Real-Time Feedback of Subjective Affect and Working Memory Load Based on Neurophysiological Activity

Sabrina Gado[1(✉)], Katharina Lingelbach[1,2(✉)], Michael Bui[1], Jochem W. Rieger[2], and Mathias Vukelić[1]

[1] Fraunhofer Institute for Industrial Engineering IAO, Stuttgart, Germany
{sabrina.gado,katharina.lingelbach,michael.bui, mathias.vukelic}@iao.fraunhofer.de

[2] Department of Psychology, Carl Von Ossietzky University, Oldenburg, Germany
jochem.rieger@uol.de

Abstract. We investigated the effects of feedback on users' performance during a cognitive task with concurrent emotional distraction. Our aim was to provide participants with insights into their current affective and cognitive state by measuring and decoding brain activity. Therefore, a real-time preprocessing, analyzing, and visualization routine was developed based on electroencephalographic (EEG) data measured during a primary study. To explore users' behavioral and neurophysiological reactions, error-tolerance as well as possibilities to improve feedback accuracy by the means of feedback-based event-related potentials (ERPs), we provided either legit or inappropriate sham feedback in a second study. The kind of feedback (legit or inappropriate) had only marginal influence on participants' subsequent performance. On a neuronal level, we did not observe differences in the ERPs evoked by the legit and inappropriate feedback. In qualitative interviews, participants evaluated the feedback as interesting but also sometimes irritating due to odd feedback trials. Our study emphasizes the importance of performance accuracy and transparency towards users regarding the underlying feedback computations.

Keywords: Brain-computer interfaces · Electroencephalography (EEG) · Feedback · Adaptive systems · State monitoring · Affect · Working memory load

1 Introduction

Identifying users' mental states is a decisive task for many human-machine applications like in industrial production, semi-autonomous vehicles, medical surgery, or in the context of learning. Providing users with insights on their current affective and cognitive state by measuring and decoding brain activity allows to foster self-regulation and stress-management [1]. It might even enhance cognitive performance via neurofeedback [2]. Especially affective states are known to be significant predictors of work performance

S. Gado and K. Lingelbach—The Authors contributed equally to this research.

C. Stephanidis et al. (Eds.): HCII 2021, CCIS 1499, pp. 80–87, 2021.
https://doi.org/10.1007/978-3-030-90179-0_11

[3] and satisfaction [4]. Whereby, workload is related to occupational exhaustion, stress, and fatigue [5]. Research simultaneously investigating interdependencies of the two constructs [6] and their decoding for adaptive application is, unfortunately, scarce [7]. For the application of (neuro-)adaptive systems, the feedback about the recognized states, its perceived appropriateness, and reliability of the system are essential factors. How users perceive and evaluate closed-loop human-machine systems, is significantly mediated by trust: Previous research revealed that trust in an agent or system is strongly influenced by its reliability in task performance and negatively correlated with perceived errors of the automated system [8, 9]. Consequently, users' acceptance and trust regarding a system is interrelated with the perceived accuracy of the system's feedback and subjective error tolerance of the user. In their research on performance monitoring feedback loops, Alder and Ambrose [10] highlighted the influence of perceived feedback accuracy and fairness as well as perceived control over the feedback (e.g., its frequency) on attitudinal reactions like satisfaction and commitment as well as behavioral outcomes. Hence, the perceived feedback appropriateness and accuracy are suggested to be critical, since these factors seem to affect the feedback's impact on users' performance, their attitude towards the system, and the perceived usefulness of the system application. To investigate how precise respective applications should be, one has to explore the error tolerance of users with respect to feedback on their current affective and cognitive states. With electroencephalographic (EEG) recordings, event-related potential (ERP) responses evoked by the feedback can be used for an automatic error correction to improve subsequent feedback cycles [11, 12]. The ERP responses differ depending on whether a feedback is appropriate or not. The feedback-related negativity (FRN), a negative deflection around 250 ms after feedback-onset (comparable with the error-related negativity), and the P300, an indicator for mismatch between internal and external representations, are sensitive to erroneous feedback [13].

Here, we investigate whether we can continuously monitor users' current states and provide an intuitive feedback of recognized states. Therefore, a real-time preprocessing, analyzing, and visualization routine was developed based on an experimental dataset of a preliminary study. In the second feedback study, on which we mainly focus here, we were interested in two aspects: 1) the effectiveness and evaluation of non-reliable feedback by investigating users' reactions to a sham feedback that was either legit (consistent with the task condition) or inappropriate (inconsistent with the task condition) [cf. 14, 15] and 2) the detection of neuronal correlates associated with erroneous feedback to improve the accuracy. Both studies were realized via a wireless, easy-to-use EEG with dry electrodes [cf. 16].

