . Vaughan, T. M. (2008). A P300-based brain–computer interface for people with amyotrophic lateral sclerosis. *Clinical Neurophysiology, 119*(8), 1909-1916.

Onishi, A., Takano, K., Kawase, T., Ora, H., & Kansaku, K. (2017). Affective stimuli for an auditory P300 brain-computer interface. *Frontiers in Neuroscience, 11*, 522.

Ouyang, G., Hildebrandt, A., Sommer, W., & Zhou, C. (2017). Exploiting the intra-subject latency variability from single-trial event-related potentials in the P3 time range: A review and comparative evaluation of methods. *Neuroscience & Biobehavioral Reviews, 75*, 1-21.

Pasqualotto, E., Matuz, T., Federici, S., Ruf, C. A., Bartl, M., Olivetti Belardinelli, M., . . . Halder, S. (2015). Usability and workload of access technology for people with severe motor impairment: A comparison of brain-computer interfacing and eye tracking. *Neurorehabilitation and Neural Repair, 29*(10), 950-957.

Patterson, J. V., Michalewski, H. J., & Starr, A. (1988). Latency variability of the components of auditory event-related potentials to infrequent stimuli in aging,

alzheimer-type dementia, and depression. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 71*(6), 450-460.

Piccione, F., Giorgi, F., Tonin, P., Priftis, K., Giove, S., Silvoni, S., . . . Beverina, F. (2006). P300-based brain computer interface: Reliability and performance in healthy and paralysed participants. *Clinical Neurophysiology, 117*(3), 531-537.

Pires, G., Nunes, U., & Castelo-Branco, M. (2011). Statistical spatial filtering for a P300-based BCI: Tests in able-bodied, and patients with cerebral palsy and amyotrophic lateral sclerosis. *Journal of Neuroscience Methods, 195*(2), 270-281.

Pires, G., Nunes, U., & Castelo-Branco, M. (2012). Comparison of a row-column speller vs. a novel lateral single-character speller: Assessment of BCI for severe motor disabled patients. *Clinical Neurophysiology, 123*(6), 1168-1181.

Poletti, B., Carelli, L., Solca, F., Lafronza, A., Pedroli, E., Faini, A., . . . Cipresso, P. (2016). Cognitive assessment in amyotrophic lateral sclerosis by means of P300-brain computer interface: A preliminary study. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 17*(7-8), 473-481.

Pupillo, E., Messina, P., Logroscino, G., Beghi, E., & SLALOM Group. (2014). Long-term survival in amyotrophic lateral sclerosis: A population-based study. *Annals of Neurology, 75*(2), 287-297.

Riccio, A., Schettini, F., Simione, L., Pizzimenti, A., Inghilleri, M., Olivetti-Belardinelli, M., . . . Cincotti, F. (2018). On the relationship between attention processing and P300-based brain computer interface control in amyotrophic lateral sclerosis. *Frontiers in Human Neuroscience, 12*, 165.

Riccio, A., Simione, L., Schettini, F., Pizzimenti, A., Inghilleri, M., Olivetti Belardinelli, M., . . . Cincotti, F. (2013). Attention and P300-based BCI performance in people with amyotrophic lateral sclerosis. *Frontiers in Human Neuroscience, 7* doi:10.3389/fnhum.2013.00732

Ron-Angevin, R., Varona-Moya, S., & da Silva-Sauer, L. (2015). Initial test of a T9-like P300-based speller by an ALS patient. *Journal of Neural Engineering, 12*(4), 046023.

Ryan, D. B., Townsend, G., Gates, N. A., Colwell, K., & Sellers, E. W. (2017). Evaluating brain-computer interface performance using color in the P300 checkerboard speller. *Clinical Neurophysiology, 128*(10), 2050-2057.

Ryan, D. B., Colwell, K. A., Throckmorton, C. S., Collins, L. M., Caves, K., & Sellers, E. W. (2018). Evaluating brain-computer interface performance in an ALS population: Checkerboard and color paradigms. *Clinical EEG and Neuroscience, 49*(2), 114-121.

Saville, C. W., Dean, R. O., Daley, D., Intriligator, J., Boehm, S., Feige, B., & Klein, C. (2011). Electrocortical correlates of intra-subject variability in reaction times: Average and single-trial analyses. *Biological Psychology, 87*(1), 74-83.

Saville, C. W., Feige, B., Kluckert, C., Bender, S., Biscaldi, M., Berger, A., . . . Klein, C. (2015). Increased reaction time variability in attention-deficit hyperactivity disorder as a response-related phenomenon: Evidence from single-trial event-related potentials. *Journal of Child Psychology and Psychiatry, 56*(7), 801-813.

Schettini, F., Riccio, A., Simione, L., Liberati, G., Caruso, M., Frasca, V., . . . Inghilleri, M. (2015). Assistive device with conventional, alternative, and brain-computer interface inputs to enhance interaction with the environment for people with amyotrophic lateral sclerosis: A feasibility and usability study. *Archives of Physical Medicine and Rehabilitation, 96*(3), S46-S53.

Schettini, F., Risetti, M., Aricò, P., Formisano, R., Babiloni, F., Mattia, D., & Cincotti, F. (2015). P300 latency jitter occurrence in patients with disorders of consciousness: Toward a better design for brain computer interface applications. Paper presented at the *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* 6178-6181.

Sellers, E. W., & Donchin, E. (2006). A P300-based brain–computer interface: Initial tests by ALS patients. *Clinical Neurophysiology, 117*(3), 538-548.

Sellers, E. W., Vaughan, T. M., & Wolpaw, J. R. (2010). A brain-computer interface for long-term independent home use. *Amyotrophic Lateral Sclerosis, 11*(5), 449-455.

Severens, M., Van der Waal, M., Farquhar, J., & Desain, P. (2014). Comparing tactile and visual gaze-independent brain–computer interfaces in patients with amyotrophic lateral sclerosis and healthy users. *Clinical Neurophysiology, 125*(11), 2297-2304.

Shahriari, Y., Besio, W., Hosni, S. I., Zisk, A. H., Borgheai, S. B., Deligani, R. J., & McLinden, J. (2020). Electroencephalography. In *Neural Interface Engineering* (pp. 1-16). Springer, Cham.

Shahriari, Y., Vaughan, T. M., Corda, D. E., Zeitlin, D., Wolpaw, J. R., & Krusienski, D. J. (2013). EEG correlates of performance during long-term use of a P300 BCI by individuals with amyotrophic lateral sclerosis. Paper presented at the *Proceedings of the Fifth International Brain-Computer Interface Meeting,*

Shahriari, Y., Vaughan, T. M., McCane, L., Allison, B. Z., Wolpaw, J. R., & Krusienski, D. J. (2019). An exploration of BCI performance variations in people with amyotrophic lateral sclerosis using longitudinal EEG data. *Journal of Neural Engineering.*

Silvoni, S., Cavinato, M., Volpato, C., Ruf, C. A., Birbaumer, N., & Piccione, F. (2013). Amyotrophic lateral sclerosis progression and stability of brain-computer interface communication. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 14*(5-6), 390-396.

Silvoni, S., Volpato, C., Cavinato, M., Marchetti, M., Priftis, K., Merico, A., . . . Piccione, F. (2009). P300-based brain-computer interface communication: Evaluation and follow-up in amyotrophic lateral sclerosis. *Frontiers in Neuroscience, 3*, 1.

Simon, N., Käthner, I., Ruf, C. A., Pasqualotto, E., Kübler, A., & Halder, S. (2015). An auditory multiclass brain-computer interface with natural stimuli: Usability evaluation with healthy participants and a motor impaired end user. *Frontiers in Human Neuroscience, 8*, 1039.

Smulders, F. T., Kenemans, J. L., & Kok, A. (1994). A comparison of different methods for estimating single-trial P300 latencies. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 92*(2), 107-114.

Speier, W., Chandravadia, N., Roberts, D., Pendekanti, S., & Pouratian, N. (2017). Online BCI typing using language model classifiers by ALS patients in their homes. *Brain-Computer Interfaces, 4*(1-2), 114-121.

Spüler, M., Bensch, M., Kleih, S., Rosenstiel, W., Bogdan, M., & Kübler, A. (2012). Online use of error-related potentials in healthy users and people with severe motor impairment increases performance of a P300-BCI. *Clinical Neurophysiology, 123*(7), 1328-1337.

Spüler, M., Walter, A., Rosenstiel, W., & Bogdan, M. (2013). Spatial filtering based on canonical correlation analysis for classification of evoked or event-related potentials in EEG data. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 22*(6), 1097-1103.

Thompson, D. E., Warschausky, S., & Huggins, J. E. (2012). Classifier-based latency estimation: A novel way to estimate and predict BCI accuracy. *Journal of Neural Engineering, 10*(1), 016006.

Togashi, R., & Washizawa, Y. (2013). Feature extraction of P300 signal using Bayesian delay time estimation. Paper presented at the *2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, 1-5.

Townsend, G., LaPallo, B. K., Boulay, C. B., Krusienski, D. J., Frye, G. E., Hauser, C., . . . Sellers, E. W. (2010). A novel P300-based brain–computer interface stimulus presentation paradigm: Moving beyond rows and columns. *Clinical Neurophysiology, 121*(7), 1109-1120.

Turner, M. R., & Swash, M. (2015). The expanding syndrome of amyotrophic lateral sclerosis: A clinical and molecular odyssey. *J Neurol Neurosurgery Psychiatry, 86*(6), 667-673.

Unsal, A., & Segalowitz, S. J. (1995). Sources of P300 attenuation after head injury: Single-trial amplitude, latency jitter, and EEG power. *Psychophysiology, 32*(3), 249-256.

Verleger, R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology, 34*(2), 131-156.

Verleger, R., Baur, N., Metzner, M. F., & Śmigasiewicz, K. (2014). The hard oddball: Effects of difficult response selection on stimulus-related P 3 and on response-related negative potentials. *Psychophysiology, 51*(11), 1089-1100.

Wolpaw, J. R., Bedlack, R. S., Reda, D. J., Ringer, R. J., Banks, P. G., Vaughan, T. M., . . . Winden, S. (2018). Independent home use of a brain-computer interface by people with amyotrophic lateral sclerosis. *Neurology, 91*(3), e258-e267.

Woody, C. D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Medical and Biological Engineering, 5*(6), 539-554.

Yorkston, K. M., & Beukelman, D. R. (1996). Sentence intelligibility test. Lincoln, NE: Tice technology services.

Yu, X. (2016). *The impact of latency jitter on the interpretation of P300 in the assessment of cognitive function*

Yu, X., da Silva-Sauer, L., & Donchin, E. (2020). Habituation of P300 in the use of P300-based brain-computer interface spellers: Individuals with amyotrophic lateral sclerosis versus age-matched controls. *Clinical EEG and Neuroscience*, 1550059420918755.

Zickler, C., Halder, S., Kleih, S. C., Herbert, C., & Kübler, A. (2013). Brain painting: Usability testing according to the user-centered design in end users with severe motor paralysis. *Artificial Intelligence in Medicine, 59*(2), 99-110.

Zickler, C., Riccio, A., Leotta, F., Hillian-Tress, S., Halder, S., Holz, E., . . . Mattia, D. (2011). A brain-computer interface as input channel for a standard assistive technology software. *Clinical EEG and Neuroscience, 42*(4), 236-244.

Zisk, A. H., Borgheai, S. B., McLinden, J., Hosni, S. M., Deligani, R. J., & Shahriari, Y. (2020). P300 latency jitter and its correlates in people with amyotrophic lateral sclerosis. *Clinical Neurophysiology,*

Zisk, A. H., Borgheai, S. B., McLinden, J., & Shahriari, Y. (2021). Improving longitudinal P300-BCI performance for people with ALS using a data augmentation and jitter correction approach. *Under Review at Journal of Neural Engineering,*

CHAPTER 2: P300 LATENCY JITTER AND ITS CORRELATES IN PEOPLE WITH AMYOTROPHIC LATERAL SCLEROSIS

Published February 2021 in *Clinical Neurophysiology*, 132(2): 632-642, available online from 19 November 2020

Alyssa Hillary Zisk [a], Seyyed Bahram Borgheai [b], John McLinden [b], Sarah M. Hosni [b], Roohollah Jafari Deligani [b], Yalda Shahriari [a,b]

[a] University of Rhode Island, Interdisciplinary Neuroscience Program, Kingston, RI, USA

[b] University of Rhode Island, Electrical, Computer, and Biomedical Engineering Department, Kingston, RI, USA

*Corresponding author: Yalda Shahriari, University of Rhode Island, Computer, Electrical, and Biomedical Engineering Department & Interdisciplinary Neuroscience Program, Kingston, RI, 02881. Email: yalda_shahriari@uri.edu

As first author, Zisk analyzed the data, intepreteted the results, and primarily wrote the text. Dr. Shahriari is the corresponding author and supervised all the aspects of this project, data recording, data analysis, the interpretation of the results, and the manuscript preparation. Dr. Borgheai and John McLinden conducted participant recruitment and data recording. All authors were involved in proofreading and revising the manuscript.

ABSTRACT

Objective: People with amyotrophic lateral sclerosis (ALS) can benefit from brain-computer interfaces (BCIs). However, users with ALS may experience significant variations in BCI performance and event-related potential (ERP) characteristics. This study investigated latency jitter and its correlates in ALS.

Methods: Electroencephalographic (EEG) responses were recorded from six people with ALS and nine neurotypical controls. ERP amplitudes and latencies were extracted. Classifier-based latency estimation was used to calculate latency jitter. ERP components and latency jitter were compared between groups using Wilcoxon rank-sum tests. Correlations between latency jitter and each of the clinical measures, ERP features, and performance measures were investigated using Spearman and repeated measures correlations.

Results: Latency jitter was significantly increased in participants with ALS and significantly negatively correlated with BCI performance in both ALS and control participants. ERP amplitudes were significantly attenuated in ALS, and significant correlations between ERP features and latency jitter were observed. There was no significant correlation between latency jitter and clinical measures.

Conclusions: Latency jitter is increased in ALS and correlates with both BCI performance and ERP features.

Significance: These results highlight the associations of latency jitter with BCI performance and ERP characteristics and could inform future BCI designs for people with ALS.

Highlights:

- People with Amyotrophic Lateral Sclerosis (ALS) have increased P300 latency jitter.
- Latency jitter correlates with BCI performance in both people with ALS and controls.
- Latency jitter does not correlate with measures of disability in ALS.

Keywords: Brain-computer interface (BCI), Event-related potentials (ERP), Amyotrophic lateral sclerosis (ALS), P300 Latency Jitter, Electroencephalography (EEG)

2.1. INTRODUCTION

The P300 response, a positive deflection which occurs approximately 300 ms after an unusual but attended event, has been widely used to control brain-computer interface (BCI) systems since its introduction several decades ago (Farwell and Donchin, 1988). However, the timing of this response is affected by a variety of factors, including age (Polich and Kok, 1995), task specifications (e.g., timing and difficulty) (McFarland et al., 2011, Verleger et al., 2014), and neurological conditions (Raggi et al., 2010, McCane et al., 2015, Kellmeyer et al., 2018). The factors that affect latency can also degrade BCI performance, and thus make these systems inefficient or impractical, particularly for end-users (Birbaumer and Cohen, 2007, Kübler and Birbaumer, 2008, Murguialday et al., 2011, Kellmeyer et al., 2018).

Among BCI users, those with severe motor impairments including amyotrophic lateral sclerosis (ALS) most require these systems due to their loss of muscle control

affecting their communication and environmental control abilities. However, people with ALS are known to exhibit trial-by-trial latency variability in their P300 responses, or latency jitter (Thompson et al., 2019). Preliminary comparisons between single participants suggest jitter may be increased in ALS compared to neurotypical users (Mowla et al., 2017). Growing research also shows that other neurological conditions including attention deficit hyperactivity disorder (Saville et al., 2015), schizophrenia (Ford et al., 1994), traumatic brain injuries (Unsal and Segalowitz, 1995), and disorders of consciousness (Schettini et al., 2015) affect P300 latencies and cause jitter, as does normal aging (Fjell et al., 2009).

The importance of latency jitter in cognitive studies has been well established, and its relevance to BCIs is of growing interest. Latency jitter is significantly correlated with classification accuracy in neurotypical participants for simple visual oddball tasks, typical row-column P300 BCIs, and an alternative P300 speller designed for more effective use without eye movements (Aricò et al., 2014). This correlation also holds in a mixed group of neurotypical participants, participants with ALS, and one participant with muscular dystrophy using a typical P300 speller (Thompson et al., 2012). Jitter has also been reported to increase overall when relying on covert attention in P300 speller use, without eye movement, as compared to overt attention, with this increased jitter contributing to reduced BCI accuracy when relying on covert attention (Aricò et al., 2014).

Given the importance of latency jitter in cognitive studies and its growing recognition in the context of BCIs, a variety of methods exist to quantify this phenomenon (Smulders et al., 1994, Ouyang et al., 2017). For example, Woody

proposed an adaptive filter for estimating latencies and realigning peaks through iterative correlations of time-shifted single-trial responses with the averaged response (Woody, 1967). Kutas and colleagues applied this Woody filter to the P300 response to provide evidence that P300 latencies are correlated with stimulus evaluation time (Kutas et al., 1977). While a study by Verleger and colleagues (Verleger et al., 2014) found that increased decision complexity led to reduced P300 amplitudes, Yu used principal component analysis (PCA) to extract spatial patterns for use with the Woody filter and reported that this apparent reduction in amplitude is instead due to an increase in jitter (Yu, 2016). Other groups utilized filtering and peak-picking methods (Ouyang et al., 2017), ranging from simple but harsh lowpass filters (Smulders et al., 1994, Jaśkowski and Verleger, 2000) to more complex methods including wavelet-based filtering (Aricò et al., 2014, Chennu et al., 2009). Aricò and colleagues used wavelet analysis and found that P300 jitter is inversely correlated with BCI accuracy (Aricò et al., 2014). Their further work suggested P300 jitter is increased in people with disorders of consciousness, is negatively correlated with signs of consciousness, and may impede effective BCI use in this population (Schettini et al., 2015).

Thompson and colleagues introduced classifier-based latency estimation as an additional way to quantify P300 jitter and thereby predict BCI performance (Thompson et al., 2012). In their proposed classifier-based latency estimation method, a classifier is applied to epochs starting at multiple time points surrounding target stimulus presentation, and the time shift corresponding to the highest classifier score is selected as the latency shift for the corresponding epoch. The authors suggested their proposed method can be utilized with any linear classification algorithm and effectively predicts

BCI performance (Thompson et al., 2012). They additionally used their classifier-based latency estimation method on simulated data to evaluate the accuracy of their proposed technique on data with known jitter and further confirmed the presence of P300 jitter in data from people with ALS (Thompson et al., 2019).

While people with ALS are a representative target population for BCI use, they have often been reported to show reduced BCI performance in comparison to neurotypical users (Birbaumer et al., 2012, Kübler and Birbaumer, 2008, Kim et al., 2017, McCane et al., 2014, Mugler et al., 2010, Townsend et al., 2010) This reduced performance and the correlation between BCI performance and latency jitter (Aricò et al., 2014, Thompson et al., 2012) together suggest that increased jitter in ALS may be a primary concern in BCI use, and thus, further investigation is warranted.

Given this concern that P300 jitter may be increased in ALS and consequently negatively affect their BCI performance, this study investigated P300 latency jitter in participants with ALS and neurotypical controls in the use of a P300-based BCI. As a continuation of our prior investigations of event-related potential (ERP) correlates of BCI performance in ALS (Shahriari et al., 2019), and to support the detection of correlations which may be population-specific and ensure correspondences are relevant to end-users, we compared P300 jitter between groups and examined relationships between jitter and both BCI performance and ERP features in each group. We additionally investigated potential correlations between latency jitter and clinical measures in participants with ALS.

2.2 METHODS

2.2.1 Participants and Experimental Protocol

A total of fifteen participants were recruited for this study, six with ALS (age 57 ± 15.7 years,1 female) (See table 1), and nine (age 62.7 ± 4.8 years, 5 female) neurotypical control (NTC) participants age-matched to our elderly participants with ALS with no neurological conditions for comparison. All participants had at least some post-secondary education. All neurotypical control participants had normal or corrected to normal vision, as did participants with ALS other than ALS-1, who was in the late stages of locked-in syndrome with significant ocular impairments. Participants with ALS had an average functional rating scale-revised (ALSFRS-R) score of 11.6 ± 9.5, with a minimum score of 0 indicating no voluntary motor functions and complete dependence on life-sustaining technologies including mechanical ventilation and a maximum score of 48 indicating normal functioning (Cedarbaum et al., 1999). Participants with ALS were diagnosed 6.5 ± 4.0 years prior to the study. Three participants had gastrostomies as well as tracheostomies. ALS-1′s sole form of communication was an idiosyncratic and error-prone yes/no pupil dilation his caregiver read subjectively, which deteriorated over the course of the recordings, losing reliability as a means of communication. Two other participants with artificial ventilation (ALS-2 and 4) used eye-tracking devices to communicate. ALS-3 could still move his index finger and make non-verbal sounds to sustain minimal communication. ALS-5 and 6 retained the ability to speak, though ALS-5 had lost non-facial movement, and ALS-6 could barely move a joystick with one hand. Participants with ALS were tested in their homes or care centers, whereas neurotypical

controls participated at the NeuralPC Lab. The study protocol was approved by the Institutional Review Board (IRB) of the University of Rhode Island (URI), and all participants provided informed consent or assent for the study and received financial compensation.

Each participant took part in three sessions of recording on three different days, except for NTC-5, who only took part in two sessions. Sessions for participants with ALS took place at least two weeks apart. Including preparation such as the application of gel electrodes and tasks, each session typically lasted 2–2.5 hours. As in the conventional P300 speller, a 6x6 matrix of characters containing letters and numbers was displayed to participants, with rows and columns intensified randomly (Farwell and Donchin, 1988). Participants were instructed to attend to the intensification of their target character, with row and column intensified 10 times for each of the 14 target characters in each session. Intensifications consisted of color images of the same face replacing the characters in a row or column (Kaufmann et al., 2011, Kaufmann et al., 2013). Intensifications lasted 93.75 ms, followed by a 62.5 ms inter-stimulus-interval (ISI). Each participant was instructed to mentally count target characters while ignoring non-target ones in the offline (copy-spelling) mode. To familiarize participants with the BCI setup, including the recording protocol and the task, each participant had a familiarization session before the main experimental recordings, in which they completed the same tasks without recording the data and were given the opportunity to get clarification about the tasks.

Participants with ALS additionally took the ALS-Cognitive Behavioral Screen (ALS-CBS), a brief cognitive screen sensitive to frontal dysfunctions for people with

ALS (Woolley et al., 2010). This cognitive testing was completed each session if possible, and average scores were reported as percentages to compensate for the fact that not all items could always be used. Because several participants with ALS had difficulty speaking or writing, the information and retrieval (fluency) section of the ALS-CBS test could not be used effectively. Consequently, only the attention, concentration, and tracking portions of the ALS-CBS test were performed. Due to their disabilities, four participants with ALS required accommodation to complete these portions of the assessment. ALS-1 completed the test once, using a P300 speller. ALS-2 used his typical eye-tracking system. ALS-3 used a printed letter board, pointing with a finger. ALS-4 initially used a letter board, but later used a Tobii eye-tracking system. ALS-5 and ALS-6 did not require accommodations to complete the ALS-CBS. Participants with ALS scored 92.1 ± 6.8% on this test, with attention subscores of 82.8 ± 18.7%.

2.2.2 Data Acquisition

Electroencephalography (EEG) data were recorded using a g.USBamp amplifier (g.tec Medical Technologies) with a 256 Hz sampling rate. Data was recorded from eight channels commonly used in P300 protocols, Fz*, Cz, P3, Pz, P4, PO7, PO8, and Oz (Krusienski et al., 2008). However, as Fz was occupied by sensors for other studies recorded in the same session as the current experiment, it was replaced by the nearest available channel, FAF2, denoted as Fz*. All experimental protocols, data acquisition, and stimulus presentation were controlled using BCI2000 software (Schalk and Mellinger, 2010).

Table 2.1. Demographic information for participants with amyotrophic lateral sclerosis (ALS).

Subject Number	*Age*	*Sex*	*Time since diagnosis (years)*	*Revised ALS Functional Rating Scale ALSFRS-R (out of 48)*	*ALSFRS-R Bulbar Subscore*	*Average ALS Cognitive Behavioral Screen (ALS-CBS) Score (%)*	*Average Attention Subscore (%)*	*Artificial Ventilation*	*Means of Communication*
ALS-01	29	M	4	0	0	100.0	100.0	Yes	No reliable means
ALS-02	55	M	11	4	0	93.3	93.3	Yes	Eye-tracking
ALS-03	70	M	8	14	5	95.2	83.3	No	Non-verbal sound
ALS-04	67	M	2	7	5	94.4	100.0	Yes	Eye-tracking
ALS-05	69	F	11	23	11	80.0	56.7	No	Verbal
ALS-06	52	M	3	22	12	89.6	63.3	No	Verbal
Mean± SD	**57.0±15.7**	-	**6.5±4.0**	**11.6±9.5**	**5.5±5.2**	**92.1±6.8**	**82.8±18.7**	-	-

2.2.3 Data Processing

All data processing was conducted in MATLAB, 2019a. EEG data were detrended and bandpass filtered at 0.5–30 Hz with a Hamming window-based zero-phase filter using the MATLAB functions *fir1* and *filtfilt*. For feature extraction and classifier training, the data were segmented into epochs of 0 to 800 milliseconds post-stimulus. These epochs were averaged over all target segments within each session. For further statistical and correlation analyses, the amplitudes and latencies corresponding to four primary ERP components (N100, P200, N200, and P300) were extracted for each channel and participant. The N100 and N200 components were respectively defined as

the minimum peaks occurring in the 80–170 ms and 220–350 ms periods, and the P200 and P300 components were respectively defined as the maximum peaks in the 190–300 ms and 300–500 ms periods.

Stepwise linear discriminant analysis (SWLDA) classifiers were used to evaluate BCI performance (Krusienski et al., 2008). All epochs (0–800 ms) were downsampled by a factor of 13 through a moving average procedure, and the downsampled data from all channels combined were treated as potential predictors of whether an epoch was a target or non-target epoch. 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 variables ($p > 0.15$) were removed. This procedure was repeated for up to 60 steps, or until no additional terms satisfied the entry/removal criteria (Krusienski et al., 2008). Data from each session were divided into five segments of approximately equal length. Four segments were used for training, and the remaining segment was used for testing. This procedure was repeated five times to test all the data. Flash-by-flash binary classification performance metrics were calculated on each test set, with average performances extracted from each session for use in correlation analysis and from each participant for between-group comparisons. In particular, binary flash accuracy, precision, recall, F-score, and character selection accuracy were calculated as measures of performance (Pal and Bandyopadhyay, 2016, Tang et al., 2017). With TP, TN, FP, and FN respectively representing the number of epochs that were classified as true positives (correct targets), true negatives (correct non-targets), false positives, and false negatives, we computed accuracy, precision, recall, and F-score as below:

$$\text{Flash Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F} - \text{score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

Character selection accuracies were also calculated in which in each of the five data segments per session, the number of correct characters was determined, and the session accuracy was the average character selection accuracy over all five segments.

Following Thompson and colleagues' (Thompson et al., 2012) work, classifier-based latency estimation (CBLE) was used to measure latency jitter. To do so, the classifier was applied to shifted epochs extracted as follows: for each stimulus, 800 ms epochs (both target and non-target) beginning at each available time point from 100 ms before the stimulus to 100 ms after the stimulus were extracted as shifted responses in steps of one sample (~4 ms), creating a total of 53 shifted epochs. Each shifted epoch was then downsampled by a factor of 13 using a moving average procedure and fed to a classifier, providing a series of 53 shifted classifier scores across each epoch, one for each time shift used. These series were extracted for all epochs. For each target flash, the time shift corresponding to the highest classifier score (the probability that the flash was a target flash) was extracted as the latency shift for the flash. The variance of these latency shifts within a session denoted as vCBLE, reliably measures latency jitter (Thompson et al., 2019). This jitter was extracted for each session and averaged over participants for both the ALS and neurotypical control groups. We additionally

averaged these classifier score series over the target and non-target epochs per participant (Mowla et al., 2017). We calculated the maximum value and kurtosis of the participant average classifier score series for target epochs, as well as the unshifted, averaged score for non-target epochs.

2.2.4 Correlation and Statistical Analysis

All statistical analyses were conducted in R version 3.6.1 (R Core Team, 2019). Within-group correlations for both participants with ALS and controls between latency jitter and performance measures (character accuracy, flash accuracy, precision, recall, and F-score) were investigated using repeated measures correlations, an analysis of covariance-based regression appropriate for measuring common (overall) intra-individual associations among measures when multiple non-independent observations are available for each participant (Bakdash and Marusich, 2017). Within-group correlations between latency jitter and ERP features (N100, P200, N200, and P300 amplitudes and latencies) were similarly investigated using repeated measures correlations. Both latency jitter (vCBLE) versus performance and latency jitter versus ERP feature correlations were performed within both the ALS and control groups.

Correlations between the average latency jitter for each participant with ALS and their clinical features, specifically time since diagnosis, ALSFRS-R scores (Cedarbaum et al., 1999), bulbar subscores of the ALSFRS-R, ALS-CBS scores (Woolley et al., 2010), and attention subscores of the ALS-CBS (Geronimo et al., 2016) were investigated using Spearman correlation. For comparisons between groups, session measures (ERP amplitudes and latencies, performance measures, and latency jitter)

were averaged within participants. These participant average measures were then compared between participants with ALS and controls using non-parametric Mann-Whitney U tests (Mann and Whitney, 1947), appropriate for non-normal distributions and small, potentially uneven samples (Siegel and Tukey, 1960). The statistics related to participant average classifier score series, specifically the maximum value and kurtosis of the average series for target epochs and the unshifted averaged score for non-target epochs, were also compared using non-parametric Mann-Whitney U tests. In order to account for multiple comparison corrections, the false discovery rate adjusted p-values ($p < 0.05$) were computed and reported (Hochberg and Benjamini, 1990) for both between-group comparisons and within-group correlations.

