Neuro-adaptive tutoring systems

Neurophysiological-based recognition of affective-emotional and cognitive states of learners for intelligent neuro-adaptive tutoring systems

Katharina Lingelbach, Sabrina Gado, Wilhelm Bauer

1. Introduction

In the last decade, a growing number of studies have focused on decoding activation patterns from neurophysiological measures to identify current cognitive or affective-emotional states (e.g., Appriou et al., 2020; Fairclough, 2009; Parasuraman & Rizzo, 2008; Picard, 2000). The detection and monitoring of learners' mental states by means of a brain-computer interface (BCI) enables a continuous assessment of currently available cognitive resources, attention, and motivation. A BCI is a technical interface between the brain and a computer. In most cases, sensors on the surface of the head or skin are used to measure signals from the peripheral and central nervous system (e.g., the electroencephalography, EEG for recording brain activity). Machine learning techniques allow to process these signals and classify current mental states of learners. In a next step, the recognized states can be transmitted to a computer, for example, an intelligent tutoring system (e.g., Wolpaw et al., 2002; Vukelić et al., 2021).

In this paper, we provide an overview of research on neuro-adaptive systems and the recognition of affective-emotional and cognitive states. Current research findings are presented and explained within the illustrative application of a neuro-adaptive tutoring system. Chapters 2 and 3 present two validation studies on the continuous neuro-physiological based recognition of affective-emotional and cognitive states. The results of the studies and their implications are discussed in Chapters 4 and 5. Finally, an outlook on future research questions and methodologies is provided.

1.1. Neuro-adaptive systems and Brain-Computer Interfaces for recognizing the learners' states

A neuro-adaptive system is a system in which (1) neurophysiological signals are recorded in a closed human-technology loop, (2) mental user states are interpreted from these signals, and (3) system behaviour is adapted, accordingly. A neuro-adaptive closed-loop system has the great potential to adapt learning content, learning speed, and learning formats to the needs and abilities of the learners via

DOI: https://doi.org/10.30844/wgab_2021_15

an intelligent tutoring system. A major achievement in the field of neuro-adaptive systems has been the development of "passive" BCIs (e.g., Zander & Kothe, 2011). The main difference between passive BCIs and the more familiar "active" BCIs is that passive BCIs record implicit and spontaneous brain activity. In active BCIs, users voluntarily and mentally transmit specific commands to a computer application – they, thus, "actively" take control (for example, in physically severely impaired stroke or locked-in patients, e.g., Bensch et al., 2007; Brauchle et al., 2015; Carlson & Millan, 2013; Leeb et al., 2015). In contrast, passive BCIs do not require voluntary actuation. Hence, the person is not disturbed in his or her current task. For state detection, different measurement methods can be combined to create a hybrid BCI (e.g., an EEG with an electrocardiography (ECG) to collect cardiac activity, additionally). The use of multiple measurement methods has the advantage of a more robust and convergent estimation of the current mental states.

In addition to the system adaption towards the learner's needs and abilities, providing feedback to him or her on the current affective-emotional and cognitive states can promote self-regulation (Yu et al., 2018) and improve cognitive performance via neurofeedback (Dessy et al., 2018; Kosuru et al., 2019). Perceived successes during learning, lead to pleasant affective-emotional states, such as pleasure, promote perceived self-efficacy, and positively impact intrinsic motivation and, thus, performance in everyday and professional contexts (Shockley et al., 2012; Niklas & Dormann, 2005).

1.2. Cognitive load and affective-emotional states during learning

For the work context, concepts of lifelong learning (LLL) and employee training and reskilling are crucial for performance and maintaining competitiveness (World Economic Forum, 2019). Therefore, an optimal fit between learners and cognitive demands of the learning unit is desirable. Mental or cognitive (work)-load is defined as the ratio of available mental resources relative to the resources required to complete a task (Hart & Staveland, 1988). The more the available resources are required by the demands of a task, the higher the (perceived) cognitive load. Too high cognitive load, e.g., due to training units that are too difficult or demanding, is associated with occupational exhaustion, stress, fatigue, and, consequently, reduced performance (Bowling et al., 2015; Gevins & Smith, 2003; see also DIN EN ISO 26800 2011). Contrarily, too little stimulation and cognitive demand can lead to underload, loss of focus, and even reduced abilities (see Young & Stanton 2002; Young et al., 2015). In the case of an optimal fit between the task difficulty and the learner's abilities, a state described in positive psychology as flow can occur. This flow state is characterized by the fact that the performance of the task is perceived as rewarding and the person is absorbed in the activity (e.g., Nakamura & Csikszentmihalyi, 2009).

1.3. Physiological correlates of cognitive load and affective-emotional states

In order to use the neuro- and peripheral-physiological signals for decoding mental states, it is relevant to identify informative patterns and dynamics. Regarding the visceral or autonomic nervous system, increased mental workload is associated with a decrease in parasympathetic nervous system activity and an increase in sympathetic nervous system activity (Babiloni, 2019). The two systems act antagonistically, with the parasympathetic nervous system being associated with regeneration and digestion ("Rest and Digest") and the sympathetic nervous system being associated with activation and readiness for action ("Fight or Flight"). These changes in the autonomic nervous system activity can be detected by various peripheral physiological signals; for example, skin conductance (electrodermal activity (EDA, e.g., Roth, 1983), heart rate and heart rate variability (Berntson et al., 1997), and pupil dilation (Pomplun & Sunkara, 2003).