2 Preliminary Study

2.1 Participants Declaration

Eight participants (three female) took part in the preliminary study (mean age 23 years, $SD = 1.12$) and seven participants (four female) in the second feedback study (mean age 25.48 years, $SD = 2.66$). All participants had normal or corrected-to-normal vision, no psychiatric history, and were free of neurological diseases. They signed an informed consent according to the recommendations of the declaration of Helsinki. The study was approved by the ethics committee of the Medical Faculty of the University of Tuebingen, Germany (ID: 827/2020BO1).

2.2 Experimental Procedure

The real-time preprocessing, analyzing, and visualization routine of the feedback was developed based on the dataset acquired in the preliminary study where participants performed elementary and complex arithmetic tasks with concurrent auditory emotional distractions (negatively, neutrally, and positively associated sounds). After each mathematical task, we asked them to rate their subjectively perceived affect and effort. We investigated neurophysiological correlates and behavioral outcomes in a 2 (low working memory load vs. high working memory load) × 3 (low valence, neutral valence, and high valence) design resulting in six experimental conditions. The experimental framework was similar for the following feedback study. The experimental setup can be seen in Fig. 1A.

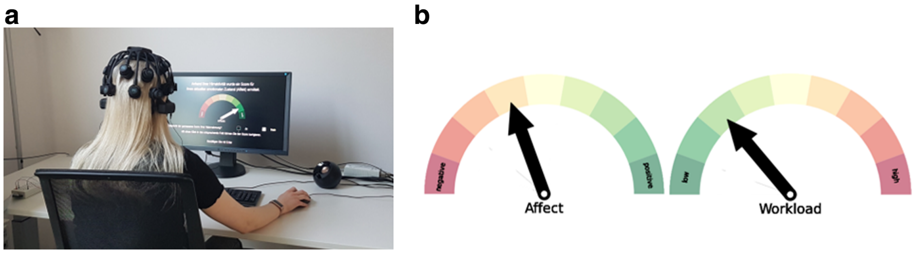

Fig. 1. **A)** Experimental setup. **B)** Visualization of the recognized affective and cognitive state based on the frontal alpha asymmetry score (left) and workload index (right).

2.3 Data Acquisition, Preprocessing and Analysis

EEG data was recorded according to the international 10–20 system using a wireless, easy-to-use EEG headset well-suited for the application in naturalistic settings with 20 electrodes and dry sensor technology. The EEG was grounded to the left mastoid, which was also used as common reference. The impedance was kept below 2,500 kΩ at the onset of the experiment. Data was acquired at a sampling rate of 500 Hz and saved via LabStreaming Layer. Data analysis was performed with custom written scripts in python™. For the online analysis, the signal was re-referenced to the Cz-electrode, detrended, and filtered with a second order zero-phase lag infinite impulse response (IIR) filter using a narrow frequency band of 0.5 to 14 Hz. During the task, the affective and cognitive states of the preceding 2 s were estimated every 0.5 s using power spectral measures computed via a modified version of the Fast Fourier Transformation (FFT), the so-called Welch's method. For the estimation of affect, the frontal alpha asymmetry coefficient was calculated [17], while working memory load estimates were derived from frontal theta and parietal alpha power [18]. A 3-min resting state measurement served as baseline for the subsequent estimation of the online scores. Substantial changes in these scores that exceeded one standard deviation of the baseline scores were concurrently translated to visual feedback. The feedback was presented on a gauge dial (see Fig. 1B).

3 Feedback Study

In the second study, we investigate neuronal and behavioral effects (i.e., response time and accuracy) of either appropriate or erroneous sham visual feedback [cf. 14, 15]. We further explore how participants accept and experience the real-time feedback. As ERP responses evoked by feedback allow to continuously improve the system's accuracy, we were interested whether we can identify neuronal correlates distinguishing appropriate and erroneous feedback [11, 12]. After each trial, we showed participants a sham feedback allegedly based on their brain activity during the task and asked them to potentially correct the score according to their own perception by clicking in the respective field. In 80% of the trials, the feedbacked score corresponded to the working memory load or emotional valence condition; in 20%, we presented an odd feedback, e.g., high cognitive load score during a rather simple task. After the experiment, we asked participants in a semi-structured qualitative interview how they perceived the feedback and whether they used it to adapt their behavior.