2.3 RESULTS

Averaged target ERPs for each of the participants with ALS (top) and neurotypical controls (bottom) at channel Cz are illustrated in Figure 2.1 (left). This figure also shows the average shifted classifier scores for each participant over both target (middle) and non-target (left) epochs, with the ALS group plots shown above and the control group's plots below. In both groups, the average classifier scores for target characters had central peaks corresponding to the unshifted epochs (0 ms shift), which decrease at increased time shifts. However, the peaks in the classifier score series were significantly ($p = 0.01$) higher in neurotypical controls (0.78 ± 0.10) than in participants with ALS (0.55 ± 0.18). The kurtosis (peakedness/sharpness) of these classifier score series trended higher in neurotypical controls (1.84 ± 0.15) than in participants with ALS (1.72 ± 0.17), supporting the sharper appearance of these peaks, although the differences were not

significant ($p = 0.18$). For non-target segments, participants with ALS showed significantly ($p = 0.005$) increased classifier scores (0.10 ± 0.04) in comparison to neurotypical controls (0.04 ± 0.03), indicating less confident classification of target and non-target segments. Within-group variation was also slightly higher in the ALS group than in neurotypical controls.

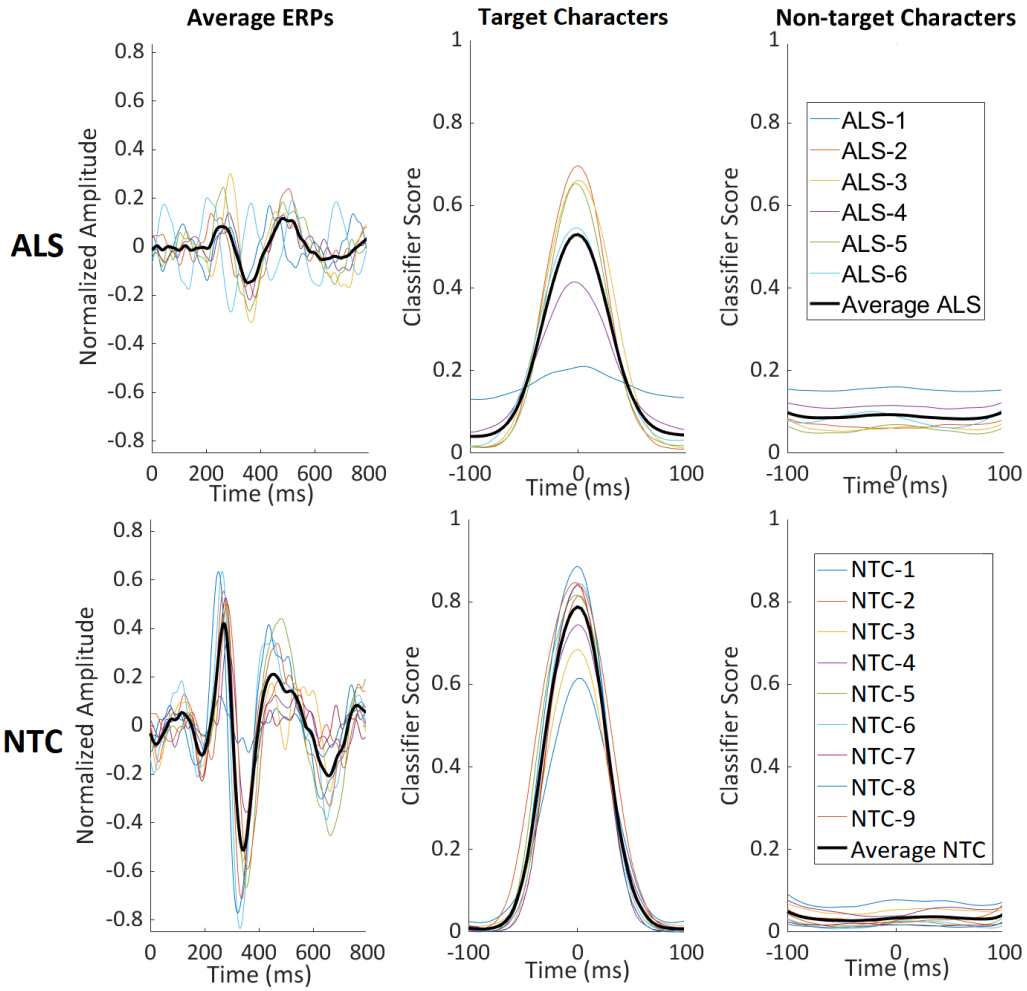

Figure 2.1. Comparison between participants with amyotrophic lateral sclerosis (ALS, top row) and neurotypical controls (NTC, bottom row) of participant averages (colors) and group averages (thick black) for grand average event-related potentials in the 800 ms following stimulus onset (left). Between-group comparisons of classifier scores as a function of time shift for target (middle) and non-target (right) epochs. Each color indicates one participant across all panels of a row.

Figure 2.2 illustrates the differences between the ALS and neurotypical control groups in ERP latencies and amplitudes. While neither N100 amplitude nor latency differed significantly between groups, significant ($p < 0.05$) attenuation was present for other ERP features as follows:

P200 amplitudes were significantly reduced in participants with ALS in the frontal (Fz), central (Cz), parietal (P3, Pz, P4), parieto-occipital (PO8), and occipital (Oz) channels ($p \leq 0.041$), with a maximum average amplitude difference of 0.30 at channel Cz and average amplitudes of 0.20 ± 0.09 and 0.50 ± 0.20 in the ALS and control groups respectively.

N200 amplitudes were significantly ($p < 0.05$) attenuated at channels P4, PO8, and Oz, with a maximum average amplitude difference of 0.42 at channel PO8 and average amplitudes of 0.22 ± 0.12 and -0.64 ± 0.28 in the ALS and control groups respectively.

P300 amplitudes were also reduced in participants with ALS at channel PO8 ($p = 0.038$), with an average amplitude difference of 0.38 and average amplitudes of 0.22 ± 0.12 and 0.50 ± 0.19 in the ALS and control groups respectively.

However, average latencies did not differ significantly between the ALS and control groups for any ERP at any channel.

Figure 2.3 compares participant average latency jitter (vCBLE) and classification performance metrics, specifically character selection accuracy, binary flash classification accuracy, precision, recall, and F-score between groups. All performance measures were significantly decreased in participants with ALS as compared to neurotypical controls ($p < 0.05$). Participants with ALS had a reduced average character selection accuracy of $82.54 \pm 30.18\%$, compared to $99.47 \pm 1.59\%$ for neurotypical

controls. Participants with ALS similarly had 88.91 ± 3.76% of flashes accurately classified as target or non-target, less than the 94.97 ± 2.23% flash accuracy for neurotypical controls. ALS participants also had a precision of 67.95 ± 16.16%, lower than neurotypical controls' precision of 88.31 ± 6.11%. Recall was again lower in the ALS group, at 51.31 ± 25.87%, compared to 80.16 ± 8.64% in neurotypical controls. Given the reductions in precision and recall, participants with ALS had lower F-scores of 0.56 ± 0.26 than neurotypical controls with F-scores of 0.84 ± 0.08. Latency jitter was significantly ($p = 0.01$) increased in participants with ALS, averaging 1350 ± 1073 ms^2 in comparison to 553 ± 224 ms^2 in neurotypical controls. Table 2.2 tabulates the individual results on these measures for each group.

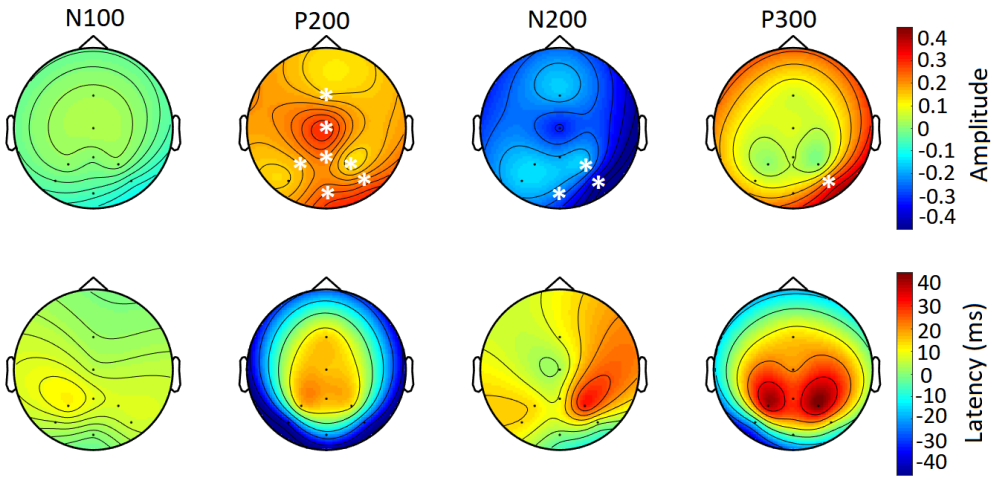

Figure 2.2 The differences in the normalized amplitude (top) and latency (bottom) between participants with amyotrophic lateral sclerosis (ALS) and controls for the N100, P200, N200, and P300 components. Shades of red and orange colors indicate more positive normalized voltages and earlier latencies in controls than in participants with ALS, whereas shades of blue indicate more negative normalized voltages and later latencies. Channels with significant differences between groups after correction for multiple comparisons ($p < 0.05$) are marked with a white asterisk.

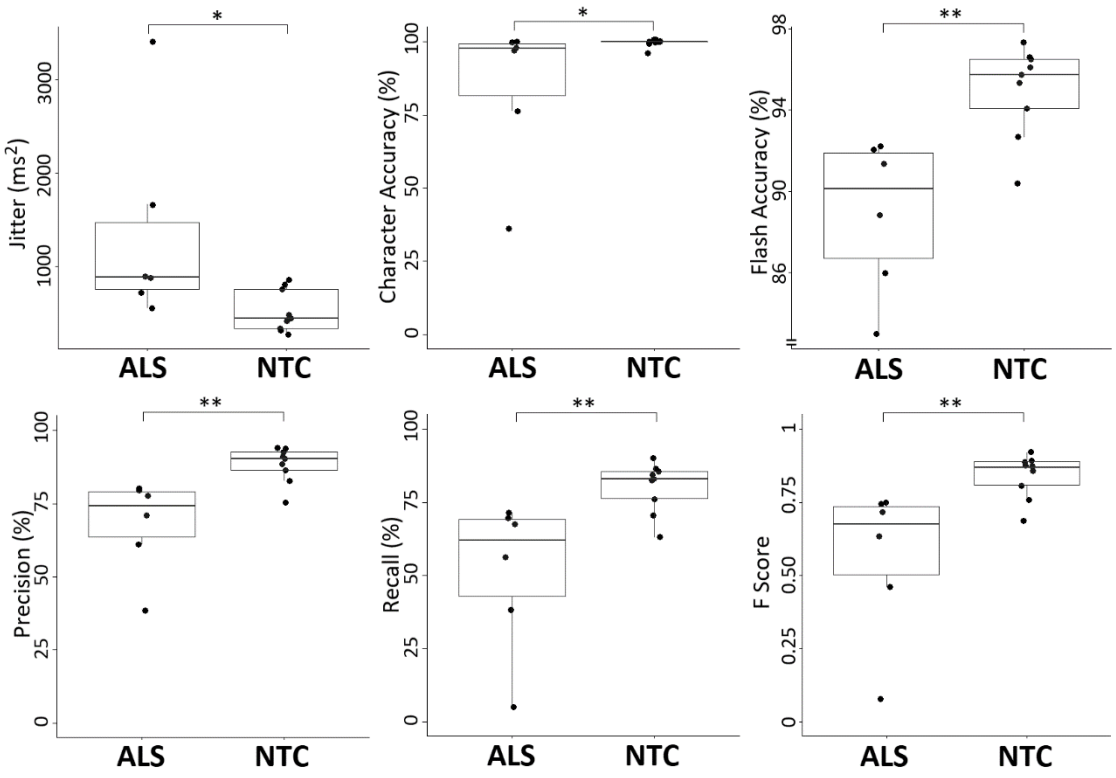

Figure 2.3. Box plots showing latency jitter (top left), character selection accuracy (top center), binary flash classification accuracy (top right), precision (bottom left), recall (bottom center), and F-score (bottom right) for all participants in both the amyotrophic lateral sclerosis (ALS) and neurotypical control (NTC) groups. The boxes show the quartiles with the median represented by a bold line through the box. Each dot shows the corresponding value for one participant (* significant at $p < 0.05$, ** significant at $p < 0.01$, Wilcoxon rank-sum test).

Figure 2.4 shows single-session single-trial classifier score series for both target and non-target segments, along with histograms of the classifier-based latency estimates from a representative participant with ALS and a representative control participant. The participant with ALS had a generally wider, lower peak in their average classifier score series for targets than the control participant, with greater variation in their individual score series apparent in the figure and greater variation in the timing of their maximum classifier score reflected in the histogram. The participant with ALS also had more

apparent peaks in their classifier scores series for non-target epochs than the control participant, leading to more misclassifications of non-target segments.

Table 2.2. Jitter, measured by the variance of classifier-based latency estimates (vCBLE), and performance metrics for all participants in both the amyotrophic lateral sclerosis (ALS) and neurotypical control (NTC) groups. Means and standard deviations (STD) are provided for each group.

		Performance Metrics					
	vCBLE (ms^2)	***Character Accuracy (%)***	***Flash Accuracy (%)***	***Precision (%)***	***Recall (%)***	***F-score***
ALS	ALS-1	3397	23.81	82.96	38.36	4.88	0.08
	ALS-2	553	97.62	92.20	79.45	71.31	0.75
	ALS-3	892	100.00	92.04	80.13	69.52	0.74
	ALS-4	1663	76.19	85.99	61.15	38.21	0.46
	ALS-5	714	97.62	91.39	77.67	67.50	0.72
	ALS-6	880	100.00	88.87	70.90	56.43	0.63
	ALS Mean±STD	**1350± 1073**	**82.54± 30.18**	**88.91± 3.76**	**67.95± 16.16**	**51.31± 25.87**	**0.56± 0.26**
NTC	NTC-1	759	100.00	90.40	75.32	62.98	0.69
	NTC-2	419	100.00	95.34	88.44	82.86	0.86
	NTC-3	807	100.00	92.66	82.64	70.60	0.76
	NTC-4	858	95.24	94.09	86.42	76.19	0.81
	NTC-5	449	100.00	95.74	91.11	82.50	0.87
	NTC-6	340	100.00	96.07	90.43	85.48	0.88
	NTC-7	316	100.00	96.51	93.97	84.40	0.89
	NTC-8	280	100.00	97.36	93.90	90.00	0.92
	NTC-9	478	100.00	96.59	92.61	86.43	0.89
	NTC Mean±STD	**553± 224**	**99.47± 1.59**	**94.97± 2.23**	**88.31± 6.11**	**80.16± 8.64**	**0.84± 0.08**

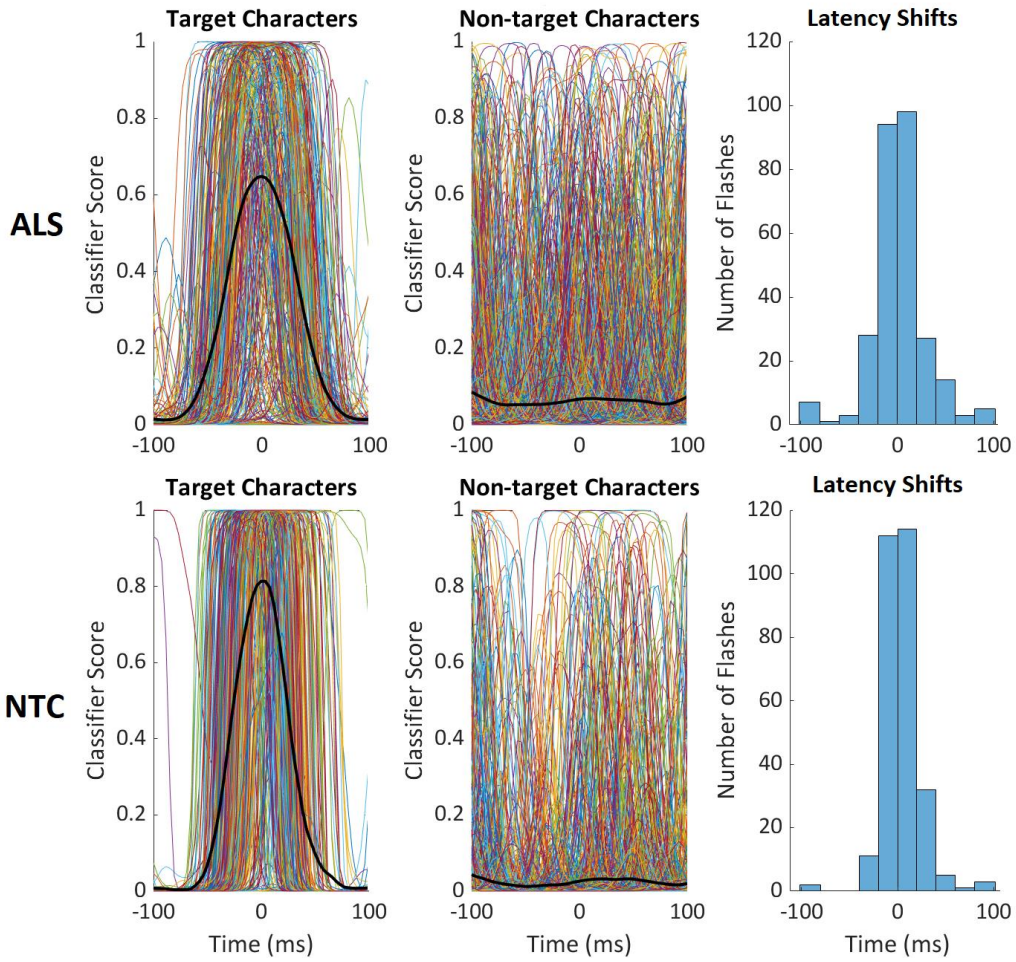

Figure 2.4 Comparison between a representative session from a representative participant with amyotrophic lateral sclerosis (ALS, top row) and a neurotypical control participant (NTC, bottom row). In each row, the classifier scores are shown as a function of time shifts for target epochs (left) and non-target epochs (middle). A histogram of classifier-based latency estimates for target epochs are shown on the right. In the left and middle panels, colored lines represent single epochs, and the bold black line represents the average classifier score series for the session.

Figure 2.5 shows the repeated measures correlations and *p*-values between latency jitter (vCBLE) and performance metrics, specifically character accuracy, flash accuracy, precision, recall, and F-score, in both the ALS and control groups. As shown, there was a significant correlation between latency jitter and four performance metrics, specifically character accuracy (r_{rm} = -0.94, $p < 0.00001$), precision (r_{rm} = -0.78,

p = 0.002), recall (r_{rm} = -0.77, p = 0.002), and F-score (r_{rm} = -0.83, p < 0.001), in participants with ALS. However, no significant correlation was observed between latency jitter and flash classification accuracy in this cohort. In contrast, the control group demonstrated a significant negative (r_{rm} < -0.85, p < 0.00001) correlation between latency jitter and all four per-flash performance metrics (binary flash classification accuracy, precision, recall, and F-score), as well as character accuracy (r_{rm} = -0.82, p < 0.0001).

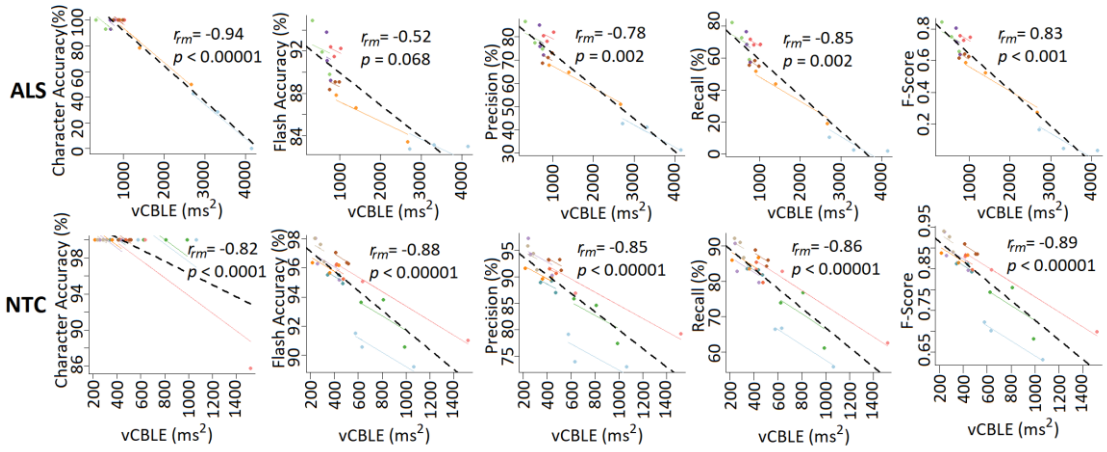

Figure 2.5 Repeated measures correlations between latency jitter (vCBLE, ms^2) and character accuracy (first column), flash accuracy (second column), precision (third column), recall (fourth column), and F-score (fifth column) in participants with amyotrophic lateral sclerosis (ALS, top row) and controls (NTC, bottom row). Each color indicates one participant, and black dashed lines show the overall trends.

Repeated measures correlations between session average ERP features and latency jitter, as measured by vCBLE, are shown in Figure 2.6. Correlations between session average N100 latencies and vCBLE were significant at parieto-occipital channels including P3 (r_{rm} = 0.69, p = 0.03), PO7 (r_{rm} = 0.68, p = 0.03), and Oz (r_{rm} = 0.80, p = 0.009) in participants with ALS. However, there was no significant correlation between N100 amplitudes and latency jitter in participants with ALS. In

contrast, neurotypical controls had significant repeated measures correlations between N100 amplitude and vCBLE at Pz (r_{rm} = 0.63, p = 0.04), P4 (r_{rm} = 0.59, p = 0.04), and PO7 (r_{rm} = 0.56, p = 0.04), but there were no significant correlations between N100 latencies and latency jitter in neurotypical control participants.

Neither P200 latencies nor amplitudes significantly correlated with latency jitter in participants with ALS. However, while P200 latencies also did not correlate with latency jitter in controls, P200 amplitudes correlated with latency jitter in controls at channels Fz (r_{rm} = -0.70, p = 0.005), Cz (r_{rm} = -0.77, p = 0.002), P3 (r_{rm} = -0.63, p = 0.02), Pz (r_{rm} = -0.60, p = 0.02), and Oz (r_{rm} = -0.55, p = 0.03). Similarly, neither N200 latencies nor amplitudes significantly correlated with latency jitter in participants with ALS. However, while N200 latencies also did not correlate with latency jitter in controls, N200 amplitudes correlated with latency jitter in controls at channels Cz (r_{rm} = 0.57, p = 0.04), PO7 (r_{rm} = 0.61, p = 0.03), and PO8 (r_{rm} = 0.76, p = 0.002).

Neither P300 latencies nor amplitudes significantly correlated with latency jitter in participants with ALS or in neurotypical controls. Spearman correlations between latency jitter, measured by vCBLE, and clinical measures, specifically ALSFRS-R scores, ALS-CBS scores, and attention subscores of the ALS-CBS in participants with ALS were not significant (p > 0.05).

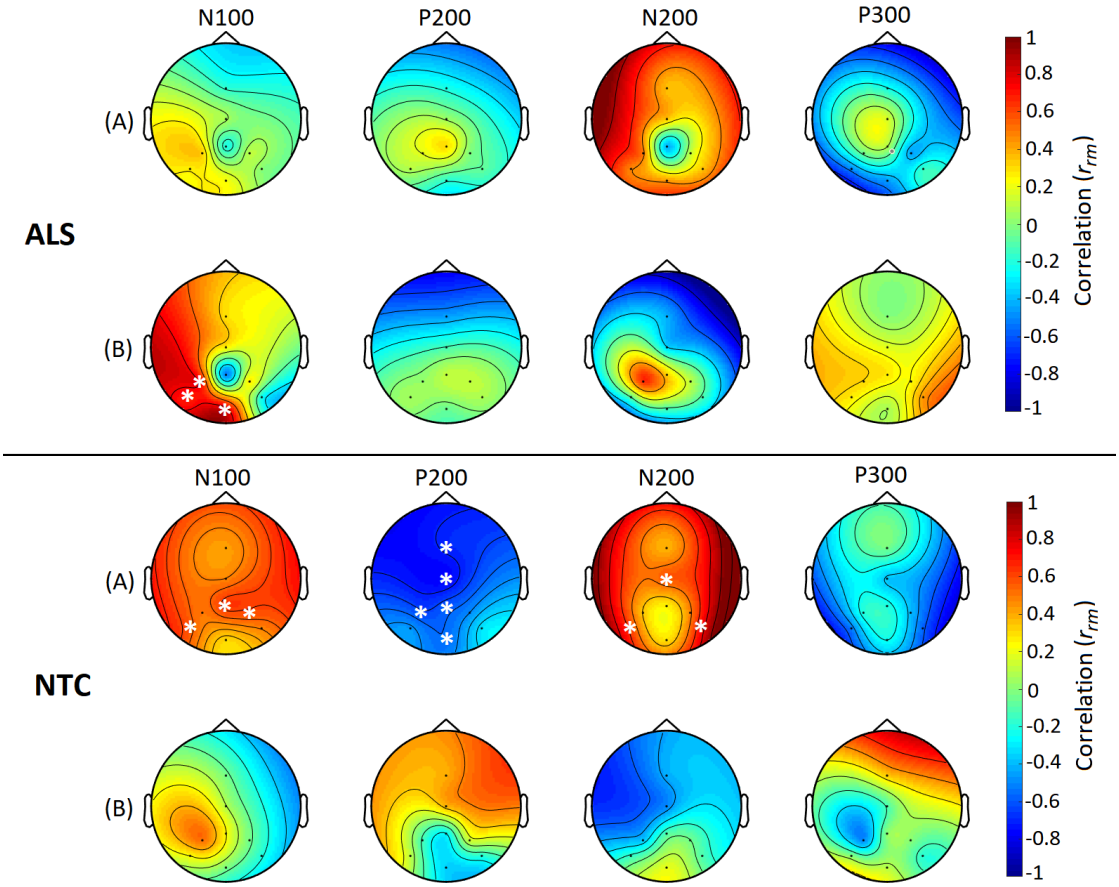

Figure 2.6. Repeated measures correlations between N100, P200, N200, and P300 (A) amplitudes and (B) latencies and latency jitter (vCBLE) in participants with amyotrophic lateral sclerosis (ALS, above) and controls (NTC, below). Shades of red and orange indicate positive correlations between ERP amplitudes or latencies and latency jitter, whereas shades of blue indicate negative correlations. White asterisks indicate channels where this correlation is significant.

2.4 DISCUSSION

This study investigated the correlates of P300 latency jitter in people with ALS and controls. Our study found that latency jitter was increased in participants with ALS, and increased jitter correlated with reduced BCI performance in both the ALS and control groups. Furthermore, we observed that latency jitter correlates significantly with N100, P200, and N200 features in both groups as well. Specifically, increased N100 latencies

were associated with increased latency jitter in ALS, and decreased N100, P200, and N200 amplitudes correlated with increased latency jitter in controls. However, no association was observed between clinical measures of ALS disease and latency jitter. This observation aligns with the findings in a study conducted by McCane and colleagues (McCane et al., 2015), which reported no significant correlations between clinical features and ERP components or BCI performance.

Our study's finding that latency jitter negatively correlates with performance measures in both participants with ALS and controls supports the findings of correlations in neurotypical participants (Aricò et al., 2014) and in mixed groups of potential end-users and controls (Thompson et al., 2012). Our study additionally found that participants with ALS had both significantly reduced BCI performance and increased latency jitter compared to controls. Given the connection between latency jitter and BCI performance and the increase in jitter in people with ALS, potential brain-computer interface users, our study suggests the importance of latency jitter in BCI design.

Generally, visual-based BCIs are not practical for people with visual impairments, as can occur in the later stages of ALS. In our study, the participant in the late locked-in state (ALS-1), who had lost eye-gaze control, had the highest latency jitter and the lowest BCI performance. In combination with increases in latency jitter in paradigms where participants are not permitted to utilize eye-gaze control (Aricò et al., 2014), this suggests latency jitter may be a contributing factor in reduced BCI performance in the absence of fine gaze-control.

To compensate for the negative effect of jitter in BCI performance in ALS, further work could investigate additional strategies to reduce BCI susceptibility to jitter. For example, paradigms that eliminate dependence on visual aspects, including those relying on auditory stimuli (Schettini et al., 2015) and the visuomental paradigm (Borgheai et al., 2019) can be further explored. Current attempts to compensate for latency jitter include the use of classifier score series (depicted in Fig. 4) as features in a secondary classifier (Mowla et al., 2017). However, no real-time implementations of this method have been reported (Mowla et al., 2017).

This study additionally revealed associations between ERP features, including attention-related features such as the N100, P200, and N200, and latency jitter in both the ALS and neurotypical control groups. Participants with ALS showed attenuated P200 and N200 amplitudes at several channels as well as increased jitter overall but did not show significant correlations between ERP amplitudes and latency jitter for any component or channel. Neurotypical participants, however, showed significant correlations between N100, P200, and N200 amplitudes and latency jitter. Previous results support associations between increased latency jitter and reduced average ERP amplitudes. Both increased jitter and attenuated ERPs are present in various conditions including schizophrenia (Ford et al., 1994), traumatic brain injuries (Unsal and Segalowitz, 1995), and normal aging (Fjell et al., 2009).