Pleasure and positive learning experiences are natural reinforcers during the learning processes that promote willingness to learn and, potentially as a consequence, learning success. Furthermore, affective-emotional states are particularly relevant, as they influence performance by means of mediators such as motivation and engagement. Due to the development of sophisticated neurophysiological recording and signal analysis methods, continuous real-time recording and decoding of emotional reactions has experienced significant progress in recent years. To decode emotions based on neuro- and peripheral physiological activity, affective-emotional states need to be operationalized first. Several approaches have been proposed in the literature. One frequently used model is the dimensional Circumplex model introduced by Russell (1980). It describes emotions with the help of the two dimensions *valence* (degree of evaluation: positive over neutral to negative) and *arousal* (degree of activation: calm to excited). While valence can be investigated via neurophysiological measurement methods, such as the EEG (Shu et al., 2018; Verma & Tiwary, 2014), peripheral-physiological measurement methods, such as the EDA and ECG, provide suitable correlates of arousal. In the past, the frontal alpha asymmetry (FAA) index has been proposed as a suitable index for decoding affective-emotional states (Smith et al., 2017). The FAA index is calculated by subtracting the EEG alpha power (i.e., oscillatory signals in the frequency range between 8 and 12 Hz) of the left hemisphere from the right hemisphere (Ahern & Schwartz, 1985). The ratio of frontal theta to parietal alpha power is used as a workload index (WL index) to detect cognitive load (e.g., Brouwer et al., 2012; Gevins & Smith, 2003).

Compared to self-reports (e.g., via questionnaires), peripheral and neurophysiological signals are stated to be more objective and unbiased when detecting cognitive and affective-emotional states. Bias in self-reports and other more subjective measures might among others occur due to social desirability or limitation of language (e.g., Nisbett & Wilson, 1977; Scherer & Ceschi, 2000).

However, especially in naturalistic settings, outside the controlled laboratory context, cognitive and affective-emotional states rarely occur separately. They are rather intertwined and interdependent (e.g., Cromheeke & Mueller, 2014; Ihme et al., 2018; Seleznov et al., 2019). In everyday life, we are confronted with complex, (socio-) emotional stimuli, demanding our cognitive resources like attention (e.g., a crying baby in a home office or laughter in an open-plan office). Previous research has shown impairing effects of task-irrelevant emotional distraction on cognitive load and working memory performance (Cromheeke & Mueller, 2014; Dolcos & Denkova, 2014; Iordan et al., 2013). There seems to be a relationship between the degree of cognitive load and affective-emotional processing. When studying interacting cognitive and affective-emotional states, identifying the neural dynamics and networks involved in order to adequately describe the interaction is a major challenge that requires further research (Morawetz et al., 2020; Okon-Singer et al., 2015; Seleznov et al., 2019; Zinchenko et al., 2020). Machine learning (ML) approaches may potentially provide a tool to identify informative correlates that decode complex cognitive and affective-emotional states and their interaction (e.g., King & Dehaene, 2014).

1.4. Neurofeedback in adaptive tutoring systems

When using neuro-adaptive systems in naturalistic applications, there are some factors that significantly influence effectiveness and acceptance: (1) feedback regarding the recognized states, (2) its perceived appropriateness, and (3) the reliability of the system. Thus, how learners perceive and evaluate feedback from neuro-adaptive tutoring systems is strongly influenced by trust: previous research has shown that trust in an agent or system is strongly affected by its reliability in task performance and negatively correlated with perceived errors of the automated system (Chen et al., 2018; Master et al., 2005). Consequently, acceptance and trust in a system are related to the perceived accuracy of the feedback and the subjective tolerance for error of the users. Alder and Ambrose's (2005) research examined the effect of perceived accuracy, fairness of feedback, and control over feedback (e.g., frequency of feedback) on satisfaction and engagement as well as behavioural measures. The authors reported that the perceived appropriateness and accuracy of feedback are critical, as these factors influence the impact of feedback on performance, attitudes toward the system, and its perceived usefulness. Using EEG, the responses evoked by the feedback can be explored in terms of Event-Related Potentials (ERPs) and used to automatically improve the feedback of the neuro-adaptive tutoring system (Ferrez & Millan, 2008; Mattout et al., 2015). ERP responses differ depending on whether feedback is perceived as appropriate or not. Two ERP responses are indicators of mental adjustment between internal and external representations and, thus, erroneously perceived feedback (Pfabigan et al., 2011): First, a negative deflection approximately 250 ms after the onset of the feedback (i.e., the Feedback-Related Negativity, FRN, which is comparable to the Error-Related Negativity) and second, a positive deflection after approximately

300 ms (e.g., P300). The indicators represent the internal process that the person perceives a discrepancy between expected and experienced feedback. To reduce the discrepancy in future interactions, expectations are adjusted based on experience.