To analyze the neurophysiological data offline, EEG signals were de-trended, band-pass filtered between 0.5 to 23 Hz using a zero-phase lag finite impulse response (FIR) filter and cut into epochs starting 200 ms before to 1 s after feedback onset. We rejected epochs containing a maximum deviation above 250 μV in any frontal EEG channels (Fp1, Fp2). Afterwards, an independent component analysis (ICA) using the extended infomax ICA algorithm [19] as implemented in the MNE-Python toolbox [20] was used to remove cardiac-related and muscular artefacts as well as ocular movement by careful visual inspection of the topography, time course, and power spectral intensity [21]. The epochs were baseline corrected by subtracting the mean amplitude of the time interval before feedback onset. For identifying differences in the ERPs between the feedback conditions (odd vs. legit for affect and workload), we used a cluster-based, non-parametric randomization approach [22]. Clusters were identified as adjacent points in space (EEG channels) and time (samples in the epoch) using a T-value based cluster-level threshold of $p < .01$ and group-level threshold of $p < .05$ (two-sided).

To investigate behavioral effects of the feedback, we performed one-way repeated measures analyses of variance (rANOVAs) with feedback (legit vs. inappropriate) as main effect and perceived correctness reflected in probability that users correct the feedbacked score as well as response time and accuracy in the subsequent trial as dependent variables. Additionally, we explored interaction effects on performance regarding the kind of feedback (legit vs. inappropriate) and experimental condition.

4 Results

4.1 Effects on Perceived Correctness

Participants were significantly more likely to correct an odd feedback compared to a reasonable one regarding the cognitive effort, $F(1, 6) = 30.82$, $p < .001$, partial $\eta^2 = .84$. Similarly, participants tended to be more likely to correct an odd feedback compared to a reasonable one regarding the affective state, $F(1, 6) = 5.14$, $p = .064$, partial $\eta^2 = .46$.

4.2 Effects on Performance

No significant effects were found of inappropriate, odd feedback on participants' performance in the subsequent trial. Interestingly, there was neither a difference between trials with previous odd or appropriate feedback for affect nor working memory load (see Fig. 2). Further, we observed no interactions between feedback and experimental condition for affect and effort. Increased working memory load did not change the perceived correctness of and probability to adjust an inappropriate feedback score.

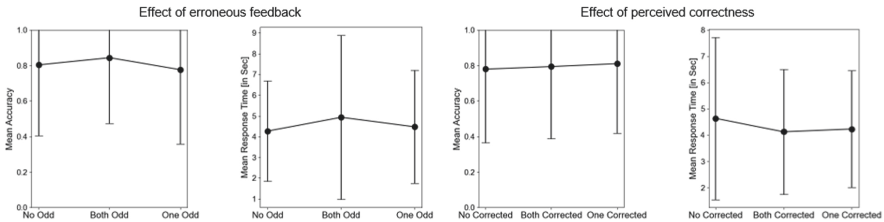

Fig. 2. Effects of odd feedback (left) and perceived correctness of the scores (right) on performance (accuracy and response time) in the subsequent trial. Error bars = standard deviation.

4.3 Results of the ERP Analysis

The cluster-based, non-parametric randomization test revealed no significant spatio-temporal cluster for the difference between the appropriated and odd feedback. Figure 3 depicts the grand average response over epochs and participants per condition and region of interest (frontal, central, and parietal).

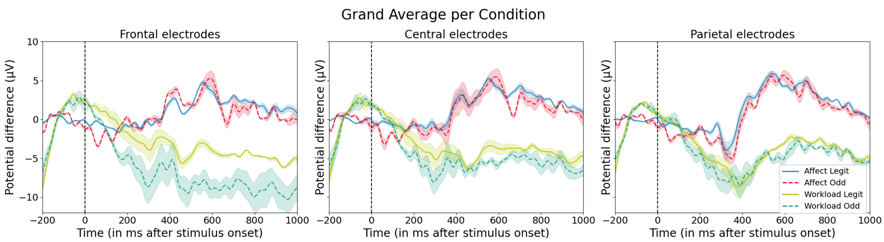

Fig. 3. Grand average over epochs and participants per condition. Dashed lines: odd feedback, solid lines: legit feedback. Shaded area: standard deviation.