The associations between ERP amplitudes and latency jitter we found in neurotypical controls align with established results. Specifically, the amplitude of the attention-related N200 component (Balconi and Canavesio, 2016, O'Brien et al., 2013), which is implicated in P300-based BCI performance (Halder et al., 2013, Mak et al.,

2012, Riccio et al., 2018), was significantly correlated with jitter in neurotypical control participants and attenuated in participants with ALS in our current study. Increased P200 amplitude is similarly associated with successful BCI sessions during longitudinal use (Shahriari et al., 2019) and with decreased latency jitter in the current study, supporting the importance of this component along with the P300. The P200, which relates to higher-order perceptual processing modulated by attention (Lijffijt et al., 2009), was also decreased overall in participants with ALS in the current study, aligning with the relevance of the P200 to BCI performance in our prior work (Shahriari et al., 2019) and other studies (Halder et al., 2013). Like the P200, the N100 has been reported to be associated with attention-modulated perceptual processing (Lijffijt et al., 2009). In our study, N100 latency correlated with latency jitter in participants with ALS, whereas its amplitude was correlated with jitter in neurotypical controls.

The P300 mediates BCI performance (Halder et al., 2013, Mak et al., 2012, Riccio et al., 2018) and can itself be mediated by attention. For example, P300 jitter is increased in neurotypical participants in BCI tasks that rely on covert attention rather than overt attention (Aricò et al., 2014). Several studies have reported attentional dysfunctions in participants with ALS (Volpato et al., 2016, Riccio et al., 2013), and other groups reported deflections in attention-related ERP components in these cohorts (Raggi et al., 2010, Vieregge et al., 1999). Thus, given the correlations between attention-related ERP features and latency jitter observed in our study, we speculate that latency jitter may also relate to attentional dysfunctions in ALS.

Overall, this study explored latency variability in the use of a typical P300-based BCI, finding that jitter is increased in participants with ALS and correlates with

performance variations within this population. These results could lead to improved BCI performance by suggesting latency jitter as a critical factor in the development of BCIs, which predict and adapt to performance variations. Furthermore, our findings enhance our understanding of the mechanisms underlying ALS which can enrich future diagnostic and prognostic techniques. Further methods to compensate for excessive jitter, possibly involving classifier score series (Mowla et al., 2017) or adaptive filters to better detect single-trial responses and correct for latency jitter (Woody, 1967), and thereby address jitter related reductions in BCI performance are worth investigating to improve BCI performance in real-time.

One limitation of the study is that vCBLE, the measure of jitter used in our study, has the theoretical limitation that it covers shifts of the entire 800 ms window and not just the P300 component (Thompson et al., 2012). However, tests with simulated data show its efficacy in reflecting P300 latency jitter (Thompson et al., 2019). Another limitation of the current study is the relatively small sample size. The use of repeated measures correlations, rather than correlations on participant averages, increases power while maintaining statistical rigor (Bakdash and Marusich, 2017). While the consistency of significant findings across all the subjects within each group supports more deterministic results, further studies with additional participants would be of value in determining if these findings can be generalized and strengthen the power of our analysis. This limitation leads to a remaining question that is whether the lack of a significant correlation between ERP amplitudes and latency jitter in participants with ALS is due to our small sample size or due to certain pathophysiological aspects of ALS. However, as latency jitter does not fully explain reductions in grand average ERP

amplitudes in other populations which experience both effects (Ford et al., 1994, Saville et al., 2015, Unsal and Segalowitz, 1995, Walhovd et al., 2008), increased jitter similarly may not fully explain grand average amplitude reductions in ALS. A final limitation of this study was our limited number of sessions resulting in a lack of longitudinal explorations of latency jitter and BCI performance variability in these cohorts. Intra-individual BCI performance variations in ALS have been previously reported by our group and others (Nijboer et al., 2010, Shahriari et al., 2019), and thus, exploring the longitudinal associations of latency jitter with BCI performance variations can support a better understanding of their respective mechanisms over time and suggest adaptive strategies to overcome this issue.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was supported by the Institutional Development Award (IdeA) Network for Biomedical Research Excellence from the NIGMS of the NIH (P20GM103430) and National Center for Adaptive Neurotechnologies (EB018783). The authors would like to thank the participants 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 for their continuous support.

REFERENCES

Aricò P, Aloise F, Schettini F, Salinari S, Mattia D, Cincotti F. Influence of P300 latency jitter on event related potential-based brain–computer interface performance. *Journal of Neural Engineering*. 2014; 11:035008.

Bakdash JZ, Marusich LR. Repeated measures correlation. *Frontiers in Psychology*. 2017; 8:456.

Balconi M, Canavesio Y. Is empathy necessary to comprehend the emotional faces? The empathic effect on attentional mechanisms (eye movements), cortical correlates (N200 event-related potentials) and facial behaviour (electromyography) in face processing. *Cognition and Emotion* 2016; 30:210-224.

Birbaumer N, Cohen LG. Brain–computer interfaces: communication and restoration of movement in paralysis. *The Journal of Physiology*. (Lond.) 2007;579:621-636.

Birbaumer N, Piccione F, Silvoni S, Wildgruber M. Ideomotor silence: the case of complete paralysis and brain–computer interfaces (BCI). *Psychological Research*. 2012; 76:183-191.

Borgheai SB, Deligani RJ, McLinden J, Zisk AH, Hosni SI, Abtahi M, et al. Multimodal exploration of non-motor neural functions in ALS patients using simultaneous EEG-fNIRS recording. *Journal of Neural Engineering*. 2019.

Cedarbaum JM, Stambler N, Malta E, Fuller C, Hilt D, Thurmond B, et al. The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). *Journal of the Neurological Sciences*. 1999; 169:13-21.

Chennu, S., Craston, P., Wyble, B., & Bowman, H. (2009). Attention increases the temporal precision of conscious perception: Verifying the neural-ST2 model. *PLoS Computational Biology*, 5(11), e1000576.

Farwell LA, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*. 1988; 70:510-523.

Fjell AM, Rosquist H, Walhovd KB. Instability in the latency of P3a/P3b brain potentials and cognitive function in aging. *Neurobiology of Aging* 2009; 30:2065-2079.

Ford JM, White P, Lim KO, Pfefferbaum A. Schizophrenics have fewer and smaller P300s: a single-trial analysis. *Biological Psychiatry* 1994; 35:96-103.

Geronimo A, Simmons Z, Schiff SJ. Performance predictors of brain–computer interfaces in patients with amyotrophic lateral sclerosis. *Journal of Neural Engineering*. 2016; 13:026002.

Halder S, Ruf CA, Furdea A, Pasqualotto E, De Massari D, van der Heiden L, et al. Prediction of P300 BCI aptitude in severe motor impairment. *PloS one* 2013; 8: e76148.

Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. *Statistics in Medicine*. 1990; 9: 811-818.

Jaśkowski P, Verleger R. An evaluation of methods for single-trial estimation of P3 latency. *Psychophysiology* 2000; 37:153-162.

Kaufmann T, Schulz SM, Grünzinger C, Kübler A. Flashing characters with famous faces improves ERP-based brain–computer interface performance. *Journal of Neural Engineering*. 2011; 8:056016.

Kaufmann T, Schulz SM, Köblitz A, Renner G, Wessig C, Kübler A. Face stimuli effectively prevent brain–computer interface inefficiency in patients with neurodegenerative disease. *Clinical Neurophysiology*. 2013; 124:893-900.

Kellmeyer P, Grosse-Wentrup M, Schulze-Bonhage A, Ziemann U, Ball T. Electrophysiological correlates of neurodegeneration in motor and non-motor brain regions in amyotrophic lateral sclerosis-implications for brain-computer interfacing. *Journal of Neural Engineering* 2018; 15:041003.

Kim HY, Seo K, Jeon HJ, Lee U, Lee H. Application of Functional Near-Infrared Spectroscopy to the Study of Brain Function in Humans and Animal Models. *Molecules and Cells* 2017; 40:523-532.

Krusienski DJ, Sellers EW, McFarland DJ, Vaughan TM, Wolpaw JR. Toward enhanced P300 speller performance. *Journal of Neuroscience Methods*. 2008; 167:15-21.

Kübler A, Birbaumer N. Brain–computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients? *Clinical Neurophysiology*. 2008; 119:2658-2666.

Kutas M, McCarthy G, Donchin E. Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science* 1977; 197:792-795.

Lijffijt M, Lane SD, Meier SL, Boutros NN, Burroughs S, Steinberg JL, et al. P50, N100, and P200 sensory gating: relationships with behavioral inhibition, attention, and working memory. *Psychophysiology* 2009; 46:1059-1068.

Mak JN, McFarland DJ, Vaughan TM, McCane LM, Tsui PZ, Zeitlin DJ, et al. EEG correlates of P300-based brain–computer interface (BCI) performance in people with amyotrophic lateral sclerosis. *Journal of Neural Engineering*. 2012; 9:026014.

Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics* 1947:50-60.

MATLAB. (2019). 9.6.0.1214997 (R2019a) Update 6. Natick, MA: The MathWorks Inc.

McCane LM, Heckman SM, McFarland DJ, Townsend G, Mak JN, Sellers EW, et al. P300-based brain-computer interface (BCI) event-related potentials (ERPs): People with amyotrophic lateral sclerosis (ALS) vs. age-matched controls. *Clinical Neurophysiology*. 2015; 126:2124-2131.

McCane LM, Sellers EW, Mcfarland DJ, Mak JN, Carmack CS, Zeitlin D, et al. Brain-computer interface (BCI) evaluation in people with amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*. 2014; 15:207-215.

McFarland DJ, Sarnacki WA, Townsend G, Vaughan T, Wolpaw JR. The P300-based brain–computer interface (BCI): effects of stimulus rate. *Clinical Neurophysiology*. 2011; 122:731-737.

Mowla MR, Huggins JE, Thompson DE. Enhancing P300-BCI performance using latency estimation. *Brain Computer Interfaces* (Abingdon) 2017; 4:137-145.

Mugler EM, Ruf CA, Halder S, Bensch M, Kubler A. Design and implementation of a P300-based brain-computer interface for controlling an internet browser. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2010; 18:599-609.

Murguialday AR, Hill J, Bensch M, Martens S, Halder S, Nijboer F, et al. Transition from the locked in to the completely locked-in state: a physiological analysis. *Clinical Neurophysiology*. 2011; 122:925-933.

Nijboer F, Birbaumer N, Kubler A. The influence of psychological state and motivation on brain–computer interface performance in patients with amyotrophic lateral sclerosis–a longitudinal study. *Frontiers in Neuroscience*. 2010; 4:55.

O'Brien JL, Edwards JD, Maxfield ND, Peronto CL, Williams VA, Lister JJ. Cognitive training and selective attention in the aging brain: An electrophysiological study. *Clinical Neurophysiology*. 2013; 124:2198-2208.

Ouyang G, Hildebrandt A, Sommer W, Zhou C. Exploiting the intra-subject latency variability from single-trial event-related potentials in the P3 time range: a review and comparative evaluation of methods. *Neuroscience & Biobehavioral Reviews*. 2017; 75:1-21.

Pal M, Bandyopadhyay S. Many-objective feature selection for motor imagery EEG signals using differential evolution and support vector machine. *International Conference on Microelectronics, Computing and Communications (MicroCom)* 2016:1-6.

Polich J, Kok A. Cognitive and biological determinants of P300: an integrative review. *Biological Psychology*. 1995; 41:103-146.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raggi A, Iannaccone S, Cappa SF. Event-related brain potentials in amyotrophic lateral sclerosis: A review of the international literature. *Amyotrophic Lateral Sclerosis* 2010; 11:16-26.

Riccio A, Schettini F, Simione L, Pizzimenti A, Inghilleri M, Olivetti-Belardinelli M, et al. On the relationship between attention processing and P300-based Brain Computer Interface control in Amyotrophic Lateral Sclerosis. *Frontiers in Human Neuroscience* 2018; 12:165.

Riccio A, Simione L, Schettini F, Pizzimenti A, Inghilleri M, Olivetti Belardinelli M, et al. Attention and P300-based BCI performance in people with amyotrophic lateral sclerosis. *Frontiers in Human Neuroscience*. 2013;7.

Saville CW, Feige B, Kluckert C, Bender S, Biscaldi M, Berger A, et al. Increased reaction time variability in attention-deficit hyperactivity disorder as a response-related phenomenon: evidence from single-trial event-related potentials. *Journal of Child Psychology and Psychiatry* 2015; 56:801-813.

Schalk G, Mellinger J. A practical guide to brain–computer interfacing with BCI2000: General-purpose software for brain-computer interface research, data acquisition, stimulus presentation, and brain monitoring. Springer Science & Business Media, 2010.

Schettini F, Risetti M, Aricò P, Formisano R, Babiloni F, Mattia D, et al. P300 latency Jitter occurrence in patients with disorders of consciousness: Toward a better design for Brain Computer Interface applications. *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* 2015:6178-6181.

Shahriari Y, Vaughan TM, McCane L, Allison BZ, Wolpaw JR, Krusienski DJ. An exploration of BCI performance variations in people with amyotrophic lateral sclerosis using longitudinal EEG data. *Journal of Neural Engineering* 2019.

Siegel S, Tukey JW. A nonparametric sum of ranks procedure for relative spread in unpaired samples. *Journal of the American Statistical Association* 1960; 55:429-445.

Smulders FT, Kenemans JL, Kok A. A comparison of different methods for estimating single-trial P300 latencies. *Electroencephalography and Clinical Neurophysiology*. 1994; 92:107-114.

Tang Z, Li C, Sun S. Single-trial EEG classification of motor imagery using deep convolutional neural networks. *Optik* 2017; 130:11-18.

Thompson DE, Mowla MR, Huggins JE. Evidence of latency variation in the P3 speller brain computer interface. Society for Neuroscience October 23, 2019. https://www.abstractsonline.com/pp8/#!/7883/presentation/69619.

Thompson DE, Warschausky S, Huggins JE. Classifier-based latency estimation: a novel way to estimate and predict BCI accuracy. *Journal of Neural Engineering*. 2012; 10:016006.

Townsend G, LaPallo BK, Boulay CB, Krusienski DJ, Frye GE, Hauser C, et al. A novel P300-based brain–computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clinical Neurophysiology*. 2010; 121:1109-1120.

Unsal A, Segalowitz SJ. Sources of P300 attenuation after head injury: Single-trial amplitude, latency jitter, and EEG power. *Psychophysiology* 1995; 32:249-256.

Verleger R, Baur N, Metzner MF, Śmigasiewicz K. The hard oddball: Effects of difficult response selection on stimulus-related P 3 and on response-related negative potentials. *Psychophysiology* 2014; 51:1089-1100.

Vieregge P, Wauschkuhn B, Heberlein I, Hagenah J, Verleger R. Selective attention is impaired in amyotrophic lateral sclerosis—a study of event-related EEG potentials. *Cognitive Brain Research* 1999; 8:27-35.

Volpato C, Prats Sedano MA, Silvoni S, Segato N, Cavinato M, Merico A, et al. Selective attention impairment in amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*. 2016; 17:236-244.

Walhovd, K. B., Rosquist, H., & Fjell, A. M. (2008). P300 amplitude age reductions are not caused by latency jitter. *Psychophysiology*, 45(4), 545-553.

Wang L, Kuroiwa Y, Li M, Wang J, Kamitani T. Do P1 and N1 evoked by the ERP task reflect primary visual processing in Parkinson's disease? *Documenta Ophthalmologica* 2001; 102:83-93.

Woody CD. Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Medical and biological engineering* 1967; 5:539-554.

Woolley SC, York MK, Moore DH, Strutt AM, Murphy J, Schulz PE, et al. Detecting frontotemporal dysfunction in ALS: utility of the ALS Cognitive Behavioral Screen (ALS-CBS). *Amyotrophic Lateral Sclerosis* 2010; 11:303-311.

Yu X. The Impact of Latency Jitter on the Interpretation of P300 in the Assessment of Cognitive Function. Master's Thesis, University of South Florida. 2016.

CHAPTER 3: IMPROVING LONGITUDINAL P300-BCI PERFORMANCE FOR PEOPLE WITH ALS USING A DATA AUGMENTATION AND JITTER CORRECTION APPROACH

Submitted March 4, 2021 to the *Journal of Neural Engineering*, under review.

Alyssa Hillary Zisk [a], Seyyed Bahram Borgheai [b], John McLinden [b],

Yalda Shahriari [ab]

[a] University of Rhode Island, Interdisciplinary Neuroscience Program, Kingston, RI, USA

[b] University of Rhode Island, Electrical, Computer, and Biomedical Engineering Department, Kingston, RI, USA

As first author, Zisk analyzed the data, intepreteted the results, and primarily wrote the text. Dr. Shahriari is the corresponding author and supervised all the aspects of this project, data recording, data analysis, the interpretation of the results, and the manuscript preparation. Dr. Borgheai and John McLinden conducted participant recruitment and data recording. All authors were involved in proofreading and revising the manuscript.

ABSTRACT

Objective. P300-based brain–computer interface (BCI) systems enable people with neuromuscular disabilities, including amyotrophic lateral sclerosis (ALS), to communicate and control using brain activity. However, variation in the P300 latency, also called latency jitter, is both increased in people with ALS and negatively associated with their performance. In this study, we proposed a classification scheme utilizing data augmentation and jitter correction to improve BCI performance for people with ALS.

Approach. Longitudinal recordings were taken while six people with ALS used a P300-based BCI. Our proposed augmentation and correction (A/C) classification scheme included data augmentation and correction for jitter, both relying on time-shifted responses with individualized parameters determined based on latency jitter. Performance metrics including character selection accuracy and binary accuracy, precision, recall, and F-score were calculated using both the proposed classification scheme and a reference classifier that did not implement data augmentation or correction. Performance was compared between the two classification methods using paired *t*-tests and investigated longitudinally using correlation analyses. Correlations between performance improvements and clinical measures were also investigated.

Main results. The proposed A/C classification scheme significantly improved character selection accuracy, required for usability, as well as recall and F-scores. However, precision was reduced, and binary accuracy was not significantly affected. Overall, BCI performance deteriorated over time with both classification methods.

Selection accuracies were more improved by the proposed A/C approach for participants with more significant physical impairments.

Significance. The proposed classification scheme improved P300 BCI performance for our participants with ALS, showing the effectiveness of both data augmentation and of taking latency jitter into consideration. While our longitudinal analysis showed decreased BCI performance and increased latency jitter over time, our proposed scheme partially mitigated deterioration in some performance metrics. These results should inform further work on improving longitudinal BCI performance and reliability for people with ALS.

Keywords: brain-computer interfaces, amyotrophic lateral sclerosis, data augmentation, latency jitter, jitter correction

3.1 INTRODUCTION

As people with amyotrophic lateral sclerosis (ALS) develop significant motor disability and lose voluntary motor control, they frequently require tools for augmentative and alternative communication. Currently available tools, including brain-computer interfaces (BCIs) support communication for people with ALS, but people with ALS have been reported to show reduced BCI performance in comparison to neurotypical users [1-5]. In addition, BCI users can experience substantial variations in BCI performance within and across days [6-8].

Given these concerns, significant attention has been paid to both understanding the correlates of BCI performance and improving classification accuracy. Much of this

research is dedicated to understanding and improving BCIs based on the visual P300 response, a positive electrical deflection occurring 250-500ms after an attended rare event [9]. In a longitudinal study of P300-based BCI home users, Shahriari and colleagues found that BCI performance was positively associated with P200 amplitude, parietal alpha-band spectral power, and occipital beta-band spectral power, but negatively correlated with occipital delta-band power [8]. Mak and colleagues found that among participants with ALS, increased event-related potential (ERP) amplitudes and theta-band spectral power were associated with increased P300 BCI performance [10]. Geronimo and colleagues found that higher cognitive scores, including scores measuring attention, were associated with both increased P300 quality and BCI performance [5, 11].

Trial to trial variation in P300 latency, known as latency jitter, has been found to be negatively associated with BCI performance in a mixed group of neurotypical participants and potential end-users [12], in neurotypical participants [13, 14], and in people with ALS [3]. For example, Zisk and colleagues recently determined that this latency jitter is elevated in people with ALS as compared to neurotypical controls [3], and latency jitter is a factor affecting BCI performance for people with ALS [3, 12, 15].

As studies have shown latency jitter can predict BCI performance [12, 14], Mowla and colleagues used latency estimation and a secondary classifier to improve BCI performance, though they did not report an online implementation of this method [16]. Togashi and Washizawa similarly utilized Bayesian latency estimation to improve P300 BCI performance [17]. Considering differences in latencies between experimental paradigms which elicited P300 responses rather than variability within participants

using a single paradigm, Iturrate and colleagues calculated the latency shift between paradigms and then trained a classifier for one paradigm using data from another, time-shifted to compensate for the latency differences between the experimental conditions [18]. They found that in cases of insufficient training data from any given paradigm, including latency-corrected training data from other paradigms improved performance.

In recent years, data augmentation for BCIs has gained attention as a strategy for improving performance [19-23]. The purpose of data augmentation is to increase the size of the training data, and thereby improve the reliability and generalizability of the classification algorithms. As electroencephalography (EEG) data varies significantly between different participants, many EEG classifiers are single-subject, though pooling data from multiple participants has also been the focus of some research [24] with a similar goal of improving generalizability and reliability. Iturrate and colleague's collected data from multiple experimental paradigms that produced P300 responses, but with different latencies. Their transfer of data between different experiments that evoke P300 responses similarly works towards the goal of improving generalizability and stability with limited training data [18]. In other studies, the use of time shifted epochs has supported the extraction of multiple segments per stimulus, providing a larger training data set [23, 25, 26]. For example, Kim and colleagues used a -100 ms shift, doubling the size of their data set in a reinforcement learning method and requiring both the shifted and unshifted epoch to be classified correctly for the classification to be considered correct [26]. In their study, they noted an improvement with this data augmentation scheme as compared to using a single time window without augmentation. Sakai and colleagues compared several data augmentation methods, including set time

shifts of ±10 ms for all participants, tripling their training data sets [25]. Their data augmentation protocol improved classification performance, with greater improvements found when the training set was smaller. Krell and colleagues similarly considered several data augmentation methods, including time-shifted data, for augmenting P300 training data [23]. In their study, single time-shifts were tested and provided improvements for some participants, but no single time-shift was reported to be consistently helpful. They then tested symmetrical time-shifts and reported that ±40 ms shifts increased the data set but did not significantly affect performance. In all three studies, unshifted epochs, beginning at the time of the stimulus, were used alongside overlapping time-shifted epochs extracted from the recorded EEG data. These three studies sought to classify responses which can vary in latency, and their use of time-shifted data both increased the number of epochs available for training and provided epochs with earlier and/or later responses of interest [23, 25, 26]. As data augmentation with time-shifted data provides time-shifted responses in the training data, this augmentation approach provides additional latency variability that may improve robustness to this same form of variability [19].

In this study we therefore proposed a correction strategy that relied on latency jitter at multiple levels. In particular, we propose to improve classification performance for P300 data longitudinally recorded from people with ALS using both data augmentation and jitter correction. The data augmentation utilizes time-shifted responses to both target and non-target trials, with individualized time shifts based on latency variations present in the training set. The jitter correction procedure was also implemented through allowing limited time-shifts of the epochs to be classified. We quantify our performance

improvements through the use of a reference classifier using neither data augmentation nor jitter correction. We then investigated longitudinal relationships between clinical measures, latency jitter, and BCI performance in our participants with ALS.

3.2 METHODS

3.2.1 Participants

Six participants with ALS (age 57±15.7 years,1 female) were recruited for this study (see Table 3.1). All participants had at least some post-secondary education. Participants other than ALS-01 had normal or corrected to normal vision, while ALS-01 was in the late stages of locked-in syndrome with significant ocular impairments. Participants were diagnosed with ALS 6.5 ± 4.0 years prior to the start of the study and had an average functional rating scale-revised (ALSFRS-R) score of 11.6 ± 9.5, with a minimum score of 0 indicating no voluntary motor functions and complete dependence on life-sustaining technologies including mechanical ventilation and a maximum score of 48 indicating normal functioning [27]. Three participants had gastrostomies as well as tracheostomies. ALS-01′s sole form of communication was an idiosyncratic and error-prone yes/no pupil dilation his caregiver read subjectively, which deteriorated over the course of the recordings, losing reliability as a means of communication. Two other participants with artificial ventilation (ALS-02 and 04) used eye-tracking devices to communicate. ALS-03 could still move his index finger and make non-verbal sounds to sustain minimal communication. ALS-05 and 06 retained the ability to speak, though ALS-05 had lost non-facial movement, and ALS-06 could barely move a joystick with one hand. Participants were tested in their homes or care centers. The study protocol

was approved by the Institutional Review Board (IRB) of the University of Rhode Island (URI), and all participants provided informed consent or assent for the study and received financial compensation.

Table 3.1. Demographic Information for Participants with amyotrophic lateral sclerosis (ALS)

Subject Number	***Age***	***Sex***	***Time since diagnosis (years)***	***Revised ALS Functional Rating Scale (ALSFRS-R) (out of 48)***	***ALSFRS-R Bulbar Subscore***	***Artificial Ventila-tion***	***Means of Communication***
ALS-01	29	M	4	0	0	Yes	No reliable means
ALS-02	55	M	11	4	0	Yes	Eye-tracking
ALS-03	70	M	8	14	5	No	Non-verbal sound
ALS-04	67	M	2	7	5	Yes	Eye-tracking
ALS-05	69	F	11	23	11	No	Verbal
ALS-06	52	M	3	22	12	No	Verbal
Mean±SD	**57.0±15.7**	**-**	**6.5±4.0**	**11.6±9.5**	**5.5±5.2**	-	-

3.2.2 Experimental Protocol

Each participant took part in 5-12 (9.5 ± 2.6) sessions of recording over 2.5-13.7 (10.9 ± 4.3) months. These sessions took place at least two weeks apart. Including preparation such as the application of gel to electrodes and impedance calibration, each session typically lasted 2-2.5 hours. To familiarize participants with the BCI setup, including the recording protocol and the task, each participant took part in a single familiarization session before the main experimental recordings, in which they

completed the same tasks without recording the data and were given the opportunity to get clarification about the experimental tasks. Each session contained one run of a standard P300 spelling protocol, in which a 6x6 matrix of characters containing letters and numbers was displayed to participants, with each row and column intensified 10 times (i.e., 10 trials) per character selection [28].

3.2.3 Data Acquisition

EEG data were recorded using a g.USBamp amplifier (g.tec Medical Technologies) with a 256 Hz sampling rate. Data were recorded from eight channels commonly used in P300 protocols, Fz*, Cz, P3, Pz, P4, PO7, PO8, and Oz [29]. However, as Fz was occupied by sensors for other studies recorded in the same session as the current experiment, it was replaced by the nearest available channel, FAF2, denoted as Fz*. All experimental protocols, data acquisition, and stimulus presentation were controlled using BCI2000 software [30].

3.2.4 Data Analysis

EEG data were detrended and bandpass filtered at 0.5-30 Hz offline. Then, the data were segmented into 100 ms pre-stimulus to 900 ms post-stimulus epochs. From these 1 s epochs, 800 ms sub-epochs were extracted using a moving window to produce epochs beginning at each available time point from approximately 100 ms pre-stimulus to 100 ms post-stimulus, producing 53 time-shifted 800 ms epochs per stimulus. These 800 ms epochs were subject to a moving average procedure, where each value was

replaced by the local mean calculated over a moving window, and then downsampled by a factor of 13, following the feature reduction procedure used in previous studies [12]. The downsampled epochs from all channels were concatenated and then treated as potential features for classification. All data processing was conducted in MATLAB R2019a.

As shown in Figure 3.1, two stepwise linear discriminant analysis (SWLDA) based classification methods with typical parameters for P300 speller applications were used to characterize performance [29, 31]. In our proposed method, a classification scheme with data augmentation and jitter correction, hereafter referred to as augmentation/correction (A/C) classification, was implemented with data augmentation on the training set and correction for latency jitter applied to the test set, which will be explained in sections 2.4.2 and 2.4.3 respectively. For comparison, reference SWLDA classifiers were trained on the same data with no data augmentation or correction for latency jitter.

To ensure that the approach could be implemented in practical environments, data from prior sessions were used to predict performance and determine correction parameters for future sessions. Beginning with each participant's third session, session performances were evaluated by taking that participant's two prior sessions as the training set. That is, classifiers were trained on data from each participant's first two sessions and then evaluated the data of their third session as its test set; then classifiers were trained on the second and third sessions to evaluate their fourth session, and so forth.

Performance metrics, including the metrics for binary classification as well as the character selection accuracy, and latency variability (i.e. latency jitter), were calculated for both the reference and A/C classification procedures. Longitudinal analysis was then performed using the outputs from both classifiers as explained in detail in section 2.5.

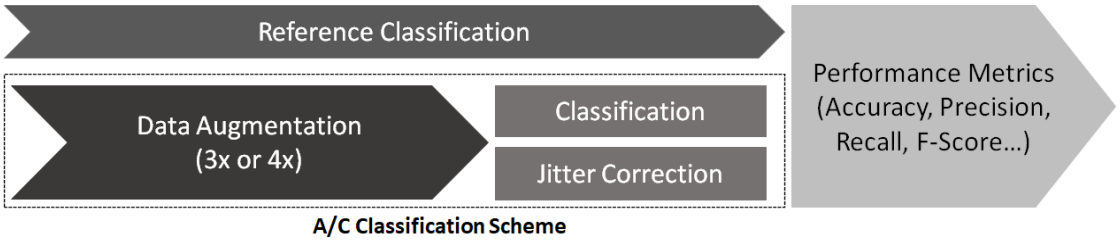

Figure 3.1. The schematic illustrates the basic steps involved in both the reference classification, and the augmented/corrected (A/C) classification method with data augmentation and jitter correction.

3.2.4.1 Latency Jitter

All calculations of latency shifts, latency jitter, and classifier score series relied on classifier-based latency estimation (CBLE), as proposed by Thompson and colleagues [12] and used in our prior investigation of latency jitter [3]. As a first step for CBLE, an SWLDA classifier was trained on either four fifths or all of the training set, depending on whether latency shifts were being calculated on the remaining portion of the training set or on the test set. Then, for each stimulus requiring a latency estimate or classifier score series, whether target or non-target, the downsampled 800 ms epochs, including all 53 time-shifted epochs starting at each available time point from approximately 100 ms pre-stimulus to 100 ms post-stimulus, were extracted and fed to the classifier. This resulted in 53 shifted classifier scores, one for each time shift used. These 53 shifted classifier scores formed the classifier score series for each stimulus. The time shift corresponding to the highest classifier score in the series, representing the highest post-

probability that the stimulus was a target stimulus, was extracted as the latency shift for that specific stimulus. Then latency jitter was defined as the variance of the latency shifts for all target stimuli and denoted as vCBLE for variance of the classifier-based latency estimates.