In this paper, we present a neurophysiological-based approach to continuously capture learners' cognitive and affective-emotional states by measuring and decoding brain activity using a passive EEG-based BCI. The described research vision of a closed-loop neuro-adaptive tutoring system allows the system to learn from and adapt to detected mental states estimated from neurophysiological activation patterns.

We focus on the following research questions: (1) How well can we decode the interaction of mental states using theoretically supported correlates? (2) Can we predict subjective appraisal using neurophysiological correlates?

In a second study, we investigate (3) what effect the feedback of recognized cognitive and affective-emotional states has on performance (i.e., reaction time and accuracy). In this study, we focus on two aspects: (a) the effectiveness and assessment of unreliable feedback examined using either legitimate (consistent with the experimental condition) or inadequate (inconsistent with the experimental condition) feedback (Enriquez-Geppert et al., 2017; Logemann et al., 2010) and (b) the detection of neural correlates associated with erroneous feedback.

2. Methods

In the following, two validation studies and their results on EEG-based continuous recognition of cognitive and affective-emotional states are presented.

2.1. Sample

Eight participants participated in the first study (three women; *M* = 23 years; *SD* = 1.12; we used data from five participants for the decoding due to strong artifact in the remaining three participants) and another seven participants participated in the follow-up study (four women; *M* = 25.48 years; *SD* = 2.66). The purpose of the first study was to develop a method for continuous estimation and visualization of mental states. In the second study, we examined the effect of continuous feedback of recognized states on performance. Participants had corrected or normal vision and reported no psychiatric or neurological disorders. The study was approved by the Ethics Committee of the Medical Faculty of the University of Tübingen (ID: 827/2020BO1) and preregistered on OSF (osf.io/gnst5). Prior to the study, participants signed an informed consent form according to the recommendations of the Declaration of Helsinki.

2.2. Experimental design and procedere

At the beginning of each session, a three-minute resting state measurement of the EEG signals was performed. During the resting state recording participants had their eyes open and fixated on a crosshair positioned in the centre of the screen. Afterwards, participants performed arithmetic tasks requiring the addition of either low, 1-digit numbers (low working memory load, LWML) or larger, 2-digit numbers (high working memory load, HWML). The participants had to add up three consecutive numbers while updating and retaining the intermediate result in their memory. At the same time, affective-emotional states were induced by auditory sounds with negative (Low Valence, LV), neutral (Neutral Valence, NV), or positive (High Valence, HV) content from the International Affective Digitized Sounds (I-ADS) database (Bradley & Lang, 2007). This results in a 2 × 3 factorial study design with the cognitive conditions LWML and HWML and affective-emotional conditions LV, NV, and HV.

In the first study, participants rated the sounds after each presentation in terms of subjectively perceived valence using the Emoij-Grid (Toet & van Erp, 2019). After completing three consecutive arithmetic operations of the same difficulty, participants were asked to rate the perceived effort using the NASA Task Load Index (Hoonakker et al., 2011). Figure 01 provides a detailed overview of the experimental procedure.

In the second trial, we no longer asked participants about their subjective evaluation of the stimuli, but instead, after completing three successive arithmetic operations, we showed them either legitimate (i.e., consistent with the experimental condition) or inadequate (i.e., inconsistent with the experimental condition) feedback related to the previous cognitive and affective-emotional states (see Figure 02). Participants could correct the rating via mouse click on the scale according to their own perception. In 80% of the runs, the feedback score was consistent and, thus, legitimate with the experimental condition; in 20%, it was inconsistent and, thus, inadequate. For example, a high recognized cognitive load was feedbacked after a simple task (inadequate feedback). After the experiment, we asked the participants in a semi-structured qualitative interview how they perceived the feedback.

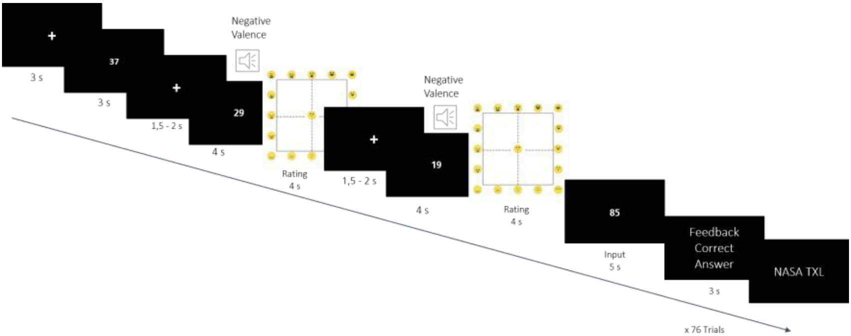

Figure 01: Experimental procedure of the first study. After each auditory stimulus and after each cognitive task, participants provided subjective ratings regarding the perceived valence and cognitive effort.

