Evaluation of ERN and FRN robustness to cross-session variability

Marcel F. Hinss[1,3*], Bertille Somon[1,2*], Frédéric Dehais[1,2], Caroline P. Carvalho Chanel[1,2], Raphaëlle N. Roy[1,2]

[1]ISAE-Supaero, Université de Toulouse, France

[2]Artificial and Natural Intelligence Toulouse Institute (ANITI), Université de Toulouse, France

[3]Maastricht University, Maastricht, the Netherlands

*These Authors contributed equally.

Keywords: ERN, FRN, EEG, Flanker task, cross-session

Introduction: Brain-Computer Interfaces (BCI) are devices that allow users to interact - explicitly or implicitly - with computers based solely on brain activity (Clerc et al., 2016). Many BCIs are electroencephalography (EEG) based, and within Neuroergonomics, BCIs may be used to detect the users' mental states. Yet, EEG-based mental state estimation pipelines still do not reach ideal classification rates suffering from signal non-stationarity, including cross-session variability (Saha & Baumert, 2020). Thus, using error-related potentials (ErrPs) extracted from EEG was proposed to help improve BCI algorithms through a type of feedback loop (Chavarriaga et al., 2014). Event-related potentials (ERP) have been shown to be more robust to time effects than spectral features (Roy et al., 2016), and error potentials (ErrPs) might also prove so. These ErrPs take the shape of two frontocentral ERPs: the Error Related Negativity (ERN) for error commission and the Feedback Related Negativity (FRN, a.k.a. Reward Positivity - RewP) for negative feedback observation. Positive parietal components accompany these peaks: respectively, the error Positivity (Pe) and the P300 (Chavarriaga et al., 2014). These potentials are also elicited by observing surprising and/or erroneous system actions (Somon et al., 2018). Chavarrigia and Millán (2010) started investigating the stability of the ERN and FRN and showed them to remain stable over time. However, some factors were not investigated. The preliminary results of this study intend to extend these results in two ways: first, we assess the effect of the congruency of stimuli presented on both the ERN and FRN over three separate sessions, all one week apart; then we observe the stability of the evolution of both the execution and feedback-related ErrPs over time.

Methods: Data for this abstract are part of an undergoing more extensive study aiming at creating a substantial pBCI database (Hinss et al., 2021). Fifteen participants performed a 10-minute arrow-based flanker task on three separate sessions spaced one week apart. Participants were presented with a total of 120 congruent or incongruent combinations of arrows (50% each; see fig. 1) and had to indicate the direction of the centre arrow by button-press. Following a 2,25-2,75 (jittered) second window allowing the participant to respond, feedback was provided (see figure 1). 9 participants were removed for this preliminary analysis as they did not perform errors in at least one of the two congruency conditions and/or one session. EEG data were recorded with 64-active Ag-AgCl electrodes (Brain Products, GmbH; sampling rate: 500 Hz). One EEG electrode (TP9) was removed from the cap and used to measure ECG concomitantly. Participants' responses were recorded with a keyboard. EEG, ECG and behavioural data were streamed through the Lab Streaming Layer (LSL) and LabRecorder (Kohte, 2013). Data were preprocessed and analysed using the EEGLAB toolbox: referenced using the right mastoid (TP10), high pass filtered (FIR filter; 1 Hz) and cleaned using ASR (with a 1-min resting-state baseline). An extended ICA was performed, and the IC-Label extension was used to automatically label components as either muscle, heart or eye components to be rejected (threshold set to 90% for each). Finally, data were low-pass FIR filtered (FIR filter; 40 Hz).

EEG data were Laplacian-transformed using the CSD toolbox (Kayser & Tenke, 2006) with a spherical spline interpolation (lambda $=10^{-5}$, head radius =10cm and spline flexibility m=4) to estimate current source densities, as this method has proven very useful to improve the signal-to-noise ratio for ErrPs by removing activities related to scalp conduction. Analysis of the data focused on two different time points: participant's response and feedback display. For both, epochs of -200 ms to 750 ms centred on each event (response or feedback) were extracted. Average ERPs were computed across participants for each condition (congruency x valence x session). Component peak latency and amplitude were computed for the ERN (most negative peak at the FCz electrode in the 25-125 ms post-response time window), the Pe (the average positivity at the Pz electrode in the 200-400 ms post-response time-window), the FRN (the most negative peak at the Fz/FCz electrode in the 200-400 ms post-feedback time window) and the P300 (the average positivity at the Pz electrode in the 200-400 ms post-feedback time window) components. Statistical analyses were carried on these components with a three-way repeated-measures ANOVA with stimulus congruency - congruent vs incongruent -, valence - error vs correct response - and session - 1, 2 or 3 - as factors.