3.2.4.2 Data Augmentation

For each subject, the training data were augmented using symmetrically time-shifted data similar to the protocol in [25], but with an adaptive time shift based on latency variations in the data. To do so, the median of the absolute latency shifts among target stimuli in each training set for each subject were used as the constant time shift for data augmentation. In this study, individualized parameters were used as participants with ALS generally experience more latency jitter than neurotypical controls (i.e. increased within-subject variability in ALS), and as latency jitter can significantly vary between participants with ALS (i.e. between-subject variability in ALS) [3].

As data augmentation calculation was done within the training set, the training data was first divided into 5 folds of approximately equal length. Then, an SWLDA classifier was trained on data from 4 folds of the original training data using the same procedure as the reference classifier, with no time shifts and was used to calculate the latency shifts for all stimuli in the fifth fold. This procedure was repeated four more times to cover all the stimuli over all folds, providing a classifier score series and an estimated latency shift for every stimulus in the training set, both target and non-target.

To determine the ultimate individualized time-shift, the median of the absolute values of the latency shifts associated with target stimuli was calculated. This median

absolute shift, ***M*** ms, was then used as both a positive and negative constant latency shift with which to augment the data. That is, for each stimulus, epochs shifted by -***M*** ms and ***M*** ms were added to the training set with the original class label for its respective stimulus. An -***M*** to -***M***+800 ms epoch, a 0 to 800 ms epoch, and an ***M*** to ***M***+800 ms epoch were thus constructed for each stimulus in the training set, tripling (3x) the original training data. However, when the latency jitter was above a threshold of 1000 ms^2, a further per-stimulus augmentation was used to correct for the excessive jitter. In this case, for each original epoch, the shifted epoch which maximized the classifier score, corresponding to the latency shift for the stimulus, was added to correct for this increased jitter on the training set. That is, for a stimulus with a latency shift of ***S*** ms based on its classifier score series that reached its maximum for the ***S*** to ***S***+800 ms epoch, this ***S*** to ***S***+800 ms epoch was then added. In this case, the number of epochs extracted from the training set was quadrupled (4x), with each stimulus providing an original 0-800 ms epoch, two symmetrically time shifted epochs, and a jitter corrected epoch.

3.2.4.3 A/C Classification Procedure

After data augmentation was performed on the training set, the parameters were determined for a jitter correction procedure. To do so, for each stimulus (in the training set), classifier score series were calculated using classifiers trained on the augmented data. Then, the maximum classifier score corresponding to an epoch within a limited range of time shifts was retained as the final classifier score. This range was limited because using a narrower range of allowable time shifts reduces the extent to which

using this maximum score increases classifier scores for non-target stimuli and characters, which is a concern when taking the maximum score over an extended range of overlapping windows [12]. The optimal maximum allowable time shift was determined using 5-fold cross-validation over the corresponding training set. The optimal range was selected out of a total of 27 possible window sizes corresponding to the central 1,3,5, …, 53 classifier scores distributed around the score for the 0-800 ms epoch, ranging from no correction to the use of the entire classifier score series. These windows provided maximum allowable time shifts of 0 ms, ±3.91 ms, ± 7.81ms, … ±101.56 ms, corresponding to intervals between data points recorded at 256 Hz.

To determine the optimal range, for each of the possible window sizes, classifier scores and class labels were assigned to each stimulus in the training set using the classifier score series calculated for that stimulus. The score for the stimulus and window of allowable time shifts was the maximum score for the stimulus within that window, and the label was assigned according to this score. In effect, if any epoch within the allowable window of time shifts would have been labeled as a target, then the stimulus was also labeled as a target. If not, then the stimulus was labeled as a non-target. This was repeated for all possible choices of windows, and the window which maximized the average F-score over the five folds was selected as the optimal range for implementation on the test set.

Finally, an SWLDA classifier was trained on the complete augmented training set and then applied to the test session. The classifier score series were calculated for each stimulus in the test session, and the estimated latencies were calculated using CBLE as before. Labels were then also applied to each flash based on the A/C classifier scores.

For comparison, the reference classifier was trained on data from the same original training sets, but without either data augmentation or jitter correction. Estimated latencies were again calculated using CBLE, but latency shifts were not used in determining reference classifier scores, class labels, or character selections.

3.2.4.4 Performance Evaluation

For both types of classifier, binary classification accuracy, precision, recall, F-score, and character selection accuracy were calculated as measures of performance [32, 33]. With TP, TN, FP, and FN respectively representing the number of epochs that were classified as true positives (correct targets), true negatives (correct non-targets), false positives, and false negatives, we computed accuracy, precision, recall, and F-score as below:

$$\text{Classification Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F} - \text{score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

Character selection accuracies were calculated as the number of characters correctly selected from a test session divided by the 14 characters in each session. To do so, the character with the highest summed classification score (either A/C or reference) over all stimuli was selected as the target character. The selection accuracies were calculated using each possible number of trials, from 1 (only the first

intensification of each row and column per character) through 10 (all 10 intensifications of each row and column per character).

3.2.5 Statistical Analysis

All statistical analyses were conducted in R version 3.6.1 [34]. Differences between the proposed A/C classification method and the reference classifier were investigated using paired *t*-tests. Both per-stimulus performance metrics, specifically classification accuracy, precision, recall, and F-score, and character selection accuracy using each possible number of trials, from 1-10, per character were averaged within participants. Participant average jitter, per-stimulus performance metrics, and character selection accuracy for all 10 trials were compared between the proposed A/C classification method and the reference classifier using paired *t*-tests.

We then tested for correlations between performance metrics and latency jitter for both classification methods using repeated measures correlations, (r_{rm}), an analysis of covariance-based regression appropriate for measuring common (overall) intra-individual associations between measures with multiple non-independent observations per participant [35].

We further investigated associations between clinical measures and performance improvements from our proposed method. To do so, we tested for spearman correlations between participant-averages in selection accuracy improvements from our proposed A/C classification procedure relative to the reference classification approach, and time since diagnosis, ALSFRS-R scores, and ALSFRS-R bulbar subscores. We also tested for correlations between selection accuracies using each method and clinical scores.

Finally, latency jitter and performance metrics were investigated longitudinally. We utilized repeated measures correlations to investigate common trends across participants. To understand possible changes in performance over time, repeated measures correlations were investigated between the number of days since the first session and latency jitter, as well as the number of days since the first session and all performance metrics. We then tested for spearman correlations between character selection accuracies and days since their first session within each participant to consider inter-individual differences in trends.

3.3 RESULTS

The symmetric shifts used for data augmentation varied between ±11.72 ms and ±54.69 ms, though shifts greater than ±30 ms were only selected for ALS-01, the participant in the locked-in state. The selected correction windows ranged from 0 (no allowable time shift) to ±101.56 ms, though windows greater than ±40 ms were also only selected for the participant in the locked-in state. The selected parameters for each combination of training and testing session numbers, specifically the symmetric shifts used in data augmentation, the relative size of the augmented training set compared to the original training data, and the jitter correction window, are available in the Appendix table A.1 for each participant and session.

Character accuracy when using all 10 trials was significantly (p=0.019) higher with the A/C classifier at 73.92±31.03% than with the reference classifier at 70.50±31.77%. In particular, these improvements were about 5.0%, 7.1%, 2.4%, 4.1%, 0.9%, and 1.0% for participants ALS-01 through ALS-06, respectively. Binary classification accuracy,

however, was not significantly different between the classification procedures. Precision was significantly (p=0.025) lower with the A/C classifier at 51.14±15.61% compared to the reference classifier with 54.21 ± 15.96% precision. The A/C classifier had a significantly (p=0.002) higher recall of 53.21±21.41% than the reference classifier at 43.46±21.13%. The A/C classifier also provided a significantly (p=0.005) higher F-score of 0.50±0.18 than the reference classifier, at 0.45±0.20. Table 3.2 tabulates the individual results on these measures, averaged over all sessions.

Improvement in character selection accuracy was also observed when fewer trials were used, as shown in Figure 3.2. Character selection accuracy was improved by an average of 5.32% using the proposed A/C classifier as compared to the reference classifier over all numbers of trials and participants. Both initial selection accuracy and the extent of the improvement varied between participants. In particular, for ALS-01, character selection accuracy was improved by 3.6% on average over all possible numbers of trials, though this improvement did not allow for successful BCI control due to poor initial performance. For ALS-02, character selection accuracy was improved by an average of 8.1% over all possible numbers of trials. Character selection accuracy first reached an acceptable level (≥70% [36]), for ALS-02 using at least four trials using the reference classifier at 70.7%, as compared to three trials using the proposed A/C classifier, at 72.1%. For ALS-03, character selection accuracy was improved by an average of 3.3% over all numbers of trials, requiring at least three trials to reach acceptable accuracy with both the reference classifier (71.4%) and the proposed A/C classifier (78.6%). ALS-04 never reached acceptable character selection accuracy, but the proposed classifier improved selection accuracy by an average of 10.3% over all

possible numbers of trials. For ALS-05, this improvement was 3.6%, first achieving an acceptable accuracy using 3 trials at 78.6% with the proposed A/C classification scheme as opposed to 4 trials at 75.9% character selection accuracy with the reference classifier. Over all possible numbers of trials used, character selection accuracy was improved for ALS-06 by 3.1% using the A/C classification scheme as compared to the reference classifier, first achieving an acceptable accuracy using 3 trials at 71.4% with the proposed A/C classification scheme as opposed to 4 trials at 74.5% character selection accuracy with the reference classifier.

Figure 3.3 shows the correlations between performance metrics and latency jitter. Using the proposed A/C classification method, there were significant correlations between latency jitter and four performance metrics, specifically character accuracy (r_{rm}=-0.87, p<0.001), binary classification accuracy (r_{rm}=-0.73, p<0.001), precision (r_{rm}=-0.86, p<0.001), and F-score (r_{rm}=-0.80, p<0.001) indicating that as the latency jitter increased, that the proposed A/C method improved performance overall but did not mitigate the negative relationship between jitter and performance. However, the correlation between latency jitter and recall using the proposed A/C classification method did not reach significance ($p > 0.05$). Using the reference classifier, latency jitter correlated significantly with character selection accuracy (r_{rm}=-0.85, p<0.001), binary classification accuracy (r_{rm}=-0.74, p<0.001), precision (r_{rm}=-0.82, p<0.001), recall (r_{rm}=-0.31, p=0.049), and F-score (r_{rm}=-0.67, p<0.001), for significant correlations with all five performance metrics.

Spearman correlations between participant average character selection accuracies and clinical features, specifically age, time since diagnosis, ALSFRS-R scores, and

ALSFRS-R bulbar subscores, were not significant for either classification method (p>0.05).

Table 3.2. Average accuracy metrics for both reference and augmentation and correction (A/C) classification schemes for each participant. *significant at p<0.05, **significant at p<0.01, paired t-test. Means and standard deviations (SD) are provided for each classification method.

	Character Accuracy (%)*		***Binary Accuracy (%)***		***Precision (%)****		***Recall (%)****		***F-score****	
Participant	Reference	A/C	Reference	A/C	Reference	A/C	Reference	A/C	Reference	A/C
ALS01	10.71	15.71	81.67	78.32	26.10	23.26	7.25	12.89	0.11	0.17
ALS02	87.14	94.29	84.41	83.85	53.59	53.85	61.57	69.75	0.56	0.59
ALS03	97.62	100.00	88.12	88.61	67.48	66.17	58.69	67.26	0.61	0.66
ALS04	60.20	64.29	79.87	77.00	47.06	43.86	28.98	46.43	0.32	0.42
ALS05	85.71	86.61	85.96	84.99	66.15	61.07	53.17	64.06	0.56	0.60
ALS06	81.63	82.65	85.76	84.96	64.85	58.63	51.07	58.88	0.56	0.58
Mean ± SD	**70.50±31.77**	**73.92±31.03**	**84.30±3.03**	**82.95 ±4.42**	**54.21±15.96**	**51.14±15.61**	**43.46±21.13**	**53.21±21.41**	**0.45±0.20**	**0.50±0.18**

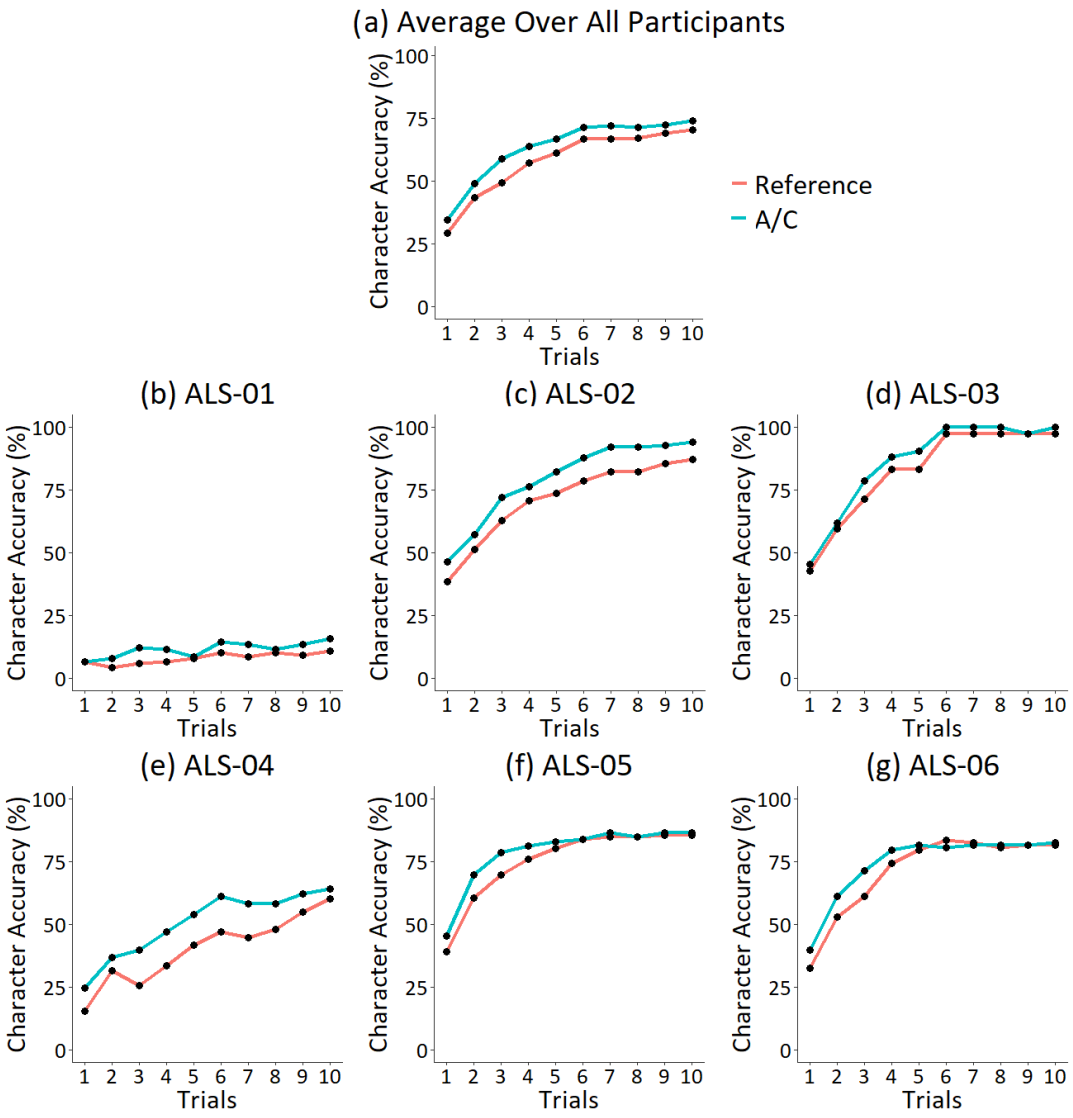

Figure 3.2. Average character selection accuracies at each number of trials used, over all participants (a) and for each participant (b-g) using both the reference and augmentation and correction (A/C) classification schemes.

However, the correlations between clinical scores and participant average *improvements* in character selection accuracy, as shown in Figure 3.4, were significant for ALSFRS-R scores (ρ=-0.94, p=0.017) and the bulbar subscore of the ALSFRS-R (ρ=-0.91, p=0.011). Spearman correlations between performance improvements and age (ρ=-0.43, p=0.419) or time since diagnosis (ρ=0, p=1) were not significant.

Repeated measures correlation plots for the longitudinal analysis of character selection accuracy, binary accuracy, and latency jitter over time are shown in Figure 3.5. Character accuracy decreased significantly over time with both the proposed A/C classification scheme (r_{rm}=-0.44, p=0.005) and the reference classifier (r_{rm}=-0.43, p=0.006). However, the negative trend in binary classification accuracy as sessions progressed was not significant for the A/C classification method (r_{rm}=-0.27, p =0.083) but it was significant for the reference classifier (r_{rm}=-0.38, p=0.015), suggesting that the longitudinal decrease in performance may be mitigated by our proposed classification scheme. Latency jitter increased over time with both the A/C classification scheme (r_{rm}=0.42, p=0.006) and the reference classifier (r_{rm}=0.50, p<0.001).

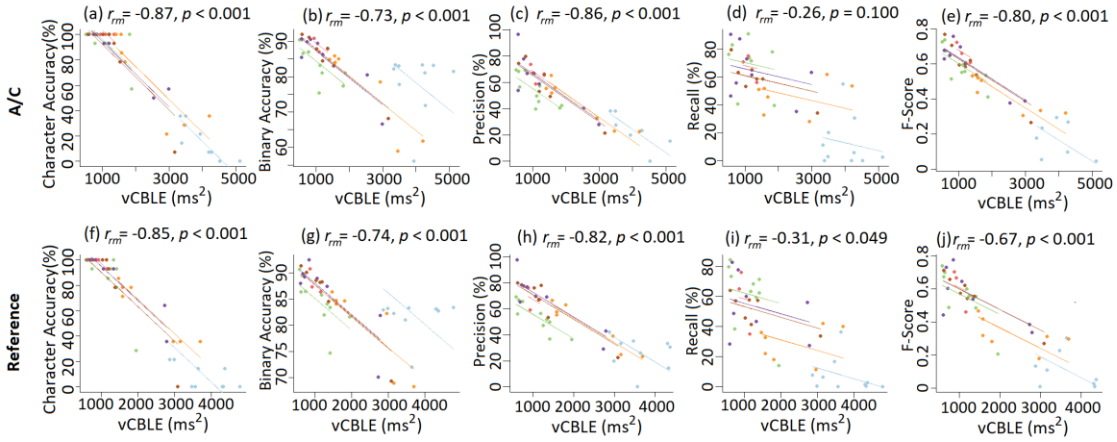

Figure 3.3 Repeated measures correlations between latency jitter (vCBLE, ms^2) and character accuracy (first column, a&d), binary classification accuracy (second column, b&g), precision (third column, c&h), recall (fourth column, d&i), and F-score (fifth column, e&j) in using the proposed A/C classification scheme (top row, a-e) and the reference classifier (bottom row, f-j). Each color indicates one participant.

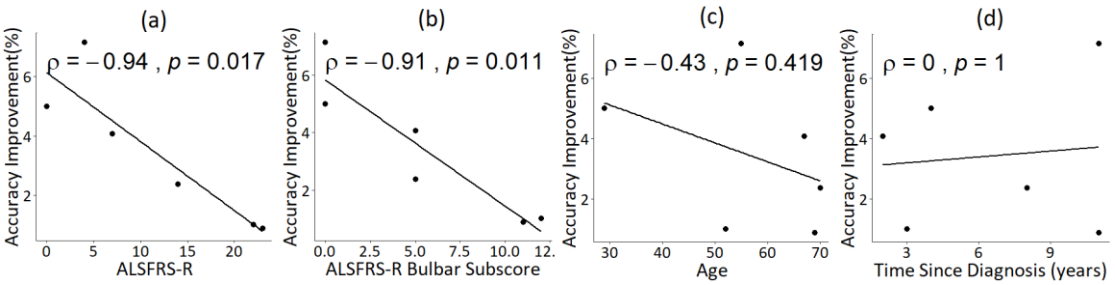

Figure 3.4 Spearman correlation plots of selection accuracy improvements from the proposed augmentation and correction (A/C) scheme with amyotrophic lateral sclerosis functional rating scale-revised (ALSFRS-S) scores (a), ALSFRS-R bulbar subscores (b), participant ages (c), and time since diagnosis (d).

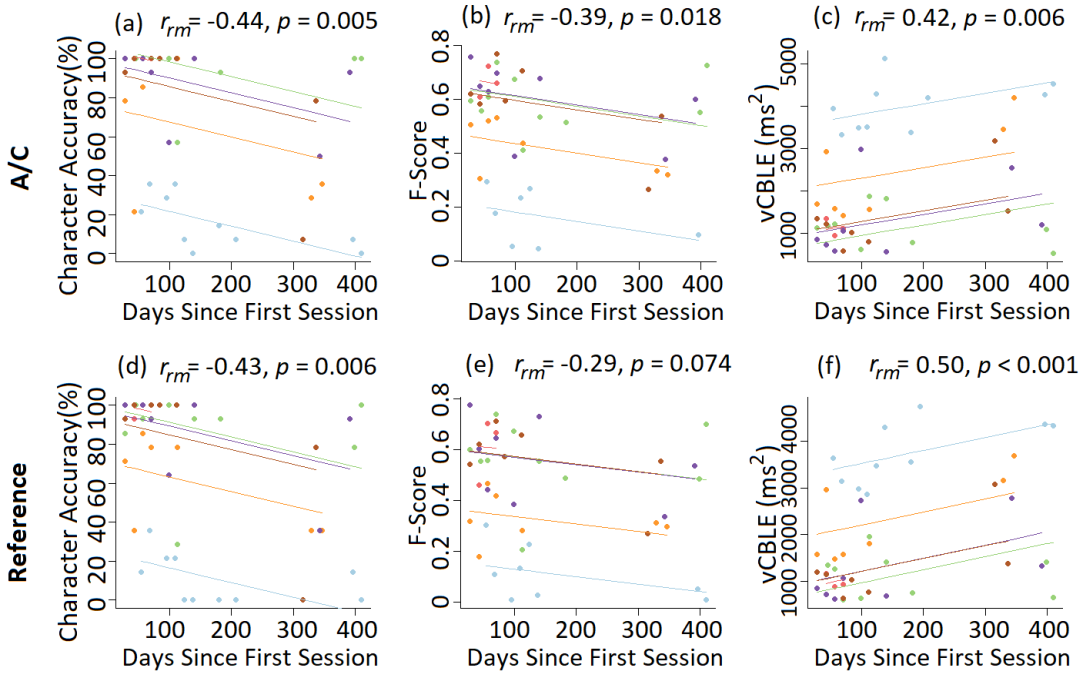

Figure 3.5. Longitudinal repeated measures correlation plots for character accuracy (left, a&d), binary classification accuracy (middle, b&e), latency jitter (vCBLE, right, c&f). Note: each color indicates one participant.

Single-participant longitudinal trends in character selection accuracy are shown in Figure 3.6. Spearman correlations between selection accuracy and the numbers of days since the first session were significant and negative in ALS-01 for both the proposed A/C classification scheme (ρ=-0.75, p=0.013) and the reference classifier (ρ=-0.65,

p=0.041). There was no significant trend in performance over time with either the A/C scheme (ρ=0.09, *p* =0.805) or the reference classifier (ρ=-0.09, *p*=0.808) for ALS-02. There was similarly no significant trend with the proposed (ρ and *p* both undefined) or reference (ρ=0.87, *p*=0.333) classification schemes for ALS-03, for whom performance metrics were only extracted from three sessions. For ALS-04, neither the correlation between accuracy with the A/C scheme (ρ=0.02, *p*=0.969) nor with the reference classifier (ρ=-0.28, *p*=0.542) and time since the first session was significant. The spearman correlations between accuracy and time were the same for both classifiers for both ALS-05 (ρ=-0.64, *p*=0.088) and ALS-06 (ρ=-0.47, *p*=0.284), not reaching significance for either participant.

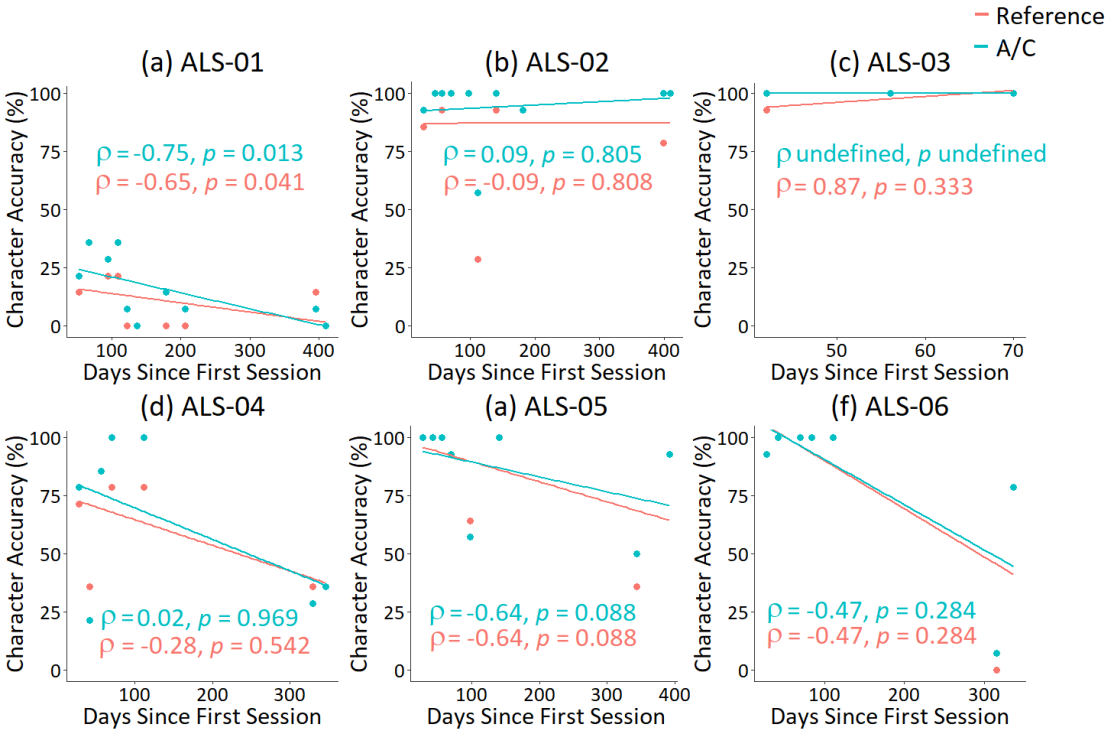

Figure 3.6. Longitudinal plots of single-session character selection accuracies over time for ALS-01 (a) through ALS-06 (f) using both the proposed augmentation and correction (A/C) classification and the reference classification schemes. Each dot represents the result from a single session. For dates where only one dot is visible, the character selection accuracies were the same with both methods.

3.4 DISCUSSION

In this study, we proposed an augmented/corrected (A/C) classification scheme that relies on latency variations at two levels, using both data augmentation and jitter correction procedures to improve P300-based BCI classification performance in people with ALS. Our proposed approach demonstrated significantly improved character selection accuracy and detection of target stimuli relative to classical reference SWLDA classifiers, with greater improvements in selection accuracy in participants with more significant motor impairments. Classification performance improvements with EEG data augmentation were reported to vary based on both tasks and augmentation methods in a recent review paper, though none of the papers covered by that review specifically addressed P300 tasks [19]. However, prior P300 studies have found some success with data augmentation. For example, Krell and colleagues considered multiple augmentation methods found improvements similar to ours using a rotational data augmentation scheme with P300 data. However, their use of one consistent symmetric time-shift to augment P300 data across all neurotypical participants, did not find significant improvement, whereas we showed performance improvements with the individualized time-shifts used in our study [23]. Synthetic oversampling of target samples near the border of target versus non-target has also been reported to improve P300 BCI performance for neurotypical participants with initially poor performance [37]. While we tested our method with an ALS cohort, Bittencourt-Villalpando and Maurits applied a variety of methods to a P300 dataset recorded from autistic adults and found the best performance with a method involving data augmentation for approximately one third of the sessions they considered [38]. Our proposed method,

comparatively, found larger and more consistent improvements in selection accuracy than some prior augmentation approaches with P300 data, and similar improvements to one. Augmentation procedures similar to ours have been implemented in prior studies with neurotypical participants, increasing the amount of training data [23, 25, 26] and thereby improving performance in the two non-P300 studies [25, 26]. These prior studies examined augmentation using constant time-shifts across participants, while the current study determined individual time shifts for each subject separately. Augmentation with symmetric time-shifts has also reported to improve performance in Sakai and colleagues' study using data recorded during an intrinsic motivation task with neurotypical participants [25]. A constant but non-symmetric shift was also used to improve feedback in the detection of error-related potentials, again with neurotypical participants [26]. However, Krell and colleagues found no significant effect on performance after augmenting P300 data with symmetric time-shifts similar in size to the larger selected shifts from the current study [23]. By individualizing the time-shifts used based on latency variations in the data, we were able to both increase the amount of training data and improve performance. We also investigated changes in performance over time to evaluate how our proposed method can facilitate robust long-term use of the P300-based BCI system. While our proposed classification procedure improved performance overall, it could not completely eliminate the decline in performance over time, likely due to the inherent disease progression.

Our jitter correction procedure relying on the maximum classifier score within a given allowable window of time shifts to correct for latency variations similarly improved selection accuracy. Considering this latency variation has also shown

improvement in P300 classification metrics in some prior studies [16, 17]. Prior investigations involving classifier-based latency estimation noted qualitatively that taking the maximum classifier score within a given range of time-shifts as our study did, increased the risk of false-positives, or detecting a P300 response for non-target stimuli, but did not quantitatively specify the size of this increase [12, 16]. Rather than using this maximum score, Mowla and colleagues used a secondary classifier relying on a wavelet transform of the classifier score series to improve performance [16]. Here, by utilizing individualized parameters in the current study, we successfully improved character selection accuracy utilizing these classifier score series without a secondary classifier despite some decrease in precision.