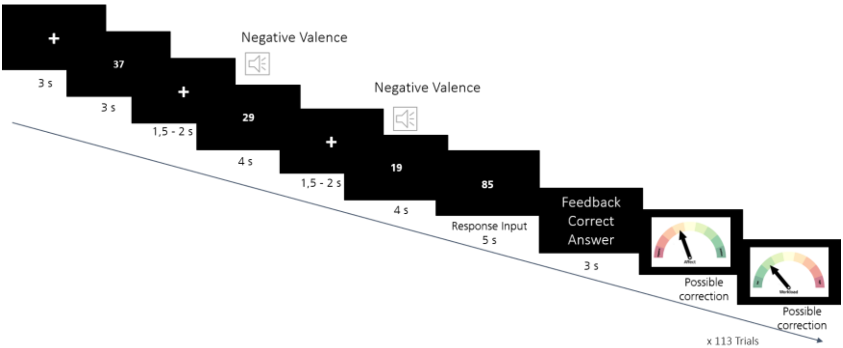

Figure 02: Experimental procedure of the second study. After each run, two scores indicating the recognized valence and cognitive effort experienced during the previous trial were presented to the participants. In 20% of the cases, the presented scores were not adequate. The participants could adjust the scores via mouse click according to their own perception.

2.3. EEG recording and analysis

EEG data were recorded according to the international 10-20 system using the Cognionics wireless EEG headset with 20 dry electrodes and a sampling rate of 500 Hz (see Figure 03).

The EEG was grounded to the left mastoid and the impedance was below 2,500 kΩ at the beginning of the experiment. During offline pre-processing, signals were decorrelated, zero-padded, and referenced to mathematically linked mastoids (Nunez & Srinivasan, 2006). A notch filter and FIR (finite impulse response) band-pass filters with cut-off frequencies of 1 and 20 Hz were applied. Then, the signal was cut into 4 s long epochs starting from the presentation of the stimuli. A 200 ms baseline was extracted per epoch from the signal before stimulus onset. Epochs that had a maximum deviation of 250 µV in one of the frontal EEG channels (Fp1, Fp2) were removed. In addition, artifacts due to cardiac and muscle activity or eye movements were removed using an independent component analysis (ICA) (Chaumon et al., 2015; Hipp & Siegel, 2013; Lee et al., 1999). The independent component analysis (ICA) computes linearly independent components (IC) from the data, which can then be classified as artifact or true EEG signal based on their topology or power spectrum. In the next step, the power in the alpha (8 - 12 Hz) and theta (4 - 7 Hz) frequency bands was calculated using the Welch method, that is a modified version of the Fast Fourier Transform (FFT).

Figure 03: State-of-the-art EEG sensors with dry electrodes, which allow easy handling during preparation.

To evaluate the ERPs for the consistent and inconsistent feedbacks, the EEG signals were filtered with an FIR bandpass filter and narrow frequency band from 0.5 to 23 Hz. In addition, smaller 1-s epochs were chosen starting from the presentation of the estimated score (i.e., the feedback). Epochs were corrected to baseline by subtracting the mean amplitude of the baseline interval (200 ms before score onset). To identify differences in ERPs between feedback conditions for valence and cognitive effort scores, we used a cluster-based nonparametric randomization approach (Maris & Oostenveld, 2007). Clusters were identified as adjacent EEG channels and time points in the epoch, using a T-value-based cluster-level threshold of $p < .01$ and a group-level threshold of $p < .05$ (two-sided).

To quantify differences between feedback conditions (consistent and inconsistent), a one-factor repeated measures analysis of variance (rmANOVAs) was used with the dependent variables (1) perceived correctness (that is, the likelihood that the person will correct the feedback score), (2) reaction time and (3) accuracy on the subsequent trial.

2.4. ML-based decoding of cognitive and affective-emotional states

For the estimation of the mental states, we used those channels and frequency bands that are proposed in the literature for the calculation of indices associated with affective-emotional and cognitive states (F3 - alpha, Fz - theta, F4 - alpha, Pz - alpha) as well as the Hjorth measures of mobility (proportion of standard deviation of the frequency spectrum) and complexity (change within the frequency band) of the respective channels as predictive features. In a second step, we corrected the hypothesized annotations (based on the experimental condition) using the subjective ratings in order to predict the subjective evaluation. In the second prediction of the subjective ratings, the same neurophysiological signals were used as features.

The conditions were classified in pairs (HV-HWML vs. LV-HWML, HV-LWML vs. LV-HWML, LV-HWML vs. HV-LWML, LV-HWML vs. LV-LWML) and in a four-class problem (HV-HWML vs. HV-LWML vs. LV-HWML vs. LV-LWML).

The following supervised ML classifiers were implemented using scikit-learn (Pedregosa et al. 2011) and explored regarding their performance: (1) Logistic Regression (LR), (2) Support Vector Machine (SVM), k-Nearest Neighbor, (4) Random Forest Classifier (RFC), (5) Gradient Boosting Classifier (GBC), and Gaussian Naive Bayes (GNB). A dummy classifier with stratification as method was trained as an empirical baseline indicating a random prediction which also considers class distributions in the data. Hyperparameters were optimized in a randomized GridSearch based on the training set and with the balanced accuracy as evaluation metric. Classification accuracy of the classifiers was evaluated within a stratified 3-fold cross-validation individually for the participants using the balanced accuracy as evaluation metric. To obtain a distribution of average classification accuracy, we used a Monte Carlo simulation (MCS) by training the classifiers 100 times, each with a new train-test split (80:20) and model initiation.

