 DOI: 10.3217/978-3-99161-050-2-150

Efficacy of Recalibration for a P300 Speller

Katherine G. Colleran[1], Guoxuan Ma[2], Jane E. Huggins[1]*

[1]*University of Michigan Medicine, Department of Physical Medicine and Rehabilitation, Ann Arbor, Michigan;*
[2]*University of Michigan Biostatistics Department, Ann Arbor, Michigan*
*325 E. Eisenhower Pkwy, Ann Arbor, MI 48109. E-mail: janeh@umich.edu

Introduction: Most P300 brain-computer interface (BCI) designs require calibration for an individual user to teach the BCI how to interpret the event-related potentials (ERP) in that user's brain activity [1]. Re-calibration can increase BCI accuracy and efficacy, however it is not clear how beneficial re-calibration may be over different periods of time. Users find re-calibration tedious, and the process takes 20-30 minutes. The time cost and tedium of calibration should be weighed against the benefit. Previous studies suggest that the time of day does not directly effect P300 ERPs, but other factors such as recent food intake do [2]. Our study was designed to compare calibrations taken on different days and at different times of day to inform recommend timeframes for recalibration.

Material and Methods: Participants were five people without physical impairments or previous experience with this BCI. Participants were two females, three males and had mean age of 45.8 years; range 26-61. There were six sessions (about one per week): one 8-hour session followed by five 2-hour sessions. Each 2-hour interval, starting on an even hour, was called a time-slot and included calibration, copying sentences with corrections, and typing novel text describing a picture, creating the test data for each time-slot (average 60.5 selections, range 31 to 116). Participants wore a 7-channel dry electrode VR300 headset (300 Hz, re-referenced to linked ears) from Wearable Sensing. A custom BCI2000 module interfaced with PRC-Saltillo augmentative and alternative communication (AAC) NuVoice software using Microsoft User Interface Automation standards to identify active keys, create stimuli over them, and activate the key selected by the user [3]. A Mann-Whitney test was conducted to evaluate if there was a significant difference in character-level accuracy between using the most recent calibration and previous calibrations, both at the individual participant level and at the group level.

Results: Recalibration showed a statistically significant improvement in character-level accuracy at the group level ($p < 0.001$). Mean improvement at the group level was 7.65 percentage points (%), ranging from -0.83% to 31.5% for individual participants. For four of five participants, recalibration produced a statistically significant improvement in accuracy ($p < 0.005$). However, for the other participant, recalibration did not produce a significant effect (mean -0.83%, range -15.5% to 22.7%). Fig. 1 shows offline analysis of how two participants' calibrations interpreted their data collected at different times and/or sessions.

Conclusion: Recalibration usually improves accuracy. However, the magnitude of improvement may not be worth the time, especially for users who already have a high accuracy. Further analysis will determine if time of day has an effect on recalibration. Future analyses will test how to predict the efficacy of recalibration and if small amounts of data can be used to select an effective prior calibration.

Acknowledgments and Disclosures: No conflicts of interest. Supported by PRC-Saltillo through National Institutes of Health grant SB1DC015142 from the National Institute of Deafness and other Communication Disorders.

References:

[1] Farwell LA, & Donchin E, "Talking off the top of your head…", *Electroencephalogr. Clin. Neurophysiol*., 70(6) (1988): 510-523.

[2] Polich J, & Geisler MW, "P300 and time of day…", *Biol. Psychol.*, 31(2) (1990): 117-136.

[3] Schalk G et al., "BCI2000: a general-purpose..." *IEEE Trans. Biomed. Eng.*, 51(6) (2004):1034-43.

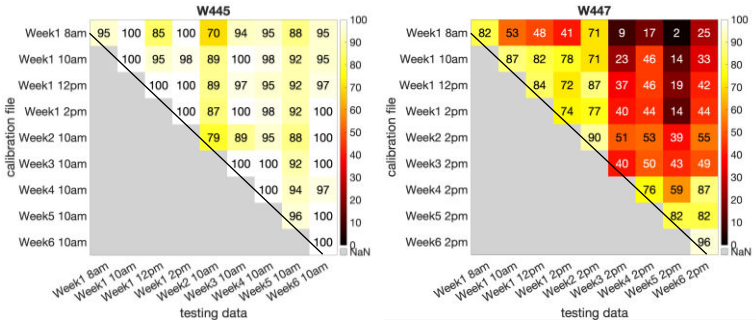

Figure 1: Heat maps show accuracy for the participants who had the median (left) and most (right) improvement with recalibration. The calibration (y-axis) for each time-slot was tested on that participant's test data for each time-slot (x-axis). Time is arranged from top-to-bottom for calibrations and left-to-right for test data. The diagonal line marks calibrations and test data for the same day and time. Lighter colors indicate better accuracy. The first four rows/columns show different time slots in session 1.

 Published by Verlag der Technischen Universität Graz