Our longitudinal analyses found that latency jitter increased over time, and performance accordingly decreased over time, using both the reference and proposed A/C classification methods, though deterioration in some metrics was partially mitigated by the proposed scheme. While participants with ALS in the completely locked in state have not often been shown to successfully use visual BCIs [1, 39], prior longitudinal studies which did not involve the completely locked-in state have not typically found BCI performance to decrease over time [6, 8, 40-42]. Several studies have, however, shown significant day-to-day variation in performance [8, 40, 42], which could affect investigations of long-term performance changes depending on the analysis methods used. One prior study found no change over time when comparing copy-spelling accuracies between the first and last several sessions [41]. Sellers 2010 BCI for home use study and Holz's 2015 brain painting study both used single-participant designs [6, 40], while another found long-term trends to vary between

participants [8]. Of our six participants, only one had a significant decline in performance over the course of the study when considered individually, two participants had consistently high performance throughout the study, and three participants appeared to have some decline in performance which did not reach significance when considered individually. It is only by considering common trends across participants with repeated measures correlations that the significant negative trend was uncovered despite both day-to-day and between-participant performance variabilities. BCIs can successfully be used for a significant period of time [6, 8], but the consistent failure of current visual P300 BCIs in the completely locked-in state [1, 39] indicates that performance must eventually decline, as we found to occur in our present study. Given the overall deterioration in performance over time detected in our current study, it is important to note that despite a continued lack of correlation between disability and BCI performance, performance improvements from our proposed A/C scheme were greater in participants with more advanced ALS, demonstrating that our proposed scheme provided greater benefits to more severely disabled participants.

Finally, while our tests of correlations between latency jitter and performance metrics were not a key feature of the study, they confirmed prior results both in our lab [3] and in others work [12-14], namely that increased latency jitter is associated with decreased BCI performance. A classification method that can reduce or eliminate this association, if possible, would likely make BCI performance more robust. However, our proposed method retained this association while improving performance overall.

3.4.1 Limitations and Future Work

One limitation of this study, common to many BCI studies of people with ALS, is the relatively low number of participants, due in part both to the rareness of the disease and the difficulties of recording from this population. We therefore did not analyze differences due to gender, though we did consider clinical features in some analyses. The longitudinal recordings we obtained from each participant, however, provides additional data points, mitigating some limitations related to small sample sizes. The proposed A/C classification method was tested on longitudinal recordings from each participant, and we reported the average results for each participant. For the longitudinal investigation, our use of repeated measures correlations, rather than separately investigating long-term trends for each participant, increased power while maintaining statistical rigor [35]. Future work could also include additional participants and recording sessions.

Another limitation to the current study is inherent to CBLE, which defines a single latency shift for the entire spatiotemporal ERP complex for each stimulus [12, 14]. While Thompson's tests with simulated data show the efficacy of CBLE in reflecting P300 latency jitter [43], future work could investigate latency variations between different ERP components.

Our analyses, while conducted offline, were designed to be appropriate for real-life settings, with all training and parameter selection procedures relying only on data from prior sessions. This would be especially important as practical environments would likely utilize information from prior sessions and/or a short amount of data from the

same session to successfully implement in any upcoming BCI experiment. The current study considers jitter in a simple way relying on individualized parameters to ensure efficacy, and so future work could include the real-time implementation of our proposed A/C method.

3.5 CONCLUSION

In this work, we proposed an augmented/corrected (A/C) classification procedure using both data augmentation and jitter correction schemes to improve P300-based BCI classification performance in people with ALS. The proposed method demonstrated an improvement in selection accuracy which was greater for participants with more significant motor impairments, but which did not show any relationship with age or time since diagnosis. Considering common trends across participants, the current work showed decreased BCI performance over time, which was suggested by BCI inefficiency in the completely locked-in state but not consistently demonstrated in the past. When participants were considered individually, however, longitudinal performance trends varied and did not consistently show decreases, which fits with prior studies. Despite improving selection accuracy and reducing the negative trend in binary classification accuracy over time, our proposed method did not fully eliminate the common downward trend in performance over time.

REFERENCES

[1] A. Kübler and N. Birbaumer, "Brain–computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients?" *Clinical Neurophysiology*, vol. 119, *(11)*, pp. 2658-2666, 2008.

[2] N. Birbaumer, F. Piccione, S. Silvoni and M. Wildgruber, "Ideomotor silence: the case of complete paralysis and brain–computer interfaces (BCI)," *Psychol. Res.*, vol. 76, *(2),* pp. 183-191, 2012.

[3] A. H. Zisk, S. B. Borgheai, J. McLinden, S. M. Hosni, R. J. Deligani and Y. Shahriari, "P300 Latency Jitter and its Correlates in People with Amyotrophic Lateral Sclerosis," Clinical *Neurophysiology,* 2020.

[4] L. M. McCane, E. W. Sellers, D. J. Mcfarland, J. N. Mak, C. S. Carmack, D. Zeitlin, J. R. Wolpaw and T. M. Vaughan, "Brain-computer interface (BCI) evaluation in people with amyotrophic lateral sclerosis," *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration,* vol. 15, *(3-4),* pp. 207-215, 2014.

[5] A. M. Geronimo and Z. Simmons, "The P300 'face' speller is resistant to cognitive decline in ALS," *Brain-Computer Interfaces,* vol. 4, *(4),* pp. 225-235, 2017.

[6] E. W. Sellers, T. M. Vaughan, and J. R. Wolpaw, "A brain-computer interface for long-term independent home use," *Amyotrophic Lat. Sclerosis.*, vol. 11, *(5),* pp. 449-455, 2010.

[7] Y. Shahriari, E. W. Sellers, L. M. McCane, T. M. Vaughan and D. J. Krusienski, "Directional brain functional interaction analysis in patients with amyotrophic lateral sclerosis," in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER),* 2015, pp. 972-975.

[8] Y. Shahriari, T. M. Vaughan, L. McCane, B. Z. Allison, J. R. Wolpaw and D. J. Krusienski, "An exploration of BCI performance variations in people with amyotrophic lateral sclerosis using longitudinal EEG data," *Journal of Neural Engineering,* 2019.

[9] J. Polich, "Updating P300: an integrative theory of P3a and P3b," *Clinical Neurophysiology,* vol. 118, *(10),* pp. 2128-2148, 2007.

[10] J. N. Mak, D. J. McFarland, T. M. Vaughan, L. M. McCane, P. Z. Tsui, D. J. Zeitlin, E. W. Sellers and J. R. Wolpaw, "EEG correlates of P300-based brain–computer interface (BCI) performance in people with amyotrophic lateral sclerosis," *Journal of Neural Engineering,* vol. 9, *(2),* pp. 026014, 2012.

[11] A. Geronimo, Z. Simmons, and S. J. Schiff, "Performance predictors of brain–computer interfaces in patients with amyotrophic lateral sclerosis," *Journal of Neural Engineering,* vol. 13, *(2),* pp. 026002, 2016.

[12] D. E. Thompson, S. Warschausky and J. E. Huggins, "Classifier-based latency estimation: a novel way to estimate and predict BCI accuracy," *Journal of Neural Engineering,* vol. 10, *(1),* pp. 016006, 2012.

[13] P. Aricò, F. Aloise, F. Schettini, S. Salinari, D. Mattia and F. Cincotti, "Influence of P300 latency jitter on event related potential-based brain–computer interface performance," *Journal of Neural Engineering,* vol. 11, *(3),* pp. 035008, 2014.

[14] M. R. Mowla, J. D. Gonzalez-Morales, J. Rico-Martinez, D. A. Ulichnie and D. E. Thompson, "A Comparison of Classification Techniques to Predict Brain-Computer Interfaces Accuracy Using Classifier-Based Latency Estimation," *Brain Sciences,* vol. 10, *(10),* pp. 734, 2020.

[15] S. Vucic, "P300 jitter latency, brain-computer interface and amyotrophic lateral sclerosis," *Clinical Neurophysiology,* 2020. Available: https://doi.org/10.1016/j.clinph.2020.11.017.

[16] M. R. Mowla, J. E. Huggins and D. E. Thompson, "Enhancing P300-BCI performance using latency estimation," *Brain-Computer Interfaces,* vol. 4, *(3),* pp. 137-145, 2017.

[17] R. Togashi and Y. Washizawa, "Feature extraction of P300 signal using Bayesian delay time estimation," in *2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference,* 2013, pp. 1-5.

[18] I. Iturrate, R. Chavarriaga, L. Montesano, J. Minguez and J. Millán, "Latency correction of event-related potentials between different experimental protocols," *Journal of Neural Engineering,* vol. 11, *(3),* pp. 036005, 2014.

[19] E. Lashgari, D. Liang and U. Maoz, "Data augmentation for deep-learning-based electroencephalography," J. *Neuroscience. Methods,* pp. 108885, 2020.

[20] S. M. Abdelfattah, G. M. Abdelrahman and M. Wang, "Augmenting the size of EEG datasets using generative adversarial networks," in *2018 International Joint Conference on Neural Networks (IJCNN),* 2018, pp. 1-6.

[21] M. RIYAD, M. KHALIL and A. Abdellah, "MI-EEGNET: A novel Convolutional Neural Network for motor imagery classification," *Journal of Neuroscience Methods,* pp. 109037, 2020.

[22] M. M. Krell and S. K. Kim, "Rotational data augmentation for electroencephalographic data," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* 2017, pp. 471-474.

[23] M. M. Krell, A. Seeland and S. K. Kim, "Data augmentation for brain-computer interfaces: Analysis on event-related potentials data," arXiv *Preprint arXiv:1801.02730,* 2018.

[24] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy and F. Yger, "A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update," *Journal of Neural Engineering,* vol. 15, *(3),* pp. 031005, 2018.

[25] A. Sakai, Y. Minoda and K. Morikawa, "Data augmentation methods for machine-learning-based classification of bio-signals," in *2017 10th Biomedical Engineering International Conference (BMEiCON)*, pp. 1-4.

[26] S. K. Kim, E. A. Kirchner, A. Stefes and F. Kirchner, "Intrinsic interactive reinforcement learning–Using error-related potentials for real world human-robot interaction," *Scientific Reports*, vol. 7, *(1)*, pp. 1-16, 2017.

[27] J. M. Cedarbaum, N. Stambler, E. Malta, C. Fuller, D. Hilt, B. Thurmond and A. Nakanishi, "The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III)," J. *Neurol. Sci.*, vol. 169, *(1-2)*, pp. 13-21, 1999.

[28] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology.*, vol. 70, *(6)*, pp. 510-523, 1988.

[29] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "Toward enhanced P300 speller performance," *Journal of Neuroscience Methods*, vol. 167, *(1)*, pp. 15-21, 2008.

[30] G. Schalk and J. Mellinger, *A Practical Guide to Brain–computer Interfacing with BCI2000: General-Purpose Software for Brain-Computer Interface Research, Data Acquisition, Stimulus Presentation, and Brain Monitoring.* Springer Science & Business Media, 2010.

[31] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayoudh, D. J. McFarland, T. M. Vaughan and J. R. Wolpaw, "A comparison of classification techniques for the P300 Speller," *Journal of Neural Engineering*, vol. 3, *(4)*, pp. 299, 2006.

[32] M. Pal and S. Bandyopadhyay, "Many-objective feature selection for motor imagery EEG signals using differential evolution and support vector machine," in *2016 International Conference on Microelectronics, Computing and Communications (MicroCom)*, 2016, pp. 1-6.

[33] Z. Tang, C. Li and S. Sun, "Single-trial EEG classification of motor imagery using deep convolutional neural networks," *Optik-International Journal for Light and Electron Optics*, vol. 130, pp. 11-18, 2017.

[34] R. C. Team, "R: A Language and Environment for Statistical Computing," *R Foundation for Statistical Computing*, vol. 3.6.1, 2019.

[35] J. Z. Bakdash and L. R. Marusich, "Repeated measures correlation," *Frontiers in Psychology*, vol. 8, pp. 456, 2017.

[36] G. Pfurtscheller, B. Z. Allison, G. Bauernfeind, C. Brunner, T. Solis Escalante, R. Scherer, T. O. Zander, G. Mueller-Putz, C. Neuper and N. Birbaumer, "The hybrid BCI," *Frontiers in Neuroscience*, vol. 4, pp. 3, 2010.

[37] T. Lee, M. Kim, and S. Kim, "Data augmentation effects using borderline-SMOTE on classification of a P300-based BCI," in *2020 8th International Winter Conference on Brain-Computer Interface (BCI),* 2020, pp. 1-4.

[38] M. Bittencourt-Villalpando and N. M. Maurits, "Linear SVM algorithm optimization for an EEG-based brain-computer interface used by high functioning autism spectrum disorder participants," in *Mediterranean Conference on Medical and Biological Engineering and Computing,* 2019, pp. 1875-1884.

[39] A. R. Murguialday, J. Hill, M. Bensch, S. Martens, S. Halder, F. Nijboer, B. Schoelkopf, N. Birbaumer and A. Gharabaghi, "Transition from the locked in to the completely locked-in state: a physiological analysis," *Clinical Neurophysiology,* vol. 122, *(5),* pp. 925-933, 2011.

[40] E. M. Holz, L. Botrel, T. Kaufmann and A. Kübler, "Long-term independent brain-computer interface home use improves quality of life of a patient in the locked-in state: a case study," *Arch. Phys. Med. Rehabil.,* vol. 96, *(3),* pp. S16-S26, 2015.

[41] J. R. Wolpaw, R. S. Bedlack, D. J. Reda, R. J. Ringer, P. G. Banks, T. M. Vaughan, S. M. Heckman, L. M. McCane, C. S. Carmack and S. Winden, "Independent home use of a brain-computer interface by people with amyotrophic lateral sclerosis," *Neurology,* vol. 91, *(3),* pp. e258-e267, 2018.

[42] S. B. Borgheai, J. McLinden, A. H. Zisk, S. I. Hosni, R. J. Deligani, M. Abtahi, K. Mankodiya and Y. Shahriari, "Enhancing communication for people in late-stage ALS using an fNIRS-based BCI system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 28, *(5),* pp. 1198-1207, 2020.

[43] D. E. Thompson, M. R. Mowla and J. E. Huggins, "Evidence of latency variation in the P3 speller brain computer interface," in *Society for Neuroscience,* October 23, 2019.

CHAPTER 4: A LONGITUDINAL STUDY OF LATENY JITTER AND DISRUPTED INTERRELATIONSHIPS IN ALS USING A WOODY FILTER APPROACH

In preparation for submission to *Clinical Neurophysiology*

Alyssa Hillary Zisk [a], Seyyed Bahram Borgheai [b], John McLinden [b], Sarah M. Hosni [b], Roohollah Jafari Deligani [b], Yalda Shahriari [a,b]

[a] University of Rhode Island, Interdisciplinary Neuroscience Program, Kingston, RI, USA

[b] University of Rhode Island, Electrical, Computer, and Biomedical Engineering Department, Kingston, RI, USA

As first author, Zisk analyzed the data, intepreteted the results, and primarily wrote the text. Dr. Shahriari is the corresponding author and supervised all the aspects of this project, data recording, data analysis, the interpretation of the results, and the manuscript preparation. Dr. Borgheai and John McLinden conducted participant recruitment and data recording. All authors were involved in proofreading and revising the manuscript.

ABSTRACT

Objective: Latency jitter is an important consideration in brain-computer interface (BCI) performance. People with ALS, who may benefit from BCIs, have increased latency jitter. This study investigated latency jitter calculated using two different methods to understand which single-trial features are represented in classifier-based latency estimation (CBLE), compare latency jitter for specific ERP features between people with ALS and neurotypical controls, and longitudinally investigate latency jitter in people with ALS.

Methods: Electroencephalographic (EEG) responses were recorded from six people with ALS and fifteen neurotypical controls. Both single-trial and session-average ERP amplitudes and latencies were extracted. Both a Woody filtering approach and CBLE were used to calculate latency jitter. ERP components and latency jitter were compared between groups using Wilcoxon rank-sum tests. Relationships between measures were investigated within and between sessions using linear regression models, Spearman correlations, and repeated measures correlations.

Results: Latency variations in the four ERP features considered, the N100, P200, N200, and P300, all contributed to whole-epoch latency variations calculated with CBLE in neurotypical participants. However, these contributions were disrupted in participants with ALS, who had increased P200, N200, P300, and whole-epoch latency jitter. Whole-epoch latency jitter increased over time in people with ALS, but N100, P200, N200, and P300 jitter did not. Neither whole-epoch nor ERP feature latency jitter correlated with clinical scores in participants with ALS.

Conclusions: CBLE not only reflects P300 latency variations, as expected, but it also reflects single-trial latency variations in other ERP features. Latency jitter is increased in ALS for several ERP features, including but not limited to the P300. Correlations between single-trial latencies for ERP features and whole-epoch latency shifts are disrupted in ALS.

Significance: The presence of latency jitter in several ERP features can inform future BCI designs meant to compensate for latency jitter in people with ALS. CBLE has now been tested against another, established method of calculating single-trial latency shifts and found to reflect latency jitter.

4.1 INTRODUCTION

P300-based brain computer interfaces can support communication for people with neuromuscular disabilities, including amyotrophic lateral sclerosis (ALS) (Allison, Kübler, & Jin, 2020; Geronimo, Andrew M. & Simmons, 2017; McCane et al., 2014; McCane et al., 2015; Wolpaw et al., 2018). However, the P300, a positive deflection that occurs approximately 300 ms after an attended, unusual stimulus, can have significant trial-to-trial latency variability, or jitter (Aricò et al., 2014; Fjell, Rosquist, & Walhovd, 2009; Yu, 2016). This variability has been studied in cognitive contexts for some time, where single-trial P300 latencies are associated with stimulus evaluation times (Kelly & O'Connell, 2013; Verleger, 1997) and single-trial reaction times (Saville et al., 2011), particularly when the focus is on accuracy rather than speed (Kutas, McCarthy, & Donchin, 1977). However, this relationship is disrupted in neurotypical participants with comparatively higher P300 latency jitter (Saville et al., 2011) and when

the task prioritizes speed (Kutas et al., 1977; Verleger, 1997). Latency variability can be studied as one form of neural variability (Magnuson, Iarocci, Doesburg, & Moreno, 2020), of which some is required for learning but excess is found in a variety of neurological conditions (Dinstein, Heeger, & Behrmann, 2015). Increased P300 jitter is found in people with ALS (Zisk et al., 2020), attention deficit hyperactivity disorder (Saville et al., 2015), schizophrenia (Ford, White, Lim, & Pfefferbaum, 1994), depression (Patterson, Michalewski, & Starr, 1988), traumatic brain injuries (Unsal & Segalowitz, 1995), disorders of consciousness (Schettini et al., 2015), and dementia (Patterson et al., 1988).

In the context of brain computer interfaces, latency jitter is important to consider because increased latency jitter is associated with decreased performance (Aricò et al., 2014; Huggins, Alcaide-Aguirre, & Hill, 2016; Mowla, Gonzalez-Morales, Rico-Martinez, Ulichnie, & Thompson, 2020; Thompson, Warschausky, & Huggins, 2012; Zisk et al., 2020; Zisk, Borgheai, McLinden, & Shahriari, 2021). Thompson and colleagues proposed classifier-based latency estimation (CBLE) as one way to estimate single-trial latency shifts, and thus latency jitter, during brain computer interface use (Thompson et al., 2012). Latency jitter calculated as the variance in the shifts calculated with CBLE is even able to predict BCI performance (Thompson et al., 2012). However, as studies using CBLE note, there are some theoretical weaknesses to CBLE in measuring single-trial latencies and latency jitter. In particular, CBLE assumes the entire event-related potential (ERP) shifts together, providing a single latency shift for all ERP features, across all channels (Mowla et al., 2020; Thompson et al., 2012; Zisk et al., 2020). While artificially added P300 jitter is reflected in CBLE estimates of jitter

(Thompson, Mowla, & Huggins, October 23, 2019), this theoretical limitation of CBLE means we do not know if latency variation in other ERP features may also be reflected in CBLE, which is possible as other ERP features including the N100, P200, and N200 are affected by "P300" speller paradigms (Allison et al., 2020). It also means that CBLE cannot be used to study latency variations between different ERP components (Thompson et al., 2012), which can vary separately in oddball paradigms (Michalewski, Prasher, & Starr, 1986) and could therefore reasonably be expected to vary separately in the use of P300-based BCIs, which rely on a version of the oddball paradigm (Farwell & Donchin, 1988). CBLE also has not yet been validated directly against other ways of measuring single-trial latencies (Thompson et al., 2012).

Other methods of measuring single-trial latencies include filtering and peak-picking methods (Ouyang, Hildebrandt, Sommer, & Zhou, 2017), ranging from low pass filters with low cut-off frequencies (Jaśkowski & Verleger, 2000; Magnuson et al., 2020; Smulders, Kenemans, & Kok, 1994) to more complex methods including wavelet-based filtering (Aricò et al., 2014; Chennu, Craston, Wyble, & Bowman, 2009) or cross-correlational filters such as Woody's adaptive filter realigning peaks through iterative correlations of time-shifted single-trial responses with the averaged responses (Woody, 1967). Both peak-picking and cross-correlational approaches can be applied to spatially filtered data (Coles, Gratton, Kramer, & Miller, 1986; Fabiani, Karis, & Donchin, 1986; Saville et al., 2011; Saville et al., 2015; Yu, 2016), and cross-correlational approaches similar to the Woody filter applied to time series spatially combining data from multiple channels are a comparatively reliable way to assess both single-trial amplitudes and latencies (Fabiani, Gratton, Karis, & Donchin, 1987). In

addition, Woody filtering approaches have successfully been used to extract single-trial latency information for several ERP features simultaneously, using appropriate segmentation of time series (Michalewski et al., 1986; Patterson et al., 1988). Both the reliability of Woody filtering methods and their successful use in extracting separate single-trial latencies for multiple ERP features make them ideal for use alongside CBLE.

In the current study, we examine which ERP features are reflected by CBLE in both a neurotypical population and participants with ALS, considering the N100, P200, N200, and P300. We also determine which component jitters are relevant to BCI performance in both populations, with a longitudinal investigation of N100, P200, N200, and P300 jitter in ALS for a more specific understanding of this neurological feature.

4.2 METHODS

4.2.1 Participants

16 neurotypical participants (62.5±4.5 years; 10 female) were recruited for this study. Neurotypical participants had normal or corrected to normal vision.

In addition, six participants with ALS (age 57±15.7 years,1 female) were recruited for this study (see Table 1). Participants with ALS other than ALS-01 had normal or corrected to normal vision, while ALS-01 was in the late stages of locked-in syndrome with significant ocular impairments. Participants were diagnosed with ALS 6.5 ± 4.0 years prior to the start of the study and had an average functional rating scale-revised (ALSFRS-R) score of 11.6 ± 9.5, with a minimum score of 0 indicating no voluntary motor functions and complete dependence on life-sustaining technologies including

mechanical ventilation and a maximum score of 48 indicating normal functioning (Cedarbaum et al., 1999). Three participants had gastrostomies as well as tracheostomies. ALS-01′s sole form of communication was an idiosyncratic and error-prone yes/no pupil dilation his caregiver read subjectively, which deteriorated over the course of the recordings, losing reliability as a means of communication. Two other participants with artificial ventilation (ALS-02 and ALS-04) used eye-tracking devices to communicate. ALS-03 could still move his index finger and make non-verbal sounds to sustain minimal communication. ALS-05 and 06 retained the ability to speak, though ALS-05 had lost non-facial movement, and ALS-06 could barely move a joystick with one hand. Participants with ALS were tested in their homes or care centers.

Table 4.1. Demographic information for participants with ALS.

Subject Number	***Age***	***Sex***	***Time since diagno-sis (years)***	***ALSFRS-R (out of 48)***	***ALSFRS-R Bulbar Subscore***	***Average ALS-CBS Score (%)***	***Average Attention Subscore (%)***	***Artificial Ventilation***	***Means of Communi-cation***
ALS-01	29	M	4	0	0	100.0	100.0	Yes	No reliable means
ALS-02	55	M	11	4	0	93.4	90.0	Yes	Eye-tracking
ALS-03	70	M	8	14	5	94.9	80.0	No	Non-verbal sound
ALS-04	67	M	2	7	5	88.9	90.0	Yes	Eye-tracking
ALS-05	69	F	11	23	11	81.3	58.6	No	Verbal
ALS-06	52	M	3	22	12	91.9	67.5	No	Verbal
Mean±SD	**57.0±15.7**	**-**	**6.5±4.0**	**11.6±9.5**	**5.5±5.2**	**92.1±6.8**	**82.8±18.7**	-	-

Both neurotypical participants and participants with ALS had at least some postsecondary education. The study protocol was approved by the Institutional Review Board (IRB) of the University of Rhode Island (URI), and all participants provided informed consent or assent for the study and received financial compensation.

4.2.2 Experimental Protocol

Neurotypical participants took part in 2-3 sessions of recordings. Participants with ALS took part in 5-12 (9.5 ± 2.6) sessions of recording over 2.5-13.7 (10.9 ± 4.3) months. Sessions took place at least two weeks apart. Including preparation such as the application of gel to electrodes and impedance calibration, each session typically lasted 2-2.5 hours. To familiarize participants with the BCI setup, including the recording protocol and the task, participants with ALS took part in a single familiarization session before the main experimental recordings, in which they completed the same tasks without recording the data and were given the opportunity to get clarification about the experimental tasks. Each session contained one run of the standard P300 spelling protocol, in which a 6x6 matrix of characters containing letters and numbers was displayed to participants, with each row and column intensified 10 times (i.e. 10 trials) per character selection (Farwell & Donchin, 1988).

Participants with ALS additionally took the ALS-Cognitive Behavioral Screen (ALS-CBS), a brief cognitive screen sensitive to frontal dysfunctions for people with ALS, when possible (Woolley et al., 2010). Both single-session and participant average scores were reported as percentages to compensate for the fact that not all items could always be used, and cognitive testing could not be completed for all sessions. Because

several participants with ALS had difficulty speaking or writing, the information and retrieval (fluency) section of the ALS-CBS test could not be used effectively. Consequently, only the attention, concentration, and tracking portions of the ALS-CBS test were performed. Due to their disabilities, four participants with ALS required accommodation to complete these portions of the assessment. ALS-01 completed the test once, using a P300 speller. ALS-02 used his typical eye-tracking system. ALS-03 used a printed letter board, pointing with a finger. ALS-04 initially used a letter board, but later used a Tobii eye-tracking system. ALS-05 and ALS-06 did not require accommodations to complete the ALS-CBS.

Data from 9 of the neurotypical participants in this study, as well as three sessions from each participant with ALS, were previously reported in Clinical Neurophysiology (Zisk et al., 2020). Longitudinal investigations of performance using the data from participants with ALS are under review at the Journal of Neural Engineering (Zisk et al., 2021).

4.2.3 Data Acquisition

Electroencephalography (EEG) data were recorded using a g.USBamp amplifier (g.tec Medical Technologies) with a 256 Hz sampling rate. Data were recorded from eight channels commonly used in P300 protocols, Fz*, Cz, P3, Pz, P4, PO7, PO8, and Oz (Krusienski, D. J., Sellers, McFarland, Vaughan, & Wolpaw, 2008). However, as Fz was occupied by sensors for other studies recorded in the same session as the current experiment, it was replaced by the nearest available channel, FAF2, denoted as Fz*. All

experimental protocols, data acquisition, and stimulus presentation were controlled using BCI2000 software (Schalk & Mellinger, 2010).

4.2.4 Data Pre-processing

Data processing was conducted in MATLAB R2019a. EEG data from each session were detrended and bandpass filtered at 0.5-30 Hz. The data were segmented into 100 ms pre-stimulus to 900 ms post-stimulus epochs. The average amplitudes and latencies corresponding to four primary ERP components (N100, P200, N200, P300) were extracted for each channel and session. The N100 and N200 components were respectively defined as the minimum peaks occurring in the 80–170 ms and 220–350 ms periods. The P200 and P300 components were respectively defined as the maximum peaks in the 190–300 ms and 300–500 ms periods.

4.2.5 Data Processing

Single-trial features were extracted in two main ways, first using classifier-based latency estimation, and then using a Woody filtering procedure applied to spatially filtered data, which provides single-trial amplitude and latency information separately for each of the N100, P200, N200, and P300 features. Figure 4.1 shows a schematic illustrating the two parallel extraction procedures. BCI performance was evaluated using the same classifiers used in classifier-based latency estimation.

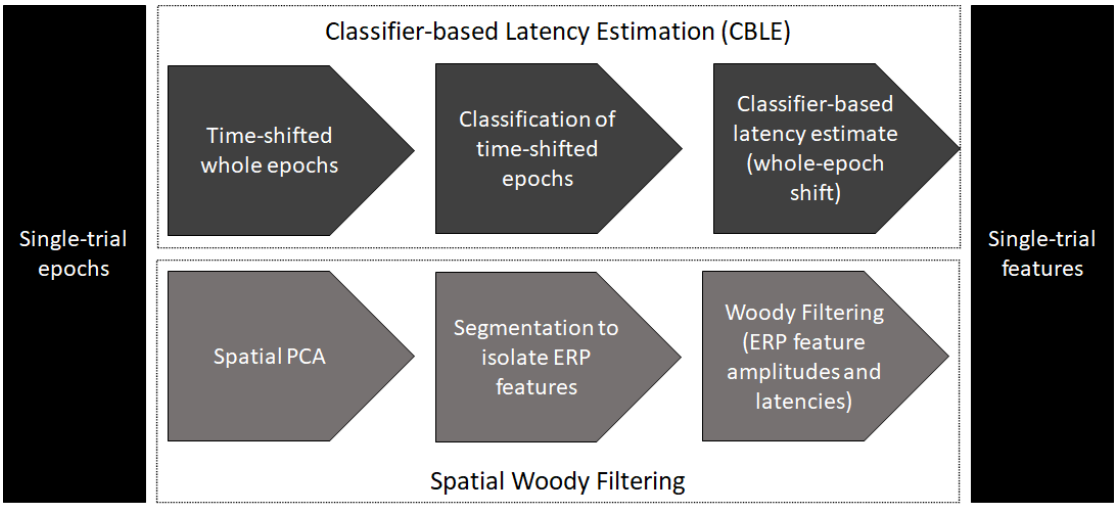

Figure 4.1. Schematic representation of the two methods of single-trial feature extraction – Classifier based latency estimation (CBLE; Thompson et al., 2012), and spatial Woody filtering.