3. Results

3.1. Decoding of neural correlates for the prediction of the conditions and subjective ratings

To answer how well we can decode the interaction of mental states using theoretically supported correlates, we compared several supervised ML methods. Our results show that we were able to discriminate experimental conditions with high classification accuracy (see Table 2 and Figure 5).

	Experimental condition			**Subjective ratings**		
Classifier	2.5th percentile	Mean	97.5th percentile	2.5th percentile	Mean	97.5th percentile
LR	0.626	0.923	1	0.15	0.344	0.542
SVM	0.642	0.917	1	0.15	0.352	0.558

KNN	0.562	0.865	1	0.133	0.353	0.592
RFC	0.597	0.866	1	0.133	0.334	0.550
GBC	0.569	0.865	1	0.117	0.327	0.558
GNB	0.532	0.835	1	0.133	0.339	0.567
Dummy	0.182	0.500	0.818	0.091	0.280	0.500

Table 1: Average classification accuracy based on the test set of classifiers compared to an empirical baseline. Left: Prediction of the experimental condition. Right: Prediction of the subjective ratings.

Thereby, the accuracy measures of the selected classifiers (LR, SVM, KNN, RFC, GBC and GNB) are significantly above an empirically estimated chance level (dummy classifier). The classifiers do not differ significantly in their classification accuracy (see Table 1). In a next step, we wanted to predict the subjective assessments, that are the subjective ratings, using the same algorithms and neurophysiological correlates. Interestingly, the classification accuracy drops to a chance level when the annotations are corrected based on the subjective ratings. Thus, influences not represented in the neurophysiological signals seem to affect a subjective assessment of experienced stimuli and perceived states.

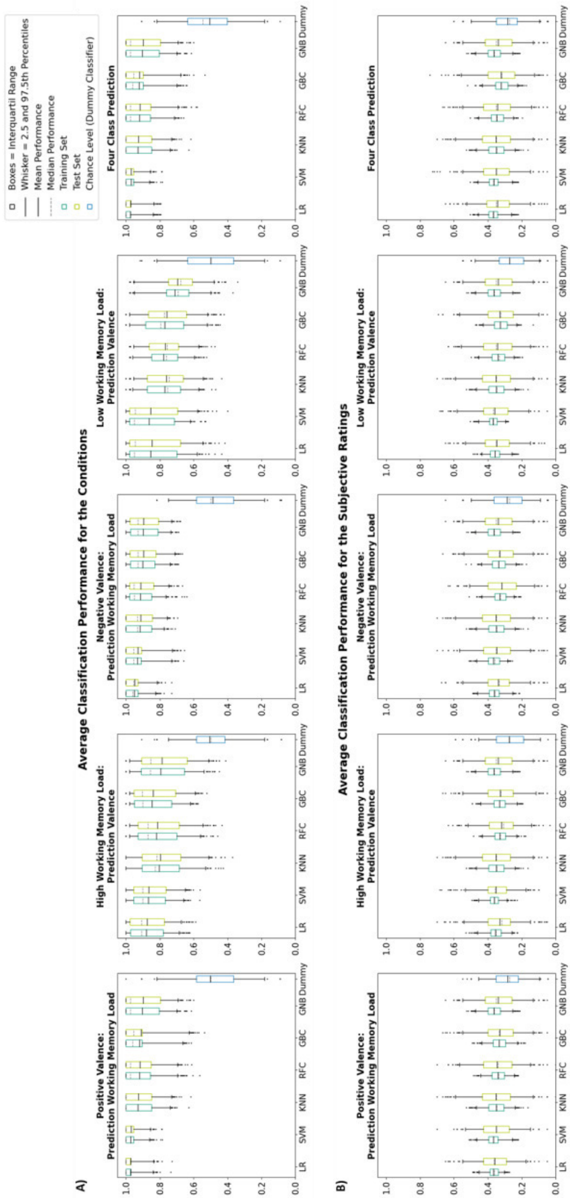

Figure 05: Average classification accuracy of the training set (green) and test set (light green) of participants and iterations compared to an empirical baseline (dummy classifier; blue). 2.5th and 97.5th percentiles of the simulation with 100 iterations. Top: Prediction of the experimental condition. Bottom: Prediction of the subjective ratings.