4.4 Insights Gained in the Interviews

Most participants evaluated the feedbacked scores positively and as interesting. However, some participants also characterized the feedback as irritating. The design and feedback format were perceived as suitable and engaging. Some participants recommended to provide detailed explanation on the underlying computations, the used data (components of the data), and recommendations for the score interpretation. About half of the participants reported that they rather did not use the provided feedback scores intentionally to change their behavior in the subsequent trial. One reported a higher intrinsic motivation to increase focus and concentration. Two responded that they had memorized the feedback-sound combination for particularly irritating sounds to suppress them more efficiently in upcoming trials. Most of the participants could not imagine using the technology in realistic environments. Especially, because some of them perceived the dry-electrode EEG device as uncomfortable over the course of time. Nevertheless, they stated that monitoring of current or potentially only critical affective and cognitive states might be interesting in safety-relevant applications, e.g., when maneuvering a car or airplane. They reported further potentials of adaptive feedback systems to enhance effectiveness in learning and training scenarios.

5 Discussion and Conclusion

Our real-time EEG-based feedback approach contributes to the development of closed-loop human-machine systems allowing to recognize users' state, provide feedback, and adapt the system parameters to individual capabilities and demands. Our study revealed two main challenges for adaptive feedback systems: 1) We observed no significant positive effect on participants' performance for appropriate compared to erroneous feedback. Since participants reported irritation in response to the inaccurate feedback and the wish to get more information regarding the score computation, it is likely that they had limited trust in the feedback system. Although, they evaluated the system in general positively, their (involuntary) trust evaluation might have mediated the feedback's impact. Probably only systems perceived as reliable and consistent, are able to induce effects on a behavioral and neuronal level [23]. Therefore, participants in our study might have ignored or suppressed the feedback without considering it as a significant cue to change behavioral strategies. An alternative explanation why participants did not actively use the feedback, might be that they did not perceive it as relevant for solving the arithmetic task. In addition, we did not provide any explicit instruction to use the feedback during the task. 2) On the neuronal level, we observed no difference in feedback-related potentials between the conditions. This absence of distinct neuronal correlates associated with erroneous feedback (FRN, P300) might be explained by either the inconsistent feedback performance or the lower signal to noise ratio due to an insufficient number of trials or the dry-electrode EEG device [16].

With this study, we investigated an approach to provide real-time insights into users' cognitive and affective states during a cognitively demanding task and the neuronal and behavioral effects of the given feedback. The described research contributed to the development of closed-loop human-machine systems and understanding of associated challenges in performance-oriented contexts.