4.2.5.1 Single-Trial Feature Extraction

CBLE was used to extract whole-epoch latency shifts and latency jitter were extracted using classifier-based latency estimation (CBLE), as proposed by Thompson and colleagues (Thompson et al., 2012) and used in our prior investigation of latency jitter (Zisk et al., 2020). As a first step for CBLE, the session data were divided into five segments of approximately equal length, and an SWLDA classifier was trained on data from four of the five segments using typical parameters for P300 speller applications (Krusienski et al., 2008; Krusienski, Dean J. et al., 2006). In particular, 0-800 ms post-stimulus sub-epochs were extracted from each of the 100 ms pre-stimulus to 900 ms post-stimulus epochs. These 800 ms epochs were subject to a moving average procedure, where each value was replaced by the local mean calculated over a moving window and then downsampled by a factor of 13, following the feature reduction procedure from (Thompson et al., 2012). The downsampled epochs from all channels were concatenated

and then treated as potential features for classification. Then, 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 variables ($p > 0.15$) were removed. This procedure was repeated for up to 60 steps, or until no additional terms satisfied the entry/removal criteria (Krusienski et al., 2008).

Then, returning to the original 100 ms pre-stimulus to 900 ms post-stimulus epochs, 800 ms sub-epochs were extracted using a moving window to produce epochs beginning at each available time point, for a total of 53 time-shifted 800 ms epochs per stimulus. Similar to the 0-800 ms epochs, these time-shifted 800 ms epochs were subject to the feature reduction procedure from (Thompson et al., 2012). These time-shifted epochs were then fed to the appropriate classifier for their corresponding stimulus, which resulted in 53 shifted classifier scores per stimulus, one for each time shift used. The time shift corresponding to the highest classifier score, representing the highest post-probability that the stimulus was a target stimulus, was extracted as the latency shift for that specific stimulus. Whole epoch-latency jitter was then defined as the variance of the whole-epoch latency shifts for all target stimuli and denoted as vCBLE for variance of the classifier-based latency estimates.

Because CBLE, despite its success in predicting BCI performance and reflection of P300 latency jitter, has the theoretical weakness of considering latency shifts of the entire epoch, we additionally extracted single-trial amplitudes and latencies for specific ERP features using another method. This second method, extracting single-trial features using a Woody (Woody, 1967) filter on the single-trial time series for selected spatial factors, supports investigation of which single-trial features are reflected in CBLE, as

well as potential relationships between single-trial latencies of the ERP features themselves. Woody filtering was selected because cross-correlational measures applied to spatially filtered EEG provide comparatively reliable estimates of single-trial amplitudes and latencies (Fabiani et al., 1987). Spatial factors were extracted using principal component analysis (PCA) using the covariance matrices calculated from the 0-800 ms post-stimulus epochs for each target stimulus and then averaged over all such stimuli in a session. The eigenvectors and eigenvalues of the average covariance matrices then constituted the channel weights for the spatial factors and the proportion of variance accounted for, respectively (Cohen, 2014). The spatial factors accounting for at least 95% of the variance in the data were retained for possible selection.

Then, appropriate spatial factors were selected for each ERP feature of interest based on correlation with templates for the respective feature over the corresponding time segment of interest (Wu et al., 2014). These templates were based on session-average ERP segments around the ERP peaks on relevant channels (Wu et al., 2014). In particular, the templates for the N100 were the session-average target responses at channels PO7 and PO8 (Espeseth, Endestad, Rootwelt, & Reinvang, 2007; Kimura, Katayama, Ohira, & Schröger, 2009) in the 90 ms surrounding the session average peak. These segments matched the 80-170 ms segment in length, but they were centered around the channel peaks. The templates for the P200 were the session-average target responses at Fz* and Cz (Shahriari et al., 2019), similarly using 110 ms segments surrounding the P200 peaks to match the 190-300 ms segment in length. The templates for the N200 were the 130 ms long session-average target responses at P3 and P4 surrounding the N200 latency (Hoffmann, Vesin, Ebrahimi, & Diserens, 2008; Kimura

et al., 2009; Shahriari et al., 2019). The templates for the P300 were the 200 ms long session-average target responses at Cz and Pz, centered at the P300 peaks for those channels (Espeseth et al., 2007; McCane et al., 2015). Each response used templates from two channels to allow for potential differences in localization between participants. This could include differences in lateralization related to handedness and/or differences in localization due to ALS (McCane et al., 2015). Channel templates for two sample sessions, one recorded from a neurotypical participant and one from a participant with ALS, are shown in Figure 4.2.

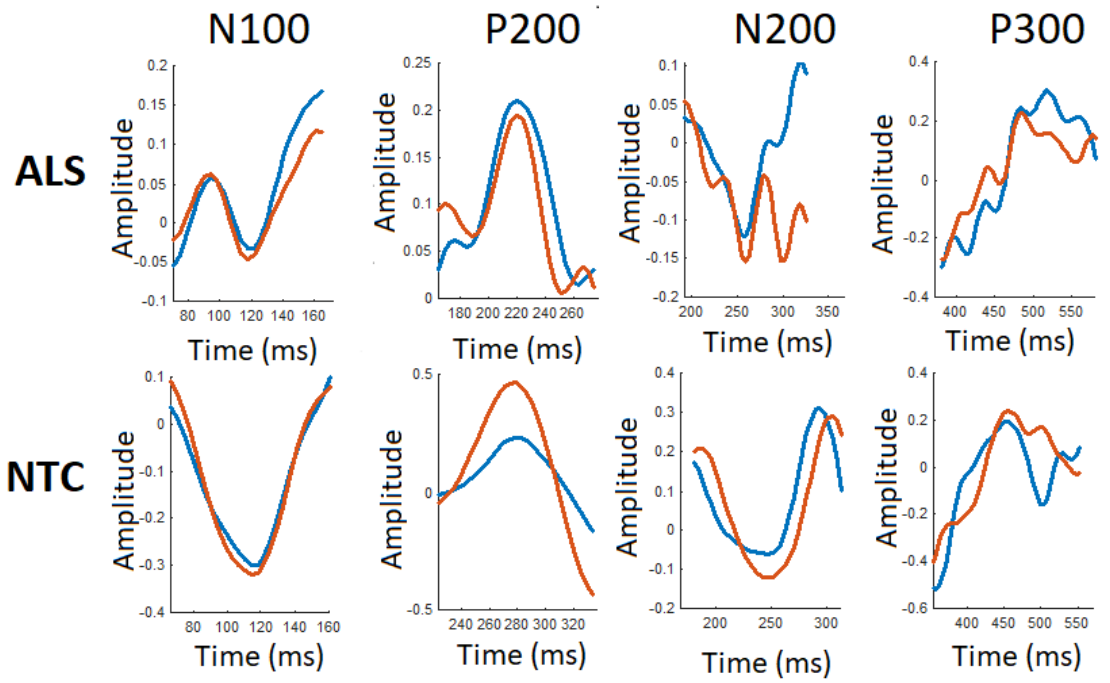

Figure 4.2. Channel templates used in the selection of spatial factors for the N100, P200, N200, and P300 responses for one session recorded from a participant with ALS (top) and one session recorded from a neurotypical participant (bottom).

Pearson correlation coefficients were calculated between each template time segment and time-shifted corresponding session-average target time segments on each spatial factor. For example, Pearson correlation coefficients were calculated between

the P300 template from channel Cz, and segments extracted from the time series of the first spatial factor, starting from 270-470 ms post stimulus through 330-530 ms post-stimulus. These shifted segments allowed time-shifts of 30 ms in either direction in the selection of the spatial factor, as used by Wu and colleagues (Wu et al., 2014). The segment with the largest absolute value of the correlation was retained. This was repeated for each spatial factor, and with the P300 template from channel Pz, the other channel of interest for the P300. The spatial factor with the greatest maximum absolute correlations was retained as the spatial factor for that ERP, in this example the P300. In the event that the two templates for any given ERP feature had their strongest time-shifted correlations with two different spatial factors, the absolute value of the sum of the correlation coefficients was taken over each of the two spatial factors, and the spatial factor with the larger of the two absolute sums was retained. Figure 4.3 shows the spatial factors and associated time series for the same sessions shown in Figure 4.2.

After spatial factors were selected for each ERP feature (N100, P200, N200, and P300), the templates for single-trial matching were extracted from the time series of their respective spatial factors. The starting and ending latencies for the segments that maximized correlation between the selected spatial factor and the channel templates were averaged over the two channels of interest, and these average values were used for the segment of the spatial factor time series. Considering the P300 as an example, if the absolute correlation between the Cz template and the time series for the selected spatial factor was maximized 300-500 ms segment, and the absolute correlation between the Pz template and the time series for the selected spatial factor was maximized for 308-508 ms segment, then the spatial factor template would use the 304-504 ms segment of

its session-average time series. Figure 4.4 shows the spatial template time series for the same sessions used in Figures 4.2 and 4.3. As sign becomes arbitrary after spatial PCA, if a selected spatial factor was inverted (indicated by a negative correlation between the time series and the channel templates), then its corresponding time segments were also inverted, such that nominally positive responses would have positive deflections and nominally negative responses would have negative deflections. After this possible transformation, average ERP amplitudes and latencies were extracted as the extreme peaks or troughs within the session-average segments.

Finally, single-trial amplitudes and latencies were extracted from the single-trial spatial factor segments using a Woody filtering procedure (Woody, 1967). For this purpose, post-stimulus single-trial segments were extracted for each ERP feature using their respective spatial factors: 80-170 ms for the N100, 190-300 ms for the P200, 220-350 ms for the N200, and 300-500 ms for the P300. For each ERP feature and trial, the cross-covariance was calculated between the appropriate spatial factor template and the single-trial segments. The latency that maximized the cross-covariance was used as the single-trial latency, and the value of the cross-covariance was used as a measure of single-trial amplitude (Fabiani et al., 1986; Fabiani et al., 1987).

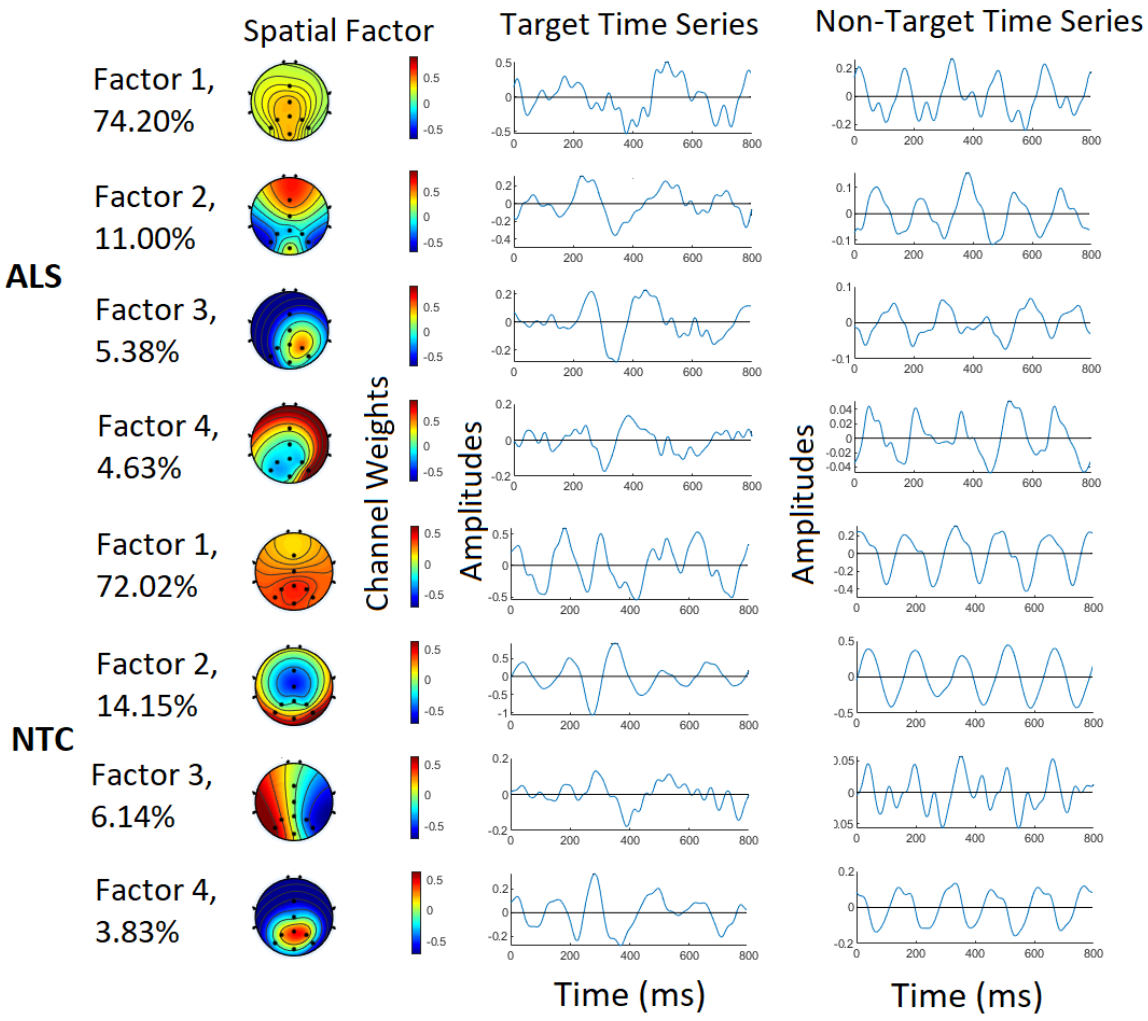

Figure 4.3. Spatial factors and associated time series for one session recorded from a participant with ALS (top) and one session recorded from a neurotypical participant (bottom). For the session recorded from a participant with ALS, factor 1 was selected for the N100, factor 3 for the P200 and P300, and factor 4 for the N200. For the session recorded from a neurotypical participant, factor 2 was selected for the N100 and N200, and factor 4 was selected with inverted sign for the P200 and P300.

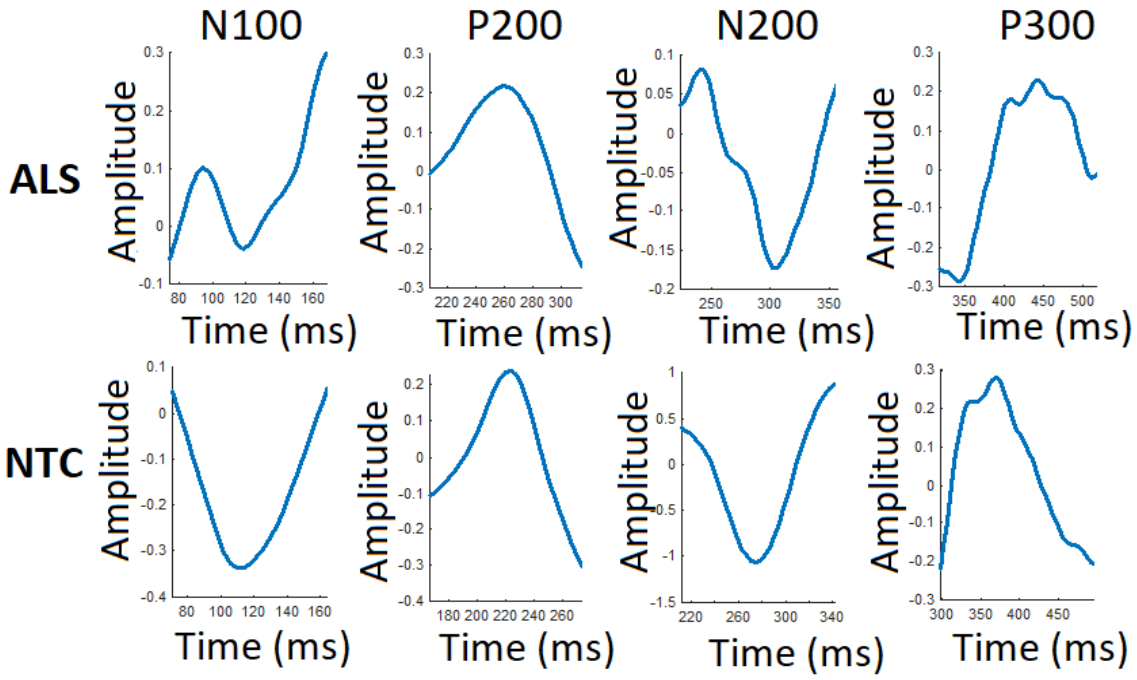

Figure 4.4. Spatial factor templates used in the Woody filtering procedure for the N100, P200, N200, and P300 responses for one session recorded from a participant with ALS (top) and one session recorded from a neurotypical participant (bottom).

4.2.5.2 BCI Performance

The same SWLDA classifiers used for CBLE were again used to investigate BCI performance. Flash-by-flash binary classification performance metrics were calculated on each test set, with average performances extracted from each session for use in correlation analysis and from each participant for between-group comparisons. In particular, binary flash accuracy, precision, recall, F-score, and character selection accuracy were calculated as measures of performance (Pal & Bandyopadhyay, 2016; Tang, Li, & Sun, 2017). We computed classification accuracy, precision, recall, and F-score below, using the numbers of true positives (TP; correct targets), true negatives

(TN; correct non-targets), false positives (FP; non-targets incorrectly classified as targets), and false negatives (FN; targets incorrectly classified as non-targets) as below:

$$\text{Classification Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F} - \text{score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

The number of correctly selected characters was determined in each of the five data segments, and the session accuracy was calculated as the average character selection accuracy over all five segments.

4.2.6 Statistical Analysis

Relationships between single-trial measures were investigated within sessions in MATLAB R2019a. Within each session, we investigated relationships between single-trial N100, P200, N200, and P300 latencies and whole-epoch CBLE shifts using linear regression models to understand how latency shifts for individual ERP features affect the calculated latency shifts for the whole epoch across all channels. We additionally investigated relationships between N100, P200, N200, and P300 amplitudes and absolute CBLE shifts to test the hypothesis that CBLE shifts may be increased when single-trial amplitudes are decreased, leading to reduced signal strength for classification. Absolute CBLE shifts were used because decreased amplitudes could lead to responses being detected either earlier or later than average due to a reduced

signal to noise ratio, rather than causing an increased latency shift in a specific direction.

In addition, linear models were constructed to investigate how the latencies of later ERP features depend on the latencies of earlier features. Specifically, we investigated how N100 latencies contribute to P200 latencies, how N100 and P200 latencies contribute to N200 latencies, and how N100, P200, and N200 latencies contribute to P300 latencies within each session, using linear regression models.

Finally, we investigated relationships between single-trial latency shifts and amplitudes using spearman correlations, again testing if smaller responses tended to be detected as also having latencies further from the center (session average) latency. Specifically, we calculated spearman correlations between the amplitudes and absolute latency shifts within each ERP feature.

As the channel weights, selected spatial factors, and magnitudes of the Woody templates varied between sessions, single-trial measures of amplitude were only used for within-session analyses. However, the variances of all single-trial latency measures were extracted from each session for use in statistical analyses combining information from multiple sessions. These analyses were then conducted in R version 3.6.1 (Team, 2019). When applicable, p-values were adjusted for multiple comparisons using the false discovery rate (Hochberg & Benjamini, 1990).

4.2.6.1 Understanding P300 BCI Latency Variations in a Neurotypical Population

Because it is not yet known if or how single-trial amplitudes and/or latencies in features other than the P300 contribute to CBLE, we first sought to understand these

relationships in a neurotypical population. The results of the within-session regressions were therefore combined to provide estimates of partial correlation coefficients (Aloe & Thompson, 2013; Aloe, 2014) for the predictors using the metafor package in R (Viechtbauer, 2010). This meta-analysis of regressions from all sessions used a mixed-effects model using a maximum likelihood estimate of heterogeneity (Viechtbauer, 2005). This meta-analysis was completed for all of the within-session linear regression models, i.e., for the model of CBLE shifts on all single-trial ERP latencies, for absolute CBLE shifts on all single-trial ERP amplitudes, for P300 latencies on N100, P200, and N200 latencies, for N200 latencies on N100 and P200 latencies, and on P200 latencies on N100 latencies.

We additionally quantified the portion of sessions in which the linear regressions models of CBLE shifts on the single-trial latencies and/or amplitudes showed significant effects. We also quantified the portion of sessions in which each individual ERP features (amplitudes and latencies for each of the N100, P200, N200, and P300) were significant to understand which ERP features contribute to classifier-based latency estimates.

We similarly quantified the portion of sessions in which linear regressions of the latencies of later responses modeled on the latencies of earlier responses were significant, and we identified which earlier response latencies were significant contributors to which later response latencies.

The spearman correlations between absolute single-trial shifts and amplitudes for each ERP feature (N100, P200, N200, and P300) were meta-analytically combined with the R package metafor (Viechtbauer, 2010). We used a mixed-effects model with a restricted maximum likelihood estimate of heterogeneity (Viechtbauer, 2005) and the

estimate of variance for spearman coefficients (Bonett & Wright, 2000). We additionally quantified the portion of correlations that were significant within each ERP feature.

Then, as whole-epoch latency jitter measured by the variance of CBLE estimates (vCBLE) is of interest largely because of its effectiveness as a predictor of BCI performance, we additionally constructed linear regression models of both vCBLE and BCI performance on session variance measures to determine which ERP variations are most relevant to BCI performance.

4.2.6.2 A Longitudinal Analysis of BCI Performance and Intra-Session Variability in People with ALS

All the analyses conducted in section 4.2.5.1 were completed over the longitudinal sessions recorded from participants with ALS as well. As participants with ALS completed several recording sessions over the course of months, we additionally investigated how certain measures change over time. Longitudinal changes in vCBLE and in BCI performance were investigated in (Zisk et al., 2021) using data from these same recording sessions. Longitudinal trends in BCI performance are therefore not repeated here. However, single-trial latencies for specific ERP features were not extracted in our prior work, and changes in these measures over time were investigated using repeated measures correlations (r_{rm}), and the longitudinal trend in vCBLE is shown for comparison. Repeated measures correlations are analysis of covariance-based regressions appropriate for measuring common (overall) intra-individual associations between measures when multiple non-independent observations are available for each

participant (Bakdash & Marusich, 2017). These repeated measures correlations were investigated between the variances of these single-trial latencies for the N100, P200, N200, and P300 and the number of days since a participant's first recording session.

As participants with ALS also participated in cognitive testing when practical, we also investigated the relationships between cognitive scores and our extracted measures. Over the sessions in which cognitive testing was completed, repeated measures correlations were therefore investigated between the variances of single-trial latencies and cognitive scores, specifically the ALS-CBS scores and attention subscores. As cognitive scores were not investigated in (Zisk et al., 2021), repeated measures correlations between cognitive scores and vCBLE, as well as between cognitive scores and BCI performance measures, were also tested.

Finally, as previous correlations between session-average amplitudes and vCBLE were significant for neurotypical participants, but the correlations were not significant in participants with ALS, for whom these amplitudes were reduced (Zisk et al., 2020), we tested for correlations between session-average amplitudes and vCBLE in this population. This was done to investigate whether the prior lack of significant correlation was due to having fewer participants with ALS than neurotypical controls in the prior study, or if the relationship between session-average amplitudes and latency jitter may truly be disrupted in people with ALS.

4.2.6.3 Comparisons Between Participants with ALS and Neurotypical Controls

To quantitatively assess the apparent disruptions in ALS, we compared measures of within-session variability, specifically variances of the single-trial latency measures,

between groups. To do so, we calculated the participant mean values of the session variances in single-trial latencies for the N100, P200, N200, P300, and classifier-based latency estimates. Participant mean values were then compared between groups using Mann-Whitney U tests (Mann & Whitney, 1947), appropriate for non-normal distributions and for small and potentially uneven samples (Siegel & Tukey, 1960).

Linear regression results were assessed for differences between groups using the meta-analytic models from before, but combining the sessions recorded from both the group of neurotypical participants and the group of participants with ALS, then testing whether group was a significant (p<0.05) moderator using the Knapp and Hartung method (Knapp & Hartung, 2003; Viechtbauer, López-López, Sánchez-Meca, & Marín-Martínez, 2015). The spearman correlations between absolute single-trial shifts and amplitudes for each ERP feature (N100, P200, N200, and P300) were similarly compared between groups using a mixed-effects model testing the significance of the group as a moderator.

For binary session measures, specifically the significance or non-significance of linear regressions, variables in the linear regressions, and spearman correlations within sessions, groups were compared using Fisher's exact test (Fisher, 1922).

4.3 RESULTS

4.3.1 Understanding P300 BCI Latency Variations in a Neurotypical Population

The linear models of single-trial whole-epoch (CBLE) latency shifts on ERP feature latency shifts (N100, P200, N200, and P300 features extracted with the Woody filtering procedure) were significant for 37 of the 41 sessions recorded from

neurotypical participants. The results for the random effects models of the influence of single-trial N100, P200, N200, and P300 latencies on the whole epoch classifier-based latency estimates, as well as the number of sessions for which the within-session models were significant for each variable, are shown in Table 4.2. On average, single-trial latencies for the N100, P200, N200, and P300 were all significant (p<0.05) contributors to whole-epoch latency shifts calculated with CBLE in neurotypical participants.

The linear models of absolute single-trial whole epoch latency shifts on single-trial amplitudes were significant for 23 of the 41 sessions recorded from neurotypical participants. The results for the random effects models of the influence of single-trial N100, P200, N200, and P300 amplitudes on absolute whole-epoch latency shifts, as well as the number of sessions for which the within-session models were significant for each variable, are shown in Table 4.3. On average, trials with larger N100 and/or smaller P300 amplitudes had larger whole epoch shifts as calculated with CBLE (p<0.05), while the association between smaller N200 amplitudes and larger latency shifts did not reach significance (p=0.064) and there was no effect of P200 amplitude in neurotypical participants.

Table 4.2. Random effects models evaluating the effects of single-trial ERP latencies on classifier-based latency estimates in neurotypical participants

	Test for heterogeneity			***Model statistics***			***Frequency of significance in single sessions***
	I2 (%)	Q	*p*	Partial correlation, r_p	95% confidence interval	*p*	
N100	0.00	34.13	0.731	0.02	(0.01, 0.04)	0.010	0/41
P200	82.85	275.24	<0.001	0.14	(0.10, 0.19)	<0.001	19/41
N200	86.33	332.24	<0.001	0.20	(0.15, 0.24)	<0.001	29/41
P300	82.26	272.98	<0.001	0.15	(0.11, 0.19)	<0.001	25/41
						Model	37/41

Table 4.3. Random effects models evaluating the effects of single-trial ERP amplitudes on classifier-based latency estimates in neurotypical participants

	Test for heterogeneity			***Model statistics***			***Frequency of significance in single sessions***
	I2 (%)	Q	*p*	Partial correlation, r_p	95% confidence interval	*p*	
N100	62.79	108.35	<0.001	0.04	(0.01, 0.07)	0.014	13/41
P200	81.26	220.84	<0.001	-0.01	(-0.05, 0.03)	0.544	18/41
N200	76.97	179.39	<0.001	-0.04	(-0.07, 0.00)	0.064	14/41
P300	65.21	116.15	<0.001	-0.04	(-0.07, -0.01)	0.015	12/41
						Model	23/41

The linear models of single-trial P300 latencies on N100, P200, and N200 latencies were significant for 17 of the 41 sessions recorded from neurotypical participants. The results for the random effects models of the associations between single-trial N100, P200, and N200 latencies and single trial P300 latencies, as well as the number of sessions for which the within-session models were significant for each variable, are shown in Table 4.4. On average, single-trial latencies for the P200 and N200 were significantly (p<0.05) associated with single-trial P300 latencies, but single-trial N100 latencies were not. The linear models of single-trial N200 latencies on N100 and P200 latencies were significant for 13 of the 41 sessions recorded from neurotypical participants. The results for the random effects models of the influence of single-trial N100 and P200 latencies on N200 latencies are also shown in Table 4.4, indicating that single-trial P200 latencies (p<0.001), but not N100 latencies (p>0.05), are significantly associated with single-trial N200 latencies. Also shown in Table 4.4, single-trial N100 latencies also were not, on average, significantly associated with single-trial P200 latencies (p>0.05).

Table 4.4. Random effects models evaluating the associations between earlier and later single-trial ERP latencies in neurotypical participants

	Test for heterogeneity			***Model statistics***			***Frequency of significance in single sessions***
	I2 (%)	Q	*p*	Partial correlation, r_p	95% confidence interval	*p*	
Associations between single-trial P300 latencies and earlier single-trial ERP latencies							
N100	26.17	53.72	0.072	-0.01	(-0.03, 0.02)	0.644	2/41
P200	56.71	93.06	<0.001	0.03	(0.00, 0.06)	0.034	8/41
N200	81.74	246.40	<0.001	0.08	(0.04, 0.12)	<0.001	15/41
						Model	17/41
Associations between single-trial N200 latencies and earlier single-trial ERP latencies							
N100	9.52	43.79	0.314	0.01	(-0.01, 0.03)	0.376	3/41
P200	90.33	517.01	<0.001	0.12	(0.07, 0.18)	<0.001	14/41
						Model	13/41
Associations between single-trial P200 latencies and earlier single-trial ERP latencies							
N100	38.10	64.36	0.009	0.01	(-0.01, 0.03)	0.412	6/41

On average, within-session spearman correlations between single-trial N100 latencies and amplitudes, between single-trial P200 latencies and amplitudes, between N200 single-trial latencies and amplitudes, and between P300 single-trial latencies and amplitudes were all significant (p<0.05), with smaller responses associated with increased latency shifts. Results from the random-effects models combining these results across sessions are in Table 4.5, while the correlation plots between single-trial amplitudes and absolute latency shifts for a single session recorded from a neurotypical participant are shown in Figure 4.5.