3.2. Influences of consistent and inconsistent feedback on performance-related measures

In a second study, we examined the effect of consistent and inconsistent feedback regarding recognized cognitive and affective-emotional states on performance (i.e., reaction time and accuracy). With regard to perceived accuracy, participants corrected inconsistent feedback of recognized values significantly more frequently, $F(1, 6) = 30.82$, $p < .001$, $p\eta^2 = .84$ (cognitive effort score) and $F(1, 6) = 5.14$, $p = .064$, $p\eta^2 = .46$ (affective-emotional state score). Inappropriate, inconsistent feedback had no significant effect on performance-related measures in the subsequent run. Experimental condition (e.g., task difficulty) had no effect on perceived accuracy of feedback and likelihood of correction. Analysis of the neurophysiological ERP responses did not identify significant clusters describing the difference between consistent and inconsistent feedback. In the semi-structured interviews, participants reported that they perceived the feedback scores as positive and interesting, but sometimes irritating. The design and feedback format, in the form of a barometer, was rated as appropriate and appealing. About half of the participants reported that they had not voluntarily used the feedback to change their strategy or behaviour. One person reported that he or she was motivated by the score and perceived it as promoting regarding the concentration. Some participants expressed a need for detailed clarification of the underlying calculations and measures used for the scores.

4. Discussion

Our results show that machine learning algorithms can distinguish different affective-emotional states and levels of cognitive load. There was no difference between the algorithms used. The finding that the hypothesized induced difficulty and valence (based on the experimental condition) can be learned and predicted with high accuracy from the neurophysiological data is of particular relevance; however, we could not predict the subjectively perceived difficulty and valence reported by participants in the questionnaires. This observation highlights the importance of objective methods for learner state recognition. Modulating effects, such as social desirability, processes of cognitive dissonance for self-image maintenance, or the capacity and ability to reflect on past experiences, can bias the self-assessment. These modulations are not represented in the neurophysiological signals measured simultaneously during the task and stimulus processing. The observed discrepancy between neurophysiological-estimated and subjectively perceived states could have relevant effects on the learner's trust in and, thus, the acceptance and effectiveness of a neuro-adaptive tutoring system. Future research on the integration of this discrepancy and the design of a tutoring system that is experienced as adequate is necessary. Furthermore, further research is needed to identify a suitable ground truth and associated calibration tasks that allow training of a tutor system.

For the implementation of neuro-adaptive tutoring systems in naturalistic environments with the goal of a high learner's acceptance of the systems, accuracy, and reliability of the online estimation of cognitive and affective-emotional states is of great importance. Even though offline methods could distinguish different affective-emotional states and levels of cognitive load with high accuracy, online state detections still showed a large variance in the accuracy of the detected states. To assess negative evaluation of inaccurate feedback, we examined the effect of inconsistent, inadequate as well as consistent, accurate feedback on a neural and behavioural level. Interestingly, we did not observe a negative effect of inconsistent or inaccurate feedback on participants' performance. However, this could be partly due to the small sample size of the exploratory study.

5. Conclusion

Our neurophysiological-based approach to capture learners' affective-emotional and cognitive states contributes to the development of closed-loop neuro-adaptive tutoring systems. These systems allow to monitor the learner's state, provide feedback, and adapt their system and learning parameters to individual abilities, needs and currently available resources (e.g., in terms of concentration). Optimal adaptation to the learner can contribute to an effective and positive learning experience. The design and validity of feedback is a major challenge for the effectiveness of feedback on performance-related measures. Therefore, influences of feedback formats should be explored in future research. In a next step, we aim to develop a methodology for robust detection and prediction of affective-emotional and cognitive states of learners during the learning and training session. This will require appropriate signal processing and artifact cleaning steps (filters, component analyses, etc.) to pre-process the signals, as well as computationally efficient, robust classifiers for real-time prediction of the experienced mental states (see Figure 06).

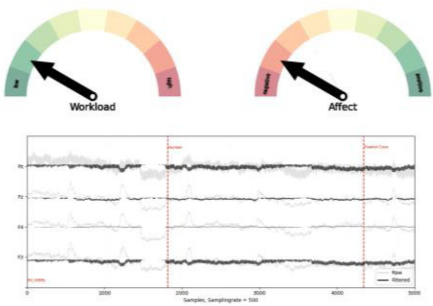

Figure 06: Outlook of a neurophysiological-based learner state recognition during a learning or training task.

References

Ahern, G. L., & Schwartz, G. E. (1985). Differential lateralization for positive and negative emotion in the human brain: EEG spectral analysis. *Neuropsychologia, 23,* 745–755. doi: 10.1016/0028-3932(85)90081-8

Alder, G. S., & Ambrose, M. L. (2005). Towards Understanding Fairness Judgments Associated with Computer Performance Monitoring: An Integration of the Feedback, Justice, and Monitoring Research. *Human Resource Management Review, 15*, 43-67. doi:10.1016/j.hrmr.2005.01.001

Appriou, A., Cichocki, A., and Lotte, F. (2020). Modern Machine-Learning Algorithms: For Classifying Cognitive and Affective States From Electroencephalography Signals. *IEEE Syst. Man Cybern. Mag.* 6, 29–38. doi: 10.1109/MSMC.2020.2968638

Babiloni, F. (2019, November). Mental workload monitoring: new perspectives from neuroscience. In *International Symposium on Human Mental Workload: Models and Applications* (pp. 3-19). Springer, Cham.