Results: Participants demonstrated the Flanker effect, as seen in an increase in reaction time and error rate during incongruent compared to congruent trials (Eriksen & Eriksen, 1974). EEG analysis revealed several tendencies. The ERN amplitude was shown to be sensitive to the participant making errors. This effect was significant at the FCz, Pz and Fz electrodes (respectively $F(1,5)=4.63$ $p<0.05$, $F(1,5)=10.58$ $p<0.01$, $F(1,5)=6.29$ $p<.001$; see fig. 2a and 2b). Post-feedback, a significant difference in amplitude on the FRN component was also observed between correct and incorrect trials at FCz and Fz ($F(1,5)= 9.42$ $p<0.05$ and $F(1,5)=15.91$ $p<0.05$ respectively; see fig. 2c and 2d). For the ERN as well as the FRN components, neither the session nor congruency had any significant effect.

Discussion: For BCIs to mature into viable systems that can work outside the lab, they have to become easy and quick to use. Usability comes from, among other things, a good transferability and generalizability of results across sessions, participants, and time. Here, it was shown that the ErrPs at response and feedback elicited during a Flanker task are consistent across congruency conditions and sessions, making them a reliable indicator of one's own or a system's error detection or surprising outcome. This consistency could improve the generalizability of BCI algorithms and lower training times for participants. Against our expectations, the P300 component at electrode Pz was not identifiable with the currently limited sample size. The Pe component was also not significantly impacted by error occurrence. However, these preliminary results have to be taken with caution due to the small sample size. We intend to expand these findings by increasing the participant number to test the viability of ErrP-based BCIs.

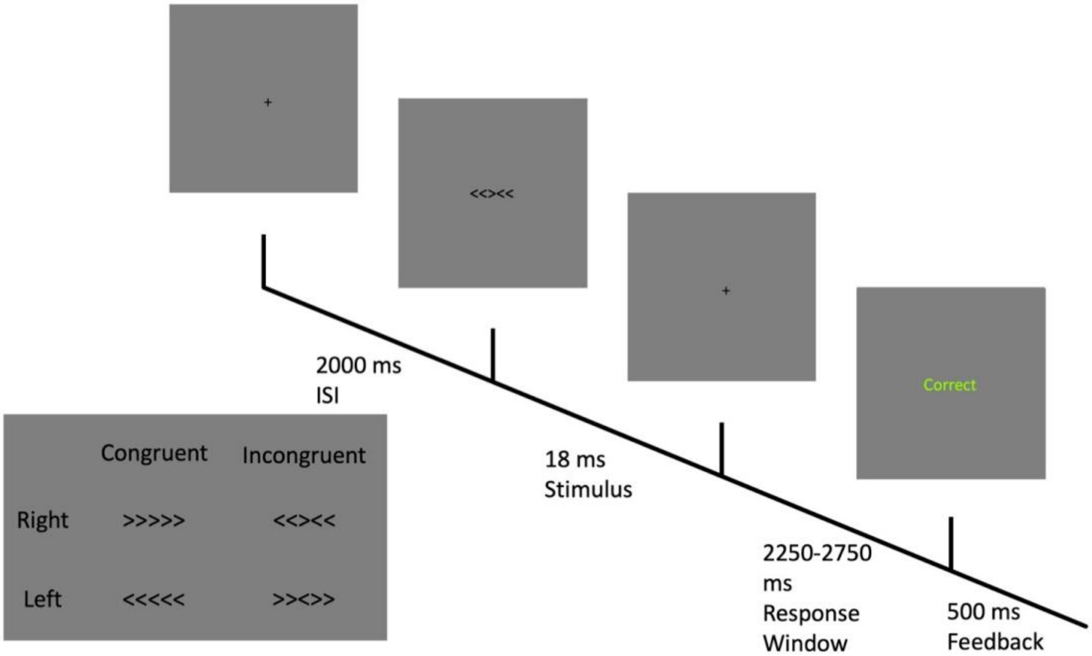

Figure 1. Flowchart of single trial as well as the 4 different stimuli.