Table 4.5. Random effects models evaluating spearman correlations between single-trial amplitudes and latencies within ERP features over sessions recorded from neurotypical participants.

	Test for heterogeneity			***Model statistics***			***Frequency of significance in single sessions***
	I2 (%)	Q	*p*	Correlation, *r*	95% confidence interval	*p*	
N100	70.61	135.82	<0.001	-0.24	(-0.28, -0.21)	<0.001	36/41
P200	64.09	110.99	<0.001	-0.25	(-0.28, -0.22)	<0.001	38/41
N200	65.91	117.13	<0.001	-0.30	(-0.34, -0.27)	<0.001	39/41
P300	74.92	159.80	<0.001	-0.30	(-0.33, -0.26)	<0.001	39/41

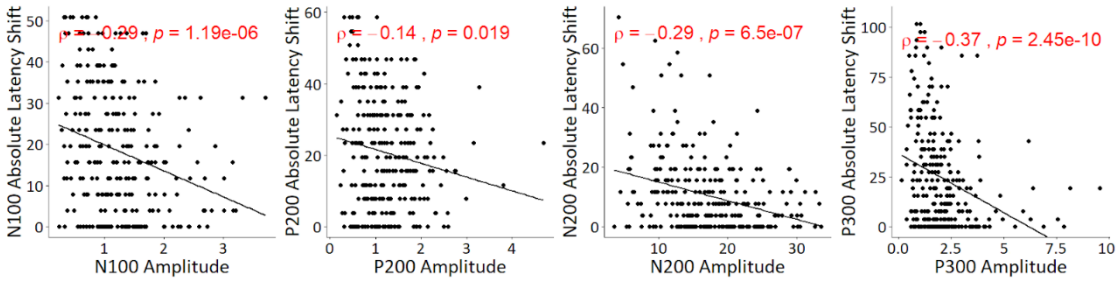

Figure 4.5. Single-trial absolute latency shifts versus single-trial amplitudes for the N100 (left), P200 (center left), N200 (center right), and P300 (right) from a single session recorded from a neurotypical participant.

Whole-epoch latency jitter, or the variance in the single-trial shifts calculated with CBLE (vCBLE), was significantly dependent on P200 jitter (p=0.038) but not on jitter in any of the other ERP components, for neurotypical participants, with results shown in Table 4.6. However, a stepwise regression using the Akaike Information Criterion (Venables & Ripley, 2002) included both P200 jitter (β=0.39, p=0.012) and P300 jitter (β=0.21, p=0.156) as predictors of vCBLE.

The model of BCI performance on individual component jitters in neurotypical participants was significant (p=0.028). However, possibly due to consistently high character selection accuracies in neurotypical participants, no single component jitter was a significant contributor to performance (p>0.05), as shown in Table 4.7. A stepwise regression using the Akaike Information Criterion retained N200 jitter (β=-0.32, p=0.0407) as a predictor of BCI performance.

Table 4.6. Linear Regression of whole-epoch jitter (vCBLE) on component jitter over sessions recorded from neurotypical participants.

	Unstandardized coefficients		***Standardized coefficients***	***t***	***p***
	B	Standard Error	β		
(Constant)	87.74	245.28		0.36	0.723
N100 Latency Jitter	-0.04	0.31	-0.02	-0.14	0.890
P200 Latency Jitter	0.36	0.17	0.34	2.16	0.038
N200 Latency Jitter	0.19	0.15	0.19	1.22	0.230
P300 Latency Jitter	0.08	0.06	0.18	1.22	0.230
Multiple R^2	***Adjusted R^2***	***Residual Standard Error***	***F***	***df***	***p***
0.26	0.18	260	3.14	(4,36)	0.026

Table 4.7. Linear Regression of performance (character selection accuracy) on component jitter over sessions recorded from neurotypical participants

	Unstandardized coefficients		***Standardized coefficients***	***t***	***p***
	B	Standard Error	β		
(Constant)	1.04	0.03		34.03	<0.001
N100 Latency Jitter	-2.81·10-5	3.85·10-5	-0.12	-0.73	0.471
P200 Latency Jitter	1.50·10-5	2.07·10-5	0.12	0.72	0.475
N200 Latency Jitter	-3.53·10-5	1.89·10-5	-0.31	-1.87	0.070
P300 Latency Jitter	-3.88·10-6	7.91·10-6	-0.08	-0.49	0.627
Multiple R^2	***Adjusted R^2***	***Residual Standard Error***	***F***	***df***	***p***
0.13	0.03	0.03	1.29	(4,36)	0.028

4.3.2 A Longitudinal Analysis of BCI Performance and Intra-Session Variability in People with ALS

The linear models of single-trial whole-epoch (CBLE) latency shifts on single-trial ERP feature latency shifts (N100, P200, N200, and P300 features extracted with the Woody filtering procedure) were significant for 29 of the 57 sessions recorded from participants with ALS. The results for the random effects models of the influence of single-trial N100, P200, N200, and P300 latencies on the whole epoch classifier-based latency estimates, as well as the number of sessions for which the within-session models were significant for each variable, are shown in Table 4.8. On average, single-trial P200,

N200, and P300 latencies contributed significantly (p<0.05) to the single-trial whole-epoch latency shifts calculated with CBLE for participants with ALS, but single-trial N100 latencies did not.

Table 4.8. Random effects models evaluating the effects of single-trial ERP latencies on classifier-based latency estimates in participants with ALS.

	Test for heterogeneity			***Model statistics***			***Frequency of significance in single sessions***
	I2 (%)	Q	*p*	Partial correlation, r_p	95% confidence interval	*p*	
N100	0.00	45.59	0.838	0.01	(-0.01, 0.02)	0.376	3/57
P200	62.71	152.68	<0.001	0.06	(0.04, 0.09)	<0.001	14/57
N200	61.70	148.61	<0.001	0.09	(0.06, 0.11)	<0.001	21/57
P300	70.85	198.40	<0.001	0.08	(0.05, 0.11)	<0.001	19/57
						Model	29/57

In addition, the linear models of absolute single-trial whole epoch latency shifts on single-trial amplitudes were significant for 20 of the 57 sessions recorded from participants with ALS. The results for the random effects models of the influence of single-trial N100, P200, N200, and P300 amplitudes on absolute whole-epoch latency shifts, as well as the number of sessions for which the within-session models were significant for each variable, are shown in Table 4.9. On average, single-trial amplitudes for the N200 and P300 were significantly (p<0.05) negatively associated with the single-trial whole-epoch latency shifts calculated with CBLE for participants with ALS, indicating that trials with smaller N200 and/or P300 amplitudes tended to show increased whole-epoch latency shifts. However, N100 and P200 amplitudes did not significantly correlate with the whole-epoch shifts in people with ALS.

The linear models of single-trial P300 latencies on N100, P200, and N200 latencies were significant for 14 of the 57 sessions recorded from participants with ALS. The results for the random effects models of associations between single-trial N100, P200,

and N200 latencies and single trial P300 latencies, as well as the number of sessions for which the within-session models were significant for each variable, are shown in Table 4.10. On average, single-trial latencies for the N200 were significantly (p<0.05) associated with single-trial P300 latencies, but single-trial N100 and P200 latencies were not.

Table 4.9. Random effects models evaluating the effects of single-trial ERP amplitudes on classifier-based latency estimates in participants with ALS.

	Test for heterogeneity			***Model statistics***			***Frequency of significance in single sessions***
	I2 (%)	Q	*p*	Partial correlation, r_p	95% confidence interval	*p*	
N100	15.98	66.09	0.168	0.01	(0.00, 0.03)	0.179	4/57
P200	28.06	77.16	0.032	-0.01	(-0.03, 0.00)	0.208	4/57
N200	59.51	138.93	<0.001	-0.03	(-0.06, 0.01)	0.009	14/57
P300	57.73	133.00	<0.001	-0.05	(-0.08, -0.03)	<0.001	18/57
						Model	20/57

The linear models of single-trial N200 latencies on N100 and P200 latencies were significant for 8 of the 57 sessions recorded from participants with ALS. The results for the random effects models of associations between single-trial N100 and P200 latencies on N200 latencies are shown in Table 4.10, indicating that single-trial P200 latencies (p=0.003), but not N100 latencies (p>0.05), are significantly associated with single-trial N200 latencies. As shown in Table 4.10, single-trial N100 latencies also were not, on average, significantly associated with single-trial P200 latencies (p>0.05).

On average, within-session spearman correlations between single-trial N100 latencies and amplitudes, between single-trial P200 latencies and amplitudes, between N200 single-trial latencies and amplitudes, and between P300 single-trial latencies and amplitudes were all significant (p<0.05), with smaller responses associated with increased latency shifts. Results from the random-effects models combining these

results across sessions are in Table 4.11, while the correlation plots between single-trial amplitudes and absolute latency shifts for a single session recorded from a participant with ALS are shown in Figure 4.6.

Table 4.10. Random effects models associations between earlier and later single-trial ERP latencies in participants with ALS.

	Test for heterogeneity			***Model statistics***			***Frequency of significance in single sessions***
	I2 (%)	Q	*p*	Partial correlation, r_p	95% confidence interval	*p*	
Associations between single-trial P300 latencies and earlier single-trial ERP latencies							
N100	8.31	60.42	0.319	0.00	(-0.02, 0.01)	0.774	3/57
P200	5.50	58.04	0.400	0.00	(-0.02, 0.01)	0.886	2/57
N200	72.55	208.43	<0.001	0.04	(0.01, 0.07)	0.006	18/57
						Model	14/57
Associations between single-trial N200 latencies and earlier single-trial ERP latencies							
N100	0.00	52.11	0.623	0.00	(-0.02, 0.01)	0.666	1/57
P200	70.45	204.86	<0.001	0.04	(0.01, 0.07)	0.003	10/57
						Model	8/57
Associations between single-trial P200 latencies and earlier single-trial ERP latencies							
N100	23.25	72.14	0.072	0.01	(-0.01, 0.03)	0.353	6/57

Table 4.11. Random effects models evaluating spearman correlations between single-trial amplitudes and latencies within ERP features over sessions recorded from participants with ALS.

	Test for heterogeneity			***Model statistics***			***Frequency of significance in single sessions***
	I2 (%)	Q	*p*	Correlation, *r*	95% confidence interval	*p*	
N100	80.44	284.20	<0.001	-0.23	(-0.27, -0.19)	<0.001	43/57
P200	75.79	228.67	<0.001	-0.26	(-0.29, -0.23)	<0.001	48/57
N200	70.14	186.23	<0.001	-0.29	(-0.32, -0.27)	<0.001	54/57
P300	72.22	200.67	<0.001	-0.27	(-0.30, -0.24)	<0.001	52/57

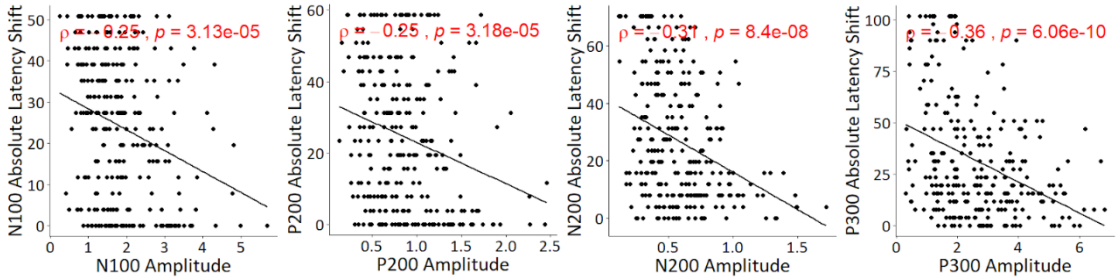

Figure 4.6. Single-trial absolute latency shifts versus single-trial amplitudes for the N100 (left), P200 (center left), N200 (center right), and P300 (right) from a single session recorded from a participant with ALS.

Whole-epoch latency jitter, or the variance in the single-trial shifts calculated with CBLE (vCBLE), was significantly dependent on P200 jitter (p=0.027) but not on jitter in any of the other ERP components, for participants with ALS, with results shown in Table 4.12. However, a stepwise regression using the Akaike Information Criterion retains P200 jitter (β=0.26, p=0.038), N200 jitter, (β=0.25, p=0.059), and P300 jitter (β=0.25, p=0.050) as predictors of whole epoch jitter (vCBLE).

The model of BCI performance on individual component jitters in participants with ALS was significant (p<0.001). P200 and N200 jitter were significantly associated with performance (p<0.05), but not N100 or P300 jitter, as shown in Table 4.13. However, a stepwise regression using the Akaike Information Criterion retained all four component jitters in its model of BCI performance.

Variance in the whole-epoch latency shifts, or vCBLE, increased significantly over time in participants with ALS (rrm=0.49, p<0.001). However, the variances of N100, P200, N200, and P300 latencies *did not* significantly increase over time in participants with ALS (p>0.05). The longitudinal trends in whole-epoch and ERP feature latency jitters are shown in Figure 4.7.

Table 4.12. Linear Regression of whole-epoch jitter (vCBLE) on component jitter over sessions recorded from participants with ALS.

	Unstandardized coefficients		***Standardized coefficients***	***t***	***p***
	B	Standard Error	β		
(Constant)	-1337.12	1294.86		-1.033	0.307
N100 Latency Jitter	-1.41	1.37	-0.02	-1.03	0.309
P200 Latency Jitter	1.76	0.77	0.34	2.28	0.027
N200 Latency Jitter	0.95	0.51	0.19	1.85	0.070
P300 Latency Jitter	0.58	0.30	0.18	1.96	0.055
Multiple R^2	***Adjusted R^2***	***Residual Standard Error***	***F***	***df***	***p***
0.28	0.22	1024	5.06	(4,52)	0.002

Table 4.13. Linear Regression of performance (character selection accuracy) on component jitter over sessions recorded participants with ALS.

	Unstandardized coefficients		***Standardized coefficients***	***t***	***p***
	B	Standard Error	β		
(Constant)	1.50	0.35		4.33	<0.001
N100 Latency Jitter	5.23·10-4	3.67·10-4	0.17	1.43	0.160
P200 Latency Jitter	-5.07·10-4	2.07·10-4	-0.30	-2.45	0.018
N200 Latency Jitter	-2.90·10-4	1.37·10-4	-0.36	-2.11	0.039
P300 Latency Jitter	-1.57·10-4	7.90·10-5	-0.24	-1.98	0.053
Multiple R^2	***Adjusted R^2***	***Residual Standard Error***	***F***	***df***	***p***
0.32	0.27	0.27	6.09	(4,52)	<0.001

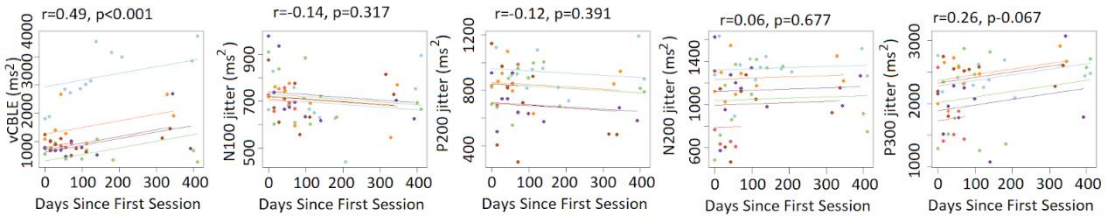

Figure 4.7. Longitudinal trends in latency jitter over time, from left to right for whole-epoch jitter (vCBLE), N100 jitter, P200 jitter, N200 jitter, and P300 jitter.

Neither overall cognitive scores nor attention subscores significantly correlated with latency jitter, whether for the whole epoch (vCBLE) or for specific ERP features (p>0.05). Repeated measures correlations between BCI performance and cognitive scores were also not significant (p>0.05).

Repeated measures correlations between session average ERP features and whole-epoch latency jitter, as measured by vCBLE, are shown in Figure 4.8. After corrections for multiple comparisons, only the correlations between N200 amplitude and vCBLE were significant, at channels Cz (r_{rm}=0.40, p=0.023), PO8 (r_{rm}=0.34, p=0.038), and Oz (r_{rm}=0.37, p=0.026). Correlations between session-average amplitudes or latencies and vCBLE were not significant for any other features or channels.

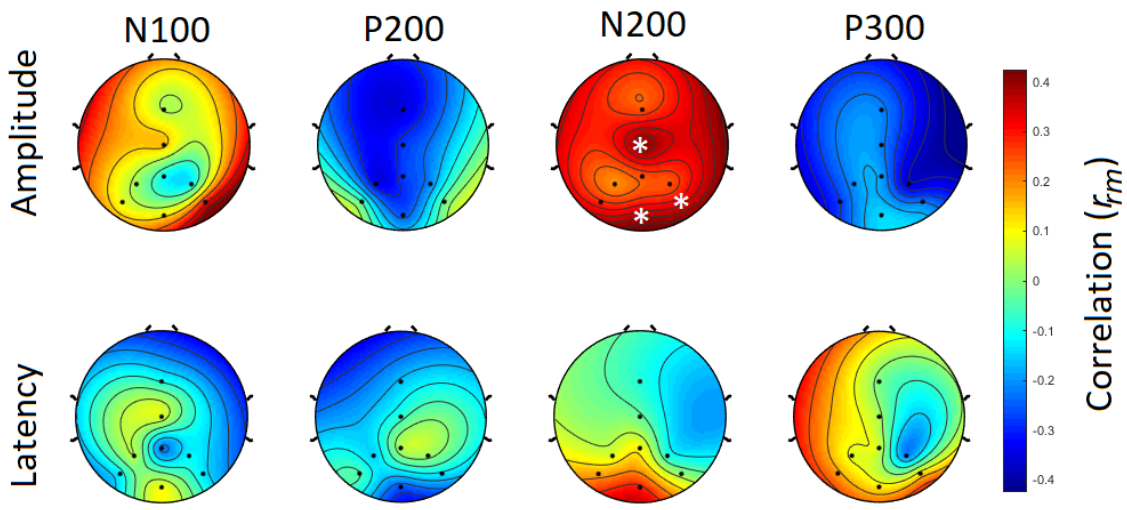

Figure 4.8. Repeated measures correlations between whole epoch latency jitter (vCBLE) and N100, P200, N200, and P300 amplitudes (top row) or latencies (bottom row) in participants with amyotrophic lateral sclerosis. Shades of red and orange indicate positive correlations between ERP amplitudes or latencies and latency jitter, whereas shades of blue indicate negative correlations. Asterisks indicate channels where this correlation is significant.

4.3.3 Comparisons Between Participants with ALS and Neurotypical Controls

Figure 4.9 compares participant average ERP feature and whole-epoch latency jitters between groups. Whole-epoch jitter (vCBLE) was significantly increased in people with ALS as compared to neurotypical controls, as were P200, N200, and P300 jitter (p<0.05). However, N100 jitter did not differ significantly between groups (p>0.05). Whole-epoch jitter was 1362 ± 993 ms^2 in participants with ALS, as compared to 521 ± 206 ms^2 in neurotypical participants (p=0.004). N100 jitter was 708 ± 12 ms^2 in participants with ALS, which did not differ significantly from the 714 ± 124 ms^2 in neurotypical participants (p=0.693). P200 jitter was 804 ± 95 ms^2 in participants with ALS, significantly (p=0.027) greater than the 563 ± 222 ms^2 in neurotypical controls.

N200 jitter was 1094 ± 192 ms^2 in participants with ALS, as compared to 747 ± 164 ms^2 in neurotypical controls (p=0.004). Finally, P300 jitter was 2197 ± 284 ms^2 in participants with ALS, significantly (p=0.010) greater than the 1595 ± 477 ms^2 in neurotypical controls. Table 4.14 tabulates the individual jitter measures for each group.

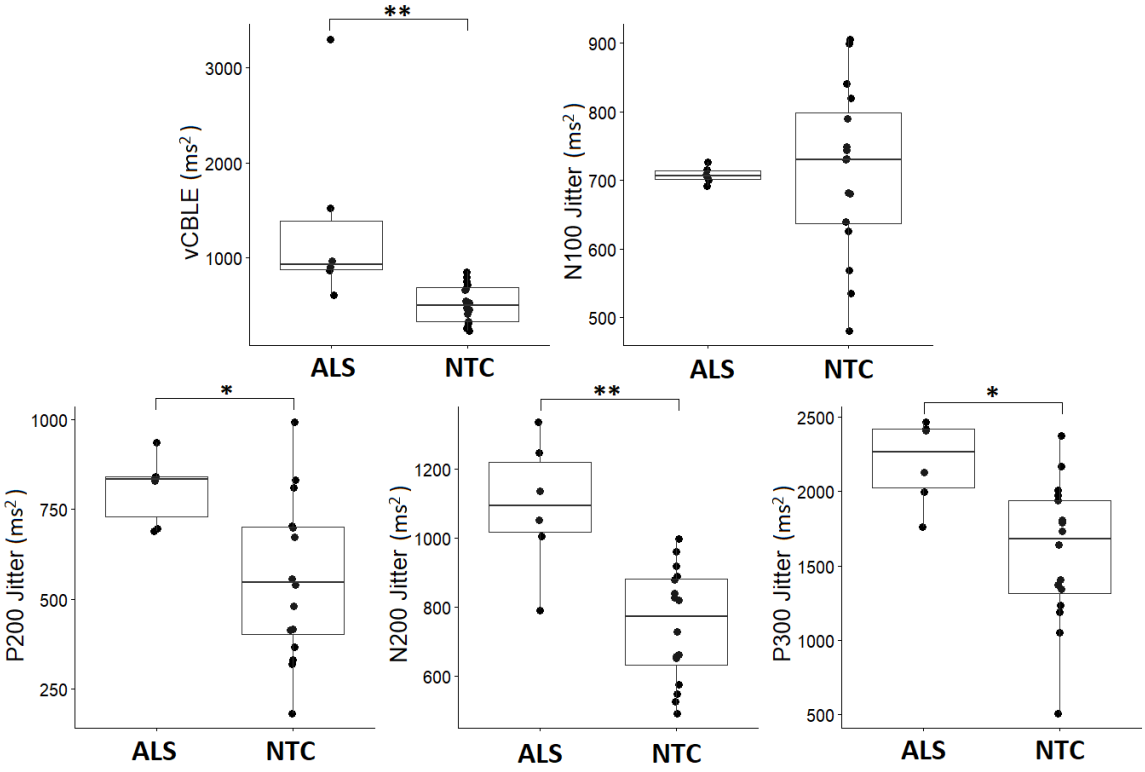

Figure 4.9. Box plots showing, from left to right, whole-epoch latency jitter (vCBLE), N100 jitter, P200 jitter, N200 jitter, and P300 jitter for all participants in both the amyotrophic lateral sclerosis (ALS) and neurotypical control (NTC) groups. The boxes show the quartiles with the median represented by a bold line through the box. Each dot shows the corresponding value for one participant (* significant at $p < 0.05$, ** significant at $p < 0.01$, Wilcoxon rank-sum test).

Individual performance measures are tabulated for each group in Table 4.15. As in our prior work using 9 of these 16 neurotypical participants and three sessions from each participant with ALS, all performance measures were significantly reduced in participants with ALS as compared to neurotypical controls (p<0.05).

Table 4.14. Jitter measures for participants with amyotrophic lateral sclerosis (ALS) and neurotypical controls (NTC), with group means and standard deviations (ms^2).

		vCBLE	*N100 Jitter*	*P200 Jitter*	*N200 Jitter*	*P300 Jitter*
ALS	ALS-1	3294	692	934	1335	2409
	ALS-2	610	716	829	1051	2130
	ALS-3	872	700	839	791	1761
	ALS-4	1525	706	840	1246	2465
	ALS-5	906	727	695	1135	1993
	ALS-6	965	709	689	1007	2422
ALS Mean±STD		**1362±993**	**708±12**	**804±95**	**1094±192**	**2197±284**
NTC	NTC-1	759	626	538	839	2372
	NTC-2	419	791	330	576	1228
	NTC-3	807	680	699	659	1369
	NTC-4	858	748	698	958	1635
	NTC-5	449	840	672	490	1794
	NTC-6	340	729	555	547	504
	NTC-7	316	731	417	995	1412
	NTC-8	280	682	480	880	1047
	NTC-9	478	743	809	654	1932
	NTC-10	527	819	703	827	2014
	NTC-11	662	899	993	886	1971
	NTC-12	259	480	181	525	1811
	NTC-13	234	906	318	650	1189
	NTC-14	718	569	365	726	1340
	NTC-15	681	640	414	819	2171
	NTC-16	551	535	831	918	1733
NTC Mean±STD		**521±206**	**714±124**	**563±222**	**747±164**	**1595±477**

Table 4.15. Performance measures for participants with amyotrophic lateral sclerosis (ALS) and neurotypical controls (NTC), with group means and standard deviations.

		Character Accuracy (%)	*Classification Accuracy (%)*	*Precision (%)*	*Recall (%)*	*F-score*
ALS	ALS-1	24.40	83.19	36.88	7.56	0.14
	ALS-2	98.81	92.36	80.03	71.49	0.75
	ALS-3	100.00	91.80	79.70	68.14	0.73
	ALS-4	73.02	85.72	60.51	36.35	0.45
	ALS-5	94.29	90.64	74.70	64.29	0.69
	ALS-6	98.41	88.97	71.26	56.27	0.63
ALS Mean±STD		**81.49±29.74**	**88.78±3.63**	**67.18±16.48**	**50.68±24.57**	**0.57±0.24**
NTC	NTC-1	100.00	90.40	75.32	62.98	0.69
	NTC-2	100.00	95.34	88.44	82.86	0.86
	NTC-3	100.00	92.66	82.64	70.60	0.76
	NTC-4	95.24	94.09	86.42	76.19	0.81
	NTC-5	100.00	95.74	91.11	82.50	0.87
	NTC-6	100.00	96.07	90.43	85.48	0.88
	NTC-7	100.00	96.51	93.97	84.40	0.89
	NTC-8	100.00	97.36	93.90	90.00	0.92
	NTC-9	100.00	96.59	92.61	86.43	0.89
	NTC-10	100.00	95.04	90.86	78.10	0.84
	NTC-11	100.00	91.90	79.47	69.29	0.74
	NTC-12	100.00	96.82	92.69	87.86	0.90
	NTC-13	96.43	96.43	91.58	86.43	0.89
	NTC-14	92.86	92.35	80.47	70.00	0.75
	NTC-15	100.00	92.65	81.00	72.50	0.76
	NTC-16	100.00	94.85	87.21	80.89	0.84
NTC Mean±STD		**99.03±2.18**	**94.67±2.09**	**87.38±5.86**	**79.16±8.02**	**0.83± 0.07**

The mixed-effects models investigating whether the effects of single-trial ERP latencies on whole-epoch CBLE shifts differed between groups were significant for the P200, N200, and P300 (p<0.05), but not for the N100. In particular, single-trial P200, N200, and P300 latencies contributed less strongly to CBLE shifts in people with ALS. This indicates that not only does CBLE reflect other factors in addition to P300 jitter, but that the relationship between classifier-based latency estimates and single-trial ERP feature latencies is likely disrupted in ALS. The mixed-effects models investigating group as a moderator of the relationship between single-trial feature latencies and whole-epoch shifts are shown in Table 4.16.

The disruption of relationships between whole-epoch shifts estimated with CBLE and single-trial feature latencies is also apparent in the portion of sessions recorded from each group in the models investigating these relationships were significant, shown in Tables 4.2 and 4.8. Notably, models of whole epoch latency shifts (CBLE) were significant for 90.2% of sessions recorded from neurotypical participants, but only 50.9% of sessions recorded from participants with ALS (p<0.001), indicating that CBLE estimates are less dependent on latency shifts and more influenced by other factors, such as reduced amplitudes or increased neural noise, in people with ALS than in neurotypical controls. P200 latencies were significantly (p=0.038) more likely to be contributors to whole-epoch latency shifts in neurotypical controls (46.3%) than in people with ALS (24.6%). N200 latencies were also more likely to be significant contributors to CBLE in neurotypical participants (70.7%) than in people with ALS (41.1%, p=0.003), as were P300 latencies, which contributed significantly to CBLE shifts in 61.0% of sessions recorded from neurotypical participants but only 33.3% of

sessions recorded from participants with ALS (p=0.013). However, the portion of sessions in which N100 latencies contributed to CBLE shifts were not significantly different between neurotypical participants (0.0%) and participants with ALS (5.3%, p=0.262).

The mixed-effects models investigating whether the effects of single-trial ERP amplitudes on whole-epoch CBLE shifts differed between groups were not significant for any ERP feature. These results are shown in Table 4.17. However, as reported in Tables 4.3 and 4.9, single-trial N100 (p=0.006) and P200 (p=0.001) amplitudes were more likely to be significantly associated with CBLE shifts for sessions recorded from neurotypical participants than from participants with ALS -- there may be some disruption in associations between single-trial amplitudes and CBLE shifts in people with ALS, but this is not conclusive.

Table 4.16. Mixed effects model evaluating differences in the effects of single-trial ERP latencies on classifier-based latency estimates between neurotypical participants and participants with ALS.

	Test for heterogeneity				***Effect of Group***		
	I2 (%)	Q	*p*	R2 (%)	Difference	95% confidence interval	*p*
N100	0.00	79.72	0.884	0.00	0.02	(0.00, 0.04)	0.132
P200	75.06	427.91	<0.001	12.99	0.08	(0.04, 0.13)	<0.001
N200	77.96	480.85	<0.001	19.82	0.11	(0.06, 0.16)	<0.001
P300	77.07	471.38	<0.001	8.61	0.07	(0.02, 0.12)	0.005

Table 4.17. Mixed effects model evaluating differences in the effects of single-trial ERP amplitudes on classifier-based latency estimates between neurotypical participants and participants with ALS.