Bensch, M., Karim, A. A., Mellinger, J., Hinterberger, T., Tangermann, M., Bogdan, M., ... & Birbaumer, N. (2007). Nessi: an EEG-controlled web browser for severely paralyzed patients. *Computational intelligence and neuroscience*, *2007*.

Berntson, G. G., Thomas Bigger Jr, J., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., ... & Van der Molen, M. W. (1997). Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology*, *34*(6), 623-648.

Blankertz, B., Schäfer, C., Dornhege, G., & Curio, G. (2002, August). Single trial detection of EEG error potentials: A tool for increasing BCI transmission rates. In *International Conference on Artificial Neural Networks* (pp. 1137-1143). Springer, Berlin, Heidelberg.

Bowling, N. A., Alarcon, G. M., Bragg, C. B., & Hartman, M. J. (2015). A meta-analytic examination of the potential correlates and consequences of workload. *Work & Stress*, *29*(2), 95-113.

Bradley, M. M., & Lang, P. J. (2007). The International Affective Digitized Sounds (IADS-2): Affective ratings of sounds and instruction manual. *University of Florida, Gainesville, FL, Tech. Rep. B-3*.

Brauchle, D., Vukelić, M., Bauer, R., & Gharabaghi, A. (2015). Brain state-dependent robotic reaching movement with a multi-joint arm exoskeleton: combining brain-machine interfacing and robotic rehabilitation. *Frontiers in human neuroscience*, *9*, 564.

Briesemeister, B. B., Tamm, S., Heine, A., and Jacobs, A. M. (2013). Approach the Good, Withdraw from the Bad—A Review on Frontal Alpha Asymmetry Measures in Applied Psychological Research. *PSYCH* 04, 261–267. doi: 10.4236/psych.2013.43A039

Brouwer, A.-M., Hogervorst, M. A., van Erp, J. B. F., Heffelaar, T., Zimmerman, P. H., and Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. *J Neural Eng* 9, 45008. doi: 10.1088/1741-2560/9/4/045008

Buttfield, A., Ferrez, P. W., & Millan, J. R. (2006). Towards a robust BCI: error potentials and online learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *14*(2), 164-168.

Carlson, T. & Millan, J. D. R. (2013). Brain-Controlled Wheelchairs: A Robotic Architecture. *IEEE Robot. Autom. Mag., 20*, 65–73. doi:10.1109/MRA.2012.2229936.

Chaumon, M., Bishop, D. V. M., and Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *J Neurosci Methods* 250, 47–63. doi: 10.1016/j.jneumeth.2015.02.025.

Chen, M., Nikolaidis, S., Soh, H., Hsu, D., & Srinivasa, S. (2018, February). Planning with trust for human-robot collaboration. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 307-315).

Cromheeke, S., & Mueller, S. C. (2014). Probing emotional influences on cognitive control: an ALE meta-analysis of cognition emotion interactions. *Brain Structure and Function*, *219*(3), 995-1008.

Enriquez-Geppert, S., Huster, R. J., Herrmann, C. S. (2017). EEG-Neurofeedback as a Tool to Modulate Cognition and Behavior: A Review Tutorial. *Frontiers in Human Neuroscience*, *11*, 1-19. doi:10.3389/fnhum.2017.00051

Dessy, E., Van Puyvelde, M., Mairesse, O., Neyt, X., & Pattyn, N. (2018). Cognitive performance enhancement: do biofeedback and neurofeedback work?. *Journal of Cognitive Enhancement, 2*, 12-42. doi:10.1007/s41465-017-0039-y

DIN EN ISO 26800:2011-11 (2011). *Ergonomie - Genereller Ansatz, Prinzipien und Konzepte (ISO_26800:2011). Deutsche Fassung EN_ISO_26800:2011.* Berlin: Beuth Verlag GmbH.

Dolcos, F., and Denkova, E. (2014). Current Emotion Research in Cognitive Neuroscience: Linking Enhancing and Impairing Effects of Emotion on Cognition. *Emotion Review* 6, 362–375. doi: 10.1177/1754073914536449

Fairclough, S. H. (2009). Fundamentals of physiological computing. *Interacting with Computers* 21, 133–145. doi: 10.1016/j.intcom.2008.10.011

Ferrez, P. W., & Millán, J. D. R. (2008). Error-related EEG potentials generated during simulated brain–computer interaction. *IEEE transactions on biomedical engineering*, *55*(3), 923-929.

Gevins, A., & Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, *4*(1-2), 113-131.

Groen, Y., Wijers, A. A., Mulder, L. J., Minderaa, R. B., & Althaus, M. (2007). Physiological correlates of learning by performance feedback in children: a study of EEG event-related potentials and evoked heart rate. *Biological Psychology*, *76*(3), 174-187.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in psychology* (vol. 52, pp. 139-183). North-Holland.