References

1. Yu, B., Funk, M., Hu, J., Wang, Q., Feijs, L.: Biofeedback for everyday stress management: a systematic review. Front. ICT **5**(23), 1–22 (2018). https://doi.org/10.3389/fict.2018.00023
2. Dessy, E., Van Puyvelde, M., Mairesse, O., Neyt, X., Pattyn, N.: Cognitive performance enhancement: do biofeedback and neurofeedback work? J. Cogn. Enhancement **2**(1), 12–42 (2017). https://doi.org/10.1007/s41465-017-0039-y
3. Shockley, K.M., Ispas, D., Rossi, M.E., Levine, E.L.: A meta-analytic investigation of the relationship between state affect, discrete emotions, and job performance. Hum. Perform. **25**(5), 377–411 (2012). https://doi.org/10.1080/08959285.2012.721832
4. Niklas, C.D., Dormann, C.: The impact of state affect on job satisfaction. Eur. J. Work Organ. Psychol. **14**(4), 367–388 (2005). https://doi.org/10.1080/13594320500348880
5. Bowling, N.A., Alarcon, G.M., Bragg, C.B., Hartman, M.J.: A meta-analytic examination of the potential correlates and consequences of workload. Work Stress **29**(2), 95–113 (2015). https://doi.org/10.1080/02678373.2015.1033037
6. Moore, M., Shafer, A.T., Bakhtiari, R., Dolcos, F., Singhal, A.: Integration of spatio-temporal dynamics in emotion-cognition interactions: a simultaneous fMRI-ERP investigation using the emotional oddball task. NeuroImage **202**, 116078 (2019). https://doi.org/10.1016/j.neuroimage.2019.116078
7. Maior, H.A., Wilson, M.L., Sharples, S.: Workload alerts - using physiological measures of mental workload to provide feedback during tasks. ACM Trans. Comput.-Hum. Interact. **25**(2), 1–25 (2018). https://doi.org/10.1145/3173380
8. Chen, M., Nikolaidis, S., Soh, H., Hsu, D., Srinivasa, S.: Planning with trust for human-robot collaboration. In: Proceedings of the Anual ACM/IEEE International Conference on Human-Robot Interaction, Chicago, IL, USA 2018, pp. 307–315. Association for Computing Machinery (2018). https://doi.org/10.1145/3171221.3171264
9. Master, R., et al.: Measurement of trust over time in hybrid inspection systems. Hum. Factors Ergon. Manuf. Serv. Ind. **15**(2), 177–196 (2005). https://doi.org/10.1002/hfm.20021
10. Alder, G.S., Ambrose, M.L.: Towards understanding fairness judgments associated with computer performance monitoring: an integration of the feedback, justice, and monitoring research. Hum. Resour. Manag. Rev. **15**(1), 43–67 (2005). https://doi.org/10.1016/j.hrmr.2005.01.001
11. Ferrez, P.W., Millan, J.d.R.: Error-related EEG potentials generated during simulated brain-computer interaction. IEEE Trans. Biomed. Eng. **55**(3), 923–929 (2008). https://doi.org/10.1109/TBME.2007.908083
12. Mattout, J., Perrin, M., Bertrand, O., Maby, E.: Improving BCI performance through co-adaptation: applications to the P300-speller. Ann. Phys. Rehabil. Med. **58**(1), 23–28 (2015). https://doi.org/10.1016/j.rehab.2014.10.006
13. Pfabigan, D.M., Alexopoulos, J., Bauer, H., Sailer, U.: Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. Psychophysiology **48**(5), 656–664 (2011). https://doi.org/10.1111/j.1469-8986.2010.01136.x
14. Enriquez-Geppert, S., Huster, R.J., Herrmann, C.S.: EEG-neurofeedback as a tool to modulate cognition and behaviour: a review tutorial. Front. Hum. Neurosci. **11**(51), 1–19 (2017). https://doi.org/10.3389/fnhum.2017.00051
15. Logemann, H.N.A., Lansbergen, M.M., Van Os, T.W.D.P., Böcker, K.B.E., Kenemans, J.L.: The Effectiveness of EEG-feedback on attention, impulsivity and EEG: a sham feedback controlled study. Neurosci. Lett. **479**(1), 49–53 (2010). https://doi.org/10.1016/j.neulet.2010.05.026

16. Guger, C., Krausz, G., Allison, B., Edlinger, G.: Comparison of dry and gel based electrodes for P300 brain-computer interfaces. Front. Neurosci. **6**(60), 1–7 (2012). https://doi.org/10.3389/fnins.2012.00060
17. Smith, E.E., Reznik, S.J., Stewart, J.L., Allen, J.J.B.: Assessing and conceptualizing frontal EEG asymmetry: an updated primer on recording, processing, analyzing, and interpreting frontal alpha asymmetry. Int. J. Psychophysiol. **111**, 98–114 (2017). https://doi.org/10.1016/j.ijpsycho.2016.11.005
18. Käthner, I., Wriessnegger, S.C., Müller-Putz, G.R., Kübler, A., Halder, S.: Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain-computer interface. Biol. Psychol. **102**, 118–129 (2014). https://doi.org/10.1016/j.biopsycho.2014.07.014
19. Lee, T.-W., Girolami, M., Sejnowski, T.J.: Independent component analysis using an extended infomax algorithm for mixed Subgaussian and Supergaussian sources. Neural Comput. **11**(2), 417–441 (1999). https://doi.org/10.1162/089976699300016719
20. Gramfort, A., et al.: MNE software for processing MEG and EEG data. NeuroImage **86**, 446–460 (2014). https://doi.org/10.1016/j.neuroimage.2013.10.027
21. Chaumon, M., Bishop, D.V.M., Busch, N.A.: A practical guide to the selection of independent components of the electroencephalogram for artifact correction. J. Neurosci. Methods **250**, 47–63 (2015). https://doi.org/10.1016/j.jneumeth.2015.02.025
22. Maris, E., Oostenveld, R.: Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods **164**(1), 177–190 (2007). https://doi.org/10.1016/j.jneumeth.2007.03.024
23. Kluger, A.N., DeNisi, A.: The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychol. Bull. **119**(2), 254–284 (1996). https://doi.org/10.1037/0033-2909.119.2.254