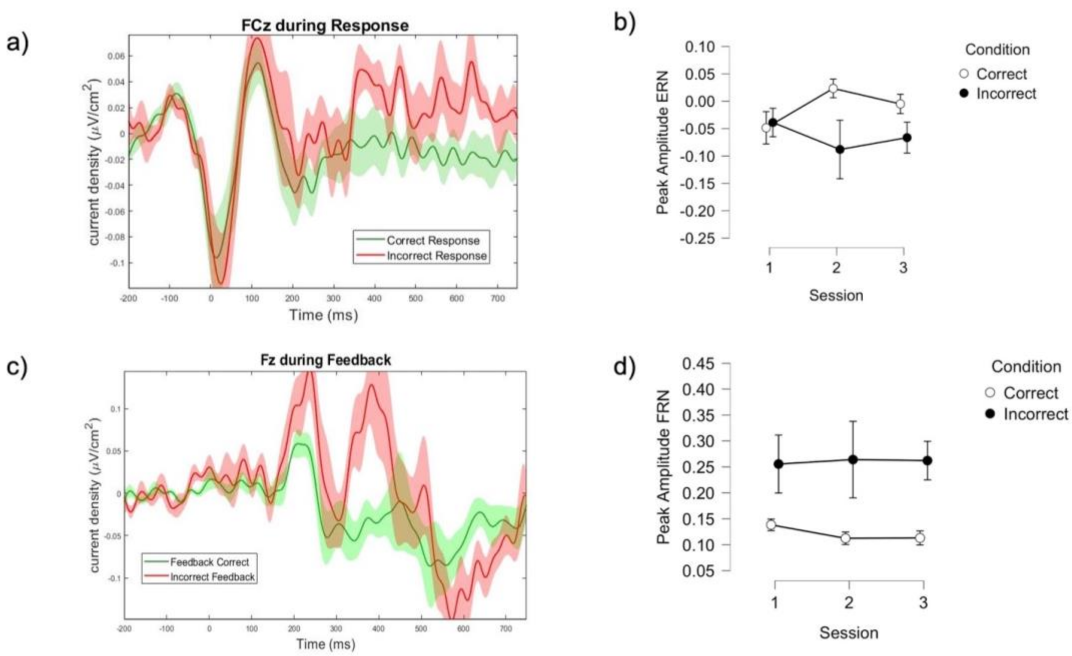

Figure 2. Grand average ERP (averaged across participants) plots with standard deviations for response at FCz (a) and feedback at Fz (c) for correct (green) and erroneous (red) trials; and statistical results across sessions for the ERN (b) and FRN (d) average amplitude across participants.

References:

Chavarriaga, R., Biasiucci, A., Forster, K., Roggen, D., Troster, G., & Millan, J. del. (2010). Adaptation of hybrid human-computer interaction systems using EEG error-related potentials. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. https://doi.org/10.1109/iembs.2010.5627376

Chavarriaga, R., Sobolewski, A., & Millan, J. Ã. del. (2014). Errare machinale est: the use of error-related potentials in brain-machine interfaces. *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00208

Clerc, M., Bougrain, L., & Lotte, F. (2016). *Brain-Computer Interfaces 1*. Wiley-ISTE.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/bf03203267

Hinss, M. F., Somon, B., Dehais, F., & Roy, R. N. (2021). Open EEG Datasets for Passive Brain-Computer Interface Applications: Lacks and Perspectives. *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*. https://doi.org/10.1109/ner49283.2021.9441214

Kothe, C., Medine, D., Boulay, C., Grivich, M., & Stenner, T. (2014). Lab streaming layer.

Roy, R. N., Charbonnier, S., Campagne, A., & Bonnet, S. (2016). Efficient mental workload estimation using task-independent EEG features. *Journal of neural engineering*, *13*(2), 026019.

Saha, S., & Baumert, M. (2020). Intra- and Inter-subject Variability in EEG-Based Sensorimotor Brain Computer Interface: A Review. *Frontiers in Computational Neuroscience*, *13*.