	Test for heterogeneity				***Effect of Group***		
	I2 (%)	Q	*p*	R2 (%)	Difference	95% confidence interval	*p*
N100	45.09	174.43	<0.001	3.67	0.03	(-0.01, 0.06)	0.111
P200	67.15	298.99	<0.001	0.00	0.00	(-0.04, 0.04)	0.951
N200	69.29	318.32	<0.001	0.00	0.00	(-0.05, 0.04)	0.875
P300	61.21	249.15	<0.001	0.00	0.02	(-0.02, 0.05)	0.427

Disruptions were also observed in associations between earlier and later single-trial ERP latencies, as shown in Table 4.18. Mixed-effects models investigating group as a moderator of the relationships between single-trial P200 latencies and single-trial N200 latencies (p=0.006) and between single-trial P200 latencies and single trial P300 latencies (p=0.041) were both significant, indicating disrupted assocciations between single-trial P200 latencies and single-trial latencies for later ERP features in people with ALS. There were not significant disruptions (p>0.05) in the associations between single-trial N100 or N200 latencies and single-trial latencies for later ERP features. The portions of sessions with significant relationships between single-trial latencies for earlier ERP features and single-trial latencies for later ERP features, found in Tables 4.4 and 4.10, were not significant for any pairs of ERP features.

Table 4.18. Mixed effects model evaluating differences in associations between earlier and later single-trial ERP latencies between neurotypical participants and participants with ALS.

	Test for heterogeneity				***Effect of Group***		
	I2 (%)	Q	*p*	R2 (%)	Difference	95% confidence interval	*p*
Differences in associations between single-trial P300 latencies and earlier single-trial ERP latencies							
N100	16.75	114.13	0.100	0.00	0.00	(-0.03, 0.02)	0.841
P200	37.08	151.10	<0.001	8.37	0.03	(0.00, 0.06)	0.041
N200	77.37	454.83	<0.001	1.53	0.03	(-0.01, 0.09)	0.151
Differences in associations between single-trial N200 latencies and earlier single-trial ERP latencies							
N100	0.84	95.90	0.484	2.33	0.01	(-0.01, 0.04)	0.323
P200	83.88	721.87	<0.001	7.77	0.08	(0.02, 0.14)	0.006
Differences in associations between single-trial P200 latencies and earlier single-trial ERP latencies							
N100	30.20	136.50	0.042	0.00	0.00	(-0.03, 0.03)	0.927

Within-session spearman correlation coefficients between the amplitudes and latencies of single trial features (e.g., N100 amplitude with N100 latency) did not differ significantly between groups (p>0.05). Results from the mixed-effects models

comparing these within-session correlations between participants with ALS and neurotypical participants are shown in Table 4.19

Table 4.19. Mixed effects model evaluating differences between participants with ALS and neurotypical participants in the spearman correlations between single-trial amplitudes and latencies within ERP features.

	Test for heterogeneity				***Effect of Group***		
	I2 (%)	Q	*p*	R2 (%)	Difference	95% confidence interval	*p*
N100	77.26	420.03	<0.001	0.00	-0.01	(-0.07, 0.04)	0.605
P200	71.98	339.66	<0.001	0.00	0.01	(-0.04, 0.06)	0.720
N200	68.50	303.37	<0.001	0.00	-0.01	(-0.05, 0.04)	0.686
P300	73.42	360.47	<0.001	0.48	-0.03	(-0.08, 0.02)	0.244

4.4 DISCUSSION

In this study, we investigated trial-to-trial variability in the N100, P200, N200, and P300, along with whole-epoch latency shifts calculated with classifier-based latency estimation (CBLE), in both neurotypical participants and participants with ALS. In doing so, we determined which ERP features contribute to classifier-based latency estimates in both groups, as well as examining the effects of single-trial latencies from earlier ERP features on the single-trial latencies of later ERP features. We also determined which component jitters are relevant to BCI performance. Finally, we longitudinally investigated N100, P200, N200, and P300 jitter in people with ALS.

We found that latency variability was increased in participants with ALS as compared to neurotypical controls for the P200, N200, and P300. We interpreted the increased whole-epoch latency jitter (vCBLE) in our prior work as indicating increased P300 latency jitter (Zisk et al., 2020). In addition to the connection between stimulus evaluation time and single-trial P300 latencies (Kelly & O'Connell, 2013; Verleger, 1997), single-trial latencies of both the P200 and N200 also appear to be connected to

perceptual decision-making (Nunez, Vandekerckhove, & Srinivasan, 2017). Our study, then, indicates increased intra-individual variability in people with ALS for several attention-related components, including but not limited to the P300.

Our investigation of the single-trial features reflected in classifier-based latency estimates revealed that single-trial latency shifts in all four ERP features we investigated, specifically N100, P200, N200, and P300, were reflected in the CBLE shifts in neurotypical participants. That is, CBLE significantly reflects P300 latency jitter, as previously shown by Thompson and colleagues (Thompson et al., October 23, 2019), and it also reflects latency jitter in other ERP features as would be expected from a method considering shifts of the entire ERP complex (Mowla et al., 2020; Thompson et al., 2012). However, the correlations between whole-epoch latency shifts and P200, N200, and P300 latency shifts were all disrupted in participants with ALS, and the correlation between N100 latency shifts and whole-epoch latency shifts did not reach significance in participants with ALS. There are multiple possible explanations for these disruptions.

First, smaller single-trial ERP amplitudes were significantly associated with greater detected latency shifts, both within ERP features (e.g., smaller N100 amplitudes on trials with greater detected N100 latency shifts) and between ERP feature amplitudes and whole-epoch latency shifts (e.g., smaller P300 amplitudes in trials where CBLE found a larger latency shift). These relationships between increased single-trial latency shifts and decreased single-trial amplitudes were present in both groups and were not disrupted in participants with ALS. These relationships are consistent with the fact that single-trial ERP latency detection is subject to greater error as the signal-to-noise ratio

decreases (Michalewski et al., 1986), which could occur from reduced ERP amplitudes commonly reported in ALS (Raggi, Iannaccone, & Cappa, 2010; Riccio et al., 2013; Vieregge, Wauschkuhn, Heberlein, Hagenah, & Verleger, 1999), including in our work with these same participants (Zisk et al., 2020). Reduced signal-to-noise ratios have previously been reported as a concern for BCI performance among participants with ALS (Geronimo, A., Simmons, & Schiff, 2016), and could be relevant to the disruption of associations between single-trial whole-epoch latency shifts and single-trial N100, P200, N200, and P300 latencies.

Second, associations between single-trial P300 latencies and reaction times are known to be disrupted under certain conditions, including tasks where the focus is on speed (Verleger, 1997) and in people who have increased reaction time variability (Saville et al., 2011). While the single-trial latencies calculated in the current study are not reaction times, as P300 BCIs are designed for use by people with no voluntary motor control, similar factors could be involved. Whole-epoch, P200, N200, and P300 latency jitter were all increased in participants with ALS as compared to neurotypical controls, similar to the case of disrupted associations in people with increased intra-individual variability. The 93.75 ms stimulus with 62.5 ms breaks between stimuli used in this studies P300 speller paradigm additionally requires participants to evaluate 6.4 stimuli per second, with a target stimulus approximately every second. While fairly typical for P300 speller applications, this is much faster than would be expected in typical cognitive studies of the P300 (Barry et al., 2020; Verleger, Baur, Metzner, & Śmigasiewicz, 2014; Vieregge et al., 1999).

Considering the effects of earlier ERP features on later features, our current study found that both single-trial P200 and N200 latencies were significant contributors to single-trial P300 latencies in neurotypical participants, but N100 latencies did not contribute significantly to P300 latencies. This aligns with prior work examining the relationships between single-trial features with an auditory stimulus (Michalewski et al., 1986), though the relationships found in our work were comparatively weaker, possibly due to more rapid stimulus presentation with overlap between epochs. In participants with ALS, single-trial N200 latencies still contributed significantly to P300 latencies, but P200 latencies did not. The relationship between single-trial P200 and N200 latencies was also disrupted in participants with ALS, though this relationship was still present in both groups. As interrelationships between single-trial ERP features are not often investigated, it is not clear what might cause this disruption. However, as these disruptions involved the same ERP features for which latency jitter was increased in ALS, the disruptions may again be a result of relatively increased neural noise.

Our longitudinal analysis in participants with ALS did not find N100, P200, N200, or P300 jitter to increase over time, though there was a non-significant positive trend in P300 jitter over time. This is of interest, as whole-epoch latency jitter measured CBLE did increase significantly over time, both in our current analysis and with classifiers trained on data from prior sessions (Zisk et al., 2021). It would be interesting to know whether this pattern holds in an independent longitudinal sample, and if so, what contributor to whole-epoch jitter is increasing over time.

In our prior work, whole-epoch jitter did not significantly correlate with session-average ERP amplitudes in people with ALS, while there were significant correlations

for neurotypical control participants (Zisk et al., 2020). As there were fewer participants with ALS and thus fewer sessions recorded from participants with ALS, we considered the possibility that this might be a result of there being fewer sessions available. However, these correlations were again not significant over the current longitudinal recordings. This indicates that, like in other conditions where both increased jitter and reduced session-average amplitudes appear, jitter likely does not explain the amplitude differences (Ford et al., 1994; Saville et al., 2015; Unsal & Segalowitz, 1995; Walhovd, Rosquist, & Fjell, 2008).

Single-session cognitive scores did not correlate significantly with BCI performance or with any measure of latency variability, aligning with prior results in which clinical scores are not correlated with BCI performance in people with ALS (McCane et al., 2015; Zisk et al., 2020)

Overall, the current study complements prior work on within-session variability in BCI use by examining several ERP features in concert with CBLE, a strong predictor of BCI performance. It also investigates these measures of variability longitudinally in people with ALS, supporting a better understanding of single-trial ERP features, their interrelationships, and their relevance to BCI performance in ALS. We have additionally worked to address a theoretical weakness of vCBLE as a measure of latency variability, showing that it reflects latency jitter in several ERP features including but not limited to the P300.

One limitation of the current study is that despite the longitudinal recordings, the number of participants with ALS is still low. Another limitation is that with spatial PCA and Woody templates determined separately for each session to allow for potential

changes in waveforms over time in participants with ALS, our single-trial amplitude measures are not well-suited to analyses combining information from multiple sessions. The use of a single set of spatial factors and Woody templates for each participant, or even for each group, could address this limitation, but with an increased risk of selecting spatial factors and templates that may not be appropriate for all participants.

REFERENCES

Allison, B. Z., Kübler, A., & Jin, J. (2020). 30 years of P300 brain–computer interfaces. *Psychophysiology, 57*(7), e13569.

Aloe, A. M. (2014). An empirical investigation of partial effect sizes in meta-analysis of correlational data. *The Journal of General Psychology, 141*(1), 47-64.

Aloe, A. M., & Thompson, C. G. (2013). The synthesis of partial effect sizes. *Journal of the Society for Social Work and Research, 4*(4), 390-405.

Aricò, P., Aloise, F., Schettini, F., Salinari, S., Mattia, D., & Cincotti, F. (2014). Influence of P300 latency jitter on event related potential-based brain–computer interface performance. *Journal of Neural Engineering, 11*(3), 035008.

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology, 8*, 456.

Barry, R. J., Steiner, G. Z., De Blasio, F. M., Fogarty, J. S., Karamacoska, D., & MacDonald, B. (2020). Components in the P300: Don't forget the novelty P3! *Psychophysiology, 57*(7), e13371.

Bonett, D. G., & Wright, T. A. (2000). Sample size requirements for estimating Pearson, Kendall, and Spearman correlations. *Psychometrika, 65*(1), 23-28.

Cedarbaum, J. M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., & Nakanishi, A. (1999). The ALSFRS-R: A revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS study group (phase III). *Journal of the Neurological Sciences, 169*(1-2), 13-21.

Chennu, S., Craston, P., Wyble, B., & Bowman, H. (2009). Attention increases the temporal precision of conscious perception: Verifying the neural-ST2 model. *PLoS Computational Biology, 5*(11), e1000576.

Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice* MIT press.

Coles, M. G., Gratton, G., Kramer, A. F., & Miller, G. A. (1986). Principles of signal acquisition and analysis. In M. G. Coles, E. Donchin & S. W. Porges (Eds.), *Psychophysiology: Systems, processes and applications* (pp. 183-221). New York: Guilford Press.

Dinstein, I., Heeger, D. J., & Behrmann, M. (2015). Neural variability: Friend or foe? *Trends in Cognitive Sciences, 19*(6), 322-328.

Espeseth, T., Endestad, T., Rootwelt, H., & Reinvang, I. (2007). Nicotine receptor gene CHRNA4 modulates early event-related potentials in auditory and visual oddball target detection tasks. *Neuroscience, 147*(4), 974-985.

Fabiani, M., Gratton, G., Karis, D., & Donchin, E. (1987). Definition, identification, and reliability of measurement of the P300 component of the event-related brain potential. *Advances in Psychophysiology, 2*(3), 1-78.

Fabiani, M., Karis, D., & Donchin, E. (1986). P300 and recall in an incidental memory paradigm. *Psychophysiology, 23*(3), 298-308.

Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology, 70*(6), 510-523.

Fisher, R. A. (1922). On the interpretation of χ 2 from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society, 85*(1), 87-94.

Fjell, A. M., Rosquist, H., & Walhovd, K. B. (2009). Instability in the latency of P3a/P3b brain potentials and cognitive function in aging. *Neurobiology of Aging, 30*(12), 2065-2079.

Ford, J. M., White, P., Lim, K. O., & Pfefferbaum, A. (1994). Schizophrenics have fewer and smaller P300s: A single-trial analysis. *Biological Psychiatry, 35*(2), 96-103.

Geronimo, A., Simmons, Z., & Schiff, S. J. (2016). Performance predictors of brain–computer interfaces in patients with amyotrophic lateral sclerosis. *Journal of Neural Engineering, 13*(2), 026002.

Geronimo, A. M., & Simmons, Z. (2017). The P300 'face' speller is resistant to cognitive decline in ALS. *Brain-Computer Interfaces, 4*(4), 225-235.

Hochberg, Y., & Benjamini, Y. (1990). More powerful procedures for multiple significance testing. *Statistics in Medicine, 9*(7), 811-818.

Hoffmann, U., Vesin, J., Ebrahimi, T., & Diserens, K. (2008). An efficient P300-based brain–computer interface for disabled subjects. *Journal of Neuroscience Methods, 167*(1), 115-125.

Huggins, J. E., Alcaide-Aguirre, R. E., & Hill, K. (2016). Effects of text generation on P300 brain-computer interface performance. *Brain-Computer Interfaces, 3*(2), 112-120.

Jaśkowski, P., & Verleger, R. (2000). An evaluation of methods for single-trial estimation of P3 latency. *Psychophysiology, 37*(2), 153-162.

Kelly, S. P., & O'Connell, R. G. (2013). Internal and external influences on the rate of sensory evidence accumulation in the human brain. *Journal of Neuroscience, 33*(50), 19434-19441.

Kimura, M., Katayama, J., Ohira, H., & Schröger, E. (2009). Visual mismatch negativity: New evidence from the equiprobable paradigm. *Psychophysiology, 46*(2), 402-409.

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine, 22*(17), 2693-2710.

Krusienski, D. J., Sellers, E. W., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2008). Toward enhanced P300 speller performance. *Journal of Neuroscience Methods, 167*(1), 15-21. doi: 10.1016/j.jneumeth.2007.07.017

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2006). A comparison of classification techniques for the P300 speller. *Journal of Neural Engineering, 3*(4), 299.

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science, 197*(4305), 792-795.

Magnuson, J. R., Iarocci, G., Doesburg, S. M., & Moreno, S. (2020). Increased Intra-Subject variability of reaction times and Single-Trial Event-Related potential components in children with autism spectrum disorder. *Autism Research, 13*(2), 221-229.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 50-60.

McCane, L. M., Heckman, S. M., McFarland, D. J., Townsend, G., Mak, J. N., Sellers, E. W., . . . Vaughan, T. M. (2015). P300-based brain-computer interface (BCI) event-related potentials (ERPs): People with amyotrophic lateral sclerosis (ALS) vs. age-matched controls. *Clinical Neurophysiology, 126*(11), 2124-2131.

McCane, L. M., Sellers, E. W., Mcfarland, D. J., Mak, J. N., Carmack, C. S., Zeitlin, D., . . . Vaughan, T. M. (2014). Brain-computer interface (BCI) evaluation in people with amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 15*(3-4), 207-215.

Michalewski, H. J., Prasher, D. K., & Starr, A. (1986). Latency variability and temporal interrelationships of the auditory event-related potentials (N1, P2, N2, and P3) in normal subjects. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 65*(1), 59-71.

Mowla, M. R., Gonzalez-Morales, J. D., Rico-Martinez, J., Ulichnie, D. A., & Thompson, D. E. (2020). A comparison of classification techniques to predict brain-computer interfaces accuracy using classifier-based latency estimation. *Brain Sciences, 10*(10), 734.

Nunez, M. D., Vandekerckhove, J., & Srinivasan, R. (2017). How attention influences perceptual decision making: Single-trial EEG correlates of drift-diffusion model parameters. *Journal of Mathematical Psychology, 76*, 117-130.

Ouyang, G., Hildebrandt, A., Sommer, W., & Zhou, C. (2017). Exploiting the intra-subject latency variability from single-trial event-related potentials in the P3 time range: A review and comparative evaluation of methods. *Neuroscience & Biobehavioral Reviews, 75*, 1-21.

Pal, M., & Bandyopadhyay, S. (2016). Many-objective feature selection for motor imagery EEG signals using differential evolution and support vector machine. Paper presented at the *2016 International Conference on Microelectronics, Computing and Communications (MicroCom)*, 1-6.

Patterson, J. V., Michalewski, H. J., & Starr, A. (1988). Latency variability of the components of auditory event-related potentials to infrequent stimuli in aging, Alzheimer-type dementia, and depression. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 71*(6), 450-460.

Raggi, A., Iannaccone, S., & Cappa, S. F. (2010). Event-related brain potentials in amyotrophic lateral sclerosis: A review of the international literature. *Amyotrophic Lateral Sclerosis: Official Publication of the World Federation of Neurology Research Group on Motor Neuron Diseases, 11*(1-2), 16-26. doi:10.3109/17482960902912399

Riccio, A., Simione, L., Schettini, F., Pizzimenti, A., Inghilleri, M., Olivetti Belardinelli, M., . . . Cincotti, F. (2013). Attention and P300-based BCI performance in people with amyotrophic lateral sclerosis. *Frontiers in Human Neuroscience, 7* doi:10.3389/fnhum.2013.00732

Saville, C. W., Dean, R. O., Daley, D., Intriligator, J., Boehm, S., Feige, B., & Klein, C. (2011). Electrocortical correlates of intra-subject variability in reaction times: Average and single-trial analyses. *Biological Psychology, 87*(1), 74-83.

Saville, C. W., Feige, B., Kluckert, C., Bender, S., Biscaldi, M., Berger, A., . . . Klein, C. (2015). Increased reaction time variability in attention-deficit hyperactivity disorder as a response-related phenomenon: Evidence from single-trial event-related potentials. *Journal of Child Psychology and Psychiatry, 56*(7), 801-813.

Schalk, G., & Mellinger, J. (2010). *A practical guide to brain–computer interfacing with BCI2000: General-purpose software for brain-computer interface research, data acquisition, stimulus presentation, and brain monitoring* Springer Science & Business Media.

Schettini, F., Risetti, M., Aricò, P., Formisano, R., Babiloni, F., Mattia, D., & Cincotti, F. (2015). P300 latency jitter occurrence in patients with disorders of consciousness: Toward a better design for brain computer interface applications. Paper presented at the *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* 6178-6181.

Shahriari, Y., Vaughan, T. M., McCane, L., Allison, B. Z., Wolpaw, J. R., & Krusienski, D. J. (2019). An exploration of BCI performance variations in people with amyotrophic lateral sclerosis using longitudinal EEG data. *Journal of Neural Engineering,*

Siegel, S., & Tukey, J. W. (1960). A nonparametric sum of ranks procedure for relative spread in unpaired samples. *Journal of the American Statistical Association, 55*(291), 429-445.

Smulders, F. T., Kenemans, J. L., & Kok, A. (1994). A comparison of different methods for estimating single-trial P300 latencies. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 92*(2), 107-114.

Tang, Z., Li, C., & Sun, S. (2017). Single-trial EEG classification of motor imagery using deep convolutional neural networks. *Optik-International Journal for Light and Electron Optics, 130*, 11-18.

Team, R. C. (2019). R: A language and environment for statistical computing [computer software]. Vienna, Austria:

Thompson, D. E., Mowla, M. R., & Huggins, J. E. (October 23, 2019). Evidence of latency variation in the P3 speller brain computer interface. Paper presented at the *Society for Neuroscience,*

Thompson, D. E., Warschausky, S., & Huggins, J. E. (2012). Classifier-based latency estimation: A novel way to estimate and predict BCI accuracy. *Journal of Neural Engineering, 10*(1), 016006.

Unsal, A., & Segalowitz, S. J. (1995). Sources of P300 attenuation after head injury: Single-trial amplitude, latency jitter, and EEG power. *Psychophysiology, 32*(3), 249-256.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.) Springer.

Verleger, R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology, 34*(2), 131-156.

Verleger, R., Baur, N., Metzner, M. F., & Śmigasiewicz, K. (2014). The hard oddball: Effects of difficult response selection on stimulus-related P 3 and on response-related negative potentials. *Psychophysiology, 51*(11), 1089-1100.

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics, 30*(3), 261-293.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48.

Viechtbauer, W., López-López, J. A., Sánchez-Meca, J., & Marín-Martínez, F. (2015). A comparison of procedures to test for moderators in mixed-effects meta-regression models. *Psychological Methods, 20*(3), 360.

Vieregge, P., Wauschkuhn, B., Heberlein, I., Hagenah, J., & Verleger, R. (1999). Selective attention is impaired in amyotrophic lateral sclerosis—a study of event-related EEG potentials. *Cognitive Brain Research, 8*(1), 27-35.

Walhovd, K. B., Rosquist, H., & Fjell, A. M. (2008). P300 amplitude age reductions are not caused by latency jitter. *Psychophysiology, 45*(4), 545-553.

Wolpaw, J. R., Bedlack, R. S., Reda, D. J., Ringer, R. J., Banks, P. G., Vaughan, T. M., . . . Winden, S. (2018). Independent home use of a brain-computer interface by people with amyotrophic lateral sclerosis. *Neurology, 91*(3), e258-e267.

Woody, C. D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Medical and Biological Engineering, 5*(6), 539-554.

Woolley, S. C., York, M. K., Moore, D. H., Strutt, A. M., Murphy, J., Schulz, P. E., & Katz, J. S. (2010). Detecting frontotemporal dysfunction in ALS: Utility of the ALS cognitive behavioral screen (ALS-CBS). *Amyotrophic Lateral Sclerosis: Official Publication of the World Federation of Neurology Research Group on Motor Neuron Diseases, 11*(3), 303-311. doi:10.3109/17482961003727954

Wu, W., Wu, C., Gao, S., Liu, B., Li, Y., & Gao, X. (2014). Bayesian estimation of ERP components from multicondition and multichannel EEG. *NeuroImage, 88*, 319-339.

Yu, X. (2016). *The impact of latency jitter on the interpretation of P300 in the assessment of cognitive function*

Zisk, A. H., Borgheai, S. B., McLinden, J., Hosni, S. M., Deligani, R. J., & Shahriari, Y. (2020). P300 latency jitter and its correlates in people with amyotrophic lateral sclerosis. *Clinical Neurophysiology,*

Zisk, A. H., Borgheai, S. B., McLinden, J., & Shahriari, Y. (2021). Improving longitudinal P300-BCI performance for people with ALS using a data augmentation and jitter correction approach. *Under Review at Journal of Neural Engineering.*

APPENDIX 1: SUPPLEMENTARY TABLE

Table A.1 shows the selected parameters for each participant and set of two training sessions with one testing session numbers. The symmetric time shifts and use or non-use of per-epoch shifts in augmentation are those determined from the procedure in section 3.2.4.2, and the correction windows are those determined as in section 3.2.4.3. Symmetric time-shifts for data augmentation were between 10 and 30 ms for five of the six participants. For these same five participants, correction windows were between 0 and 40 ms. For the remaining participant, ALS-01, symmetric time-shifts reached a maximum of 54.69 ms and correction windows had a maximum of 101.56 ms. Note: the individualized parameters vary both between participants and between training sets.

Table A.1. Parameter selections for each participant and set of training and testing session numbers.

Session		***Participant***								
		ALS01			ALS02			ALS03		
Training Sessions	Test Session	Augmentation Shift	Data Augmentation	Correction Window	Augmentation Shift	Data Augmentation	Correction Window	Augmentation Shift	Data Augmentation	Correction Window
1-2	3	±19.53ms	4x	±35.16ms	±19.53ms	4x	±15.63ms	±15.63ms	3x	±11.72ms
2-3	4	±27.34ms	4x	±50.78ms	±15.63ms	4x	±15.63ms	±19.53ms	4x	±7.81ms
3-4	5	±39.06ms	4x	±93.75ms	±15.63ms	3x	±3.91ms	±15.63ms	3x	±0ms
4-5	6	±35.16ms	4x	±66.41ms	±15.63ms	3x	±0ms	-	-	-
5-6	7	±35.16ms	4x	±101.56ms	±11.72ms	3x	±0ms	-	-	-
6-7	8	±39.06ms	4x	±89.84ms	±15.63ms	3x	±15.63ms	-	-	-
7-8	9	±50.78ms	4x	±0ms	±23.44ms	4x	±27.34ms	-	-	-
8-9	10	±54.69ms	4x	±23.44ms	±19.53ms	4x	±3.91ms	-	-	-
9-10	11	±42.97ms	4x	±85.94ms	±19.53ms	4x	±19.53ms	-	-	-
10-11	12	±46.88ms	4x	±85.94ms	±15.63ms	3x	±0ms	-	-	-
Mean ± STD		**39.06 ± 10.58ms**	-	**63.28 ± 34.44ms**	**17.19 ± 3.29ms**	-	**10.16 ± 9.78ms**	**16.92 ± 2.26ms**	-	**6.51 ± 5.97ms**

Session		***Participant***								
		ALS04			ALS05			ALS06		
Training Sessions	Test Session	Augmentation Shift	Data Augmentation	Correction Window	Augmentation Shift	Data Augmentation	Correction Window	Augmentation Shift	Data Augmentation	Correction Window
1-2	3	±15.63ms	4x	±23.44ms	±11.72ms	3x	±11.72ms	±27.34ms	4x	±35.16ms
2-3	4	±19.53ms	4x	±27.34ms	±11.72ms	3x	±3.91ms	±23.44ms	4x	±15.63ms
3-4	5	±23.44ms	4x	±35.16ms	±15.63ms	3x	±15.63ms	±19.53ms	4x	±15.63ms
4-5	6	±23.44ms	4x	±23.44ms	±15.63ms	3x	±11.72ms	±15.63ms	3x	±3.91ms
5-6	7	±15.63ms	4x	±15.63ms	±15.63ms	3x	±7.81ms	±15.63ms	3x	±3.91ms
6-7	8	±19.53ms	4x	±19.53ms	±15.63ms	3x	±11.72ms	±15.63ms	3x	±0ms
7-8	9	±27.34ms	4x	±23.44ms	±15.63ms	3x	±7.81ms	±15.63ms	3x	±3.91ms
8-9	10	-	-	-	±19.53ms	4x	±19.53ms	-	-	-
Mean ± STD		**20.64 ± 4.35ms**	-	**24.00 ± 6.15ms**	**15.14 ± 2.50ms**	-	**11.23 ± 4.87ms**	**18.97 ± 4.75ms**	-	**11.16 ± 12.23ms**

ABBREVIATED VITA

Education
University of Rhode Island, Kingston, RI 02881

Doctoral Candidate, Interdisciplinary Neuroscience Program (Expected May 2021)
Major professor: Dr. Yalda Shahriari
Co-major professor: Dr. Kunal Mankodiya
Dissertation: Quantifying and Compensating for P300 Variations in Amyotrophic Lateral Sclerosis

M.S., Mathematics	May 2016
B.S., Mechanical Engineering; B.A., Chinese	May 2015
B.S., Mathematics	December 2012

Selected Publications

Zisk, A. H., Borgheai, S. B., McLinden, J., Hosni, S. M., Deligani, R. J., & Shahriari, Y. (2020). P300 latency jitter and its correlates in people with amyotrophic lateral sclerosis. *Clinical Neurophysiology.*

Zisk, A. H., Borgheai, S. B., McLinden, J., & Shahriari, Y. (2021). Improving longitudinal P300-BCI performance for people with ALS using a data augmentation and jitter correction approach. Under Review at the *Journal of Neural Engineering.*

Borgheai, S. B., Deligani, R. J., McLinden, J., **Zisk, A. H**., Hosni, S. I., Abtahi, M., Mankodiya, K., Shahriari, Y. (2019). Multimodal exploration of non-motor neural functions in ALS patients using simultaneous EEG-fNIRS recording. *Journal of Neural Engineering.*

Borgheai, S. B., McLinden, J., **Zisk, A. H**., Hosni, S. I., Deligani, R. J., Abtahi, M., Mankodiya, K., Shahriari, Y. (2020). Enhancing communication for people in late-stage ALS using an fNIRS-based BCI system. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28*(5), 1198-1207.

Deligani, R. J., Hosni, S. I., Borgheai, S. B., McLinden, J., **Zisk, A. H.,** Mankodiya, K., & Shahriari, Y. (2020). Electrical and Vascular Neural Functions in People with ALS: An EEG-fNIRS Resting-State Study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* 28(12): 3129-3139.

Hosni, S. I., Deligani, R. J. **Zisk, A.,** McLinden, J., Borgheai, S. B., & Shahriari, Y. (2020). An Exploration of Neural Dynamics of Motor Imagery for People with Amyotrophic Lateral Sclerosis. *Journal of Neural Engineering*, 17:016005. (15pp)

Shahriari, Y., Besio, W., Hosni, S. I., **Zisk, A. H**., Borgheai, S. B., Deligani, R. J., & McLinden, J. (2020). Electroencephalography. In Guo, L. (Ed.) *Neural Interface Engineering: Linking the Physical World and the Nervous System*, 1-16. Springer Nature.