Hipp, J. F., and Siegel, M. (2013). Dissociating neuronal gamma-band activity from cranial and ocular muscle activity in EEG. *Front Hum Neurosci* 7, 338. doi: 10.3389/fnhum.2013.00338

Hoonakker, P., Carayon, P., Gurses, A. P., Brown, R., Khunlertkit, A., McGuire, K., & Walker, J. M. (2011). Measuring workload of ICU nurses with a questionnaire survey: the NASA Task Load Index (TLX). *IIE transactions on healthcare systems engineering*, *1*, 131-143.

Ihme, K., Unni, A., Zhang, M., Rieger, J. W. & Jipp, M. (2018) Recognizing Frustration of Drivers From Face Video Recordings and Brain Activation Measurements With Functional Near-Infrared Spectroscopy. *Front. Hum. Neurosci. 12,* 327. doi: 10.3389/fnhum.2018.00327

Iordan, A. D., Dolcos, S., and Dolcos, F. (2013). Neural signatures of the response to emotional distraction: a review of evidence from brain imaging investigations. *Front Hum Neurosci* 7, 200. doi: 10.3389/fnhum.2013.00200

King, J. R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: the temporal generalization method. *Trends in cognitive sciences*, *18*(4), 203-210.

Kosuru, R. K., Lingelbach, K., Bui, M., & Vukelić, M. (2019). MindTrain: How to train your mind with interactive technologies. In *Proceedings of Mensch und Computer 2019* (pp. 643-647). doi:10.1145/3340764.3344884.

Lee, T. W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput* 11, 417–441. doi: 10.1162/089976699300016719

Leeb, R., Tonin, L., Rohm, M., Desideri, L., Carlson, T., & Millan, J. D. R. (2015). Towards independence: a BCI telepresence robot for people with severe motor disabilities. *Proceedings of the IEEE*, *103*(6), 969-982. doi: 10.1109/JPROC.2015.2419736.

Logemann, H. N. A., Lansbergen, M. M., Van Os, T. W. D. P., Böcker, K. B. E., & Kenemans, J. L. (2010). The Effectiveness of EEG-Feedback on Attention, Impulsivity and EEG: A Sham Feedback Controlled Study. *Neuroscience Letters, 479*, 49-53. doi:10.1016/j.neulet.2010.05.026

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods*, *164*(1), 177-190.

Master, R., Jiang, X., Khasawneh, M. T., Bowling, S. R., Grimes, L., Gramopadhye, A. K., & Melloy, B. J. (2005). Measurement of trust over time in hybrid inspection systems. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *15*(2), 177-196.

Mattout, J., Perrin, M., Bertrand, O., & Maby, E. (2015). Improving BCI performance through co-adaptation: Applications to the P300-speller. *Annals of physical and rehabilitation medicine*, *58*(1), 23-28.

Morawetz, C., Riedel, M. C., Salo, T., Berboth, S., Eickhoff, S. B., Laird, A. R., et al. (2020). Multiple large-scale neural networks underlying emotion regulation. *Neurosci Biobehav Rev* 116, 382–395. doi: 10.1016/j.neubiorev.2020.07.001

Nakamura, J., & Csikszentmihalyi, M. (2009). Flow theory and research. *Handbook of positive psychology*, 195-206.

Niklas, C. D., & Dormann, C. (2005). The impact of state affect on job satisfaction. *European Journal of Work and Organizational Psychology, 14*, 367-388. https://doi.org/10.1080/13594320500348880.

Nisbett, R. E. & Wilson, T., D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-258.

Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain: The neurophysics of EEG*. Oxford: Oxford Univ. Press.

Okon-Singer, H., Hendler, T., Pessoa, L., & Shackman, A. J. (2015). The neurobiology of emotion-cognition interactions: fundamental questions and strategies for future research. *Front Hum Neurosci* 9, 58. doi: 10.3389/fnhum.2015.00058

Parasuraman, R., & Rizzo, M. (2008). *Human Technology Interaction Series: Neuroergonomics*. Oxford: Oxford University Press.

Pedregosa, F., Varoquaux, G., Gramfort, A., Vincent, M., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., & Weiss, R. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*(12), 2825–2830.

Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. T. A. (2011). 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.

Picard, R. W. (2000). *Affective computing*. Cambridge, Massachusetts, London: The MIT Press.

Pomplun, M., & Sunkara, S. (2003, June). Pupil dilation as an indicator of cognitive workload in human-computer interaction. In *Proceedings of the International Conference on HCI* (Vol. 273).

Roth, W. T. (1983). A Comparison of P300 and Skin Conductance Response. In *Advances in Psychology* (Vol. 10, pp. 177-199). North-Holland.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Scherer, K. R., & Ceschi, G. (2000). Studying affective communication in the airport: The case of lost baggage claims. *Personality and Social Psychology Bulletin*, *26*, 327-339.

Seleznov, I., Zyma, I., Kiyono, K., Tukaev, S., Popov, A, Chernykh, M. & Shpenkov, O. (2019) Detrended Fluctuation, Coherence, & Spectral Power Analysis of Activation Rearrangement in EEG Dynamics During Cognitive Workload. *Front. Hum. Neurosci., 13,* 270. doi: 10.3389/fnhum.2019.00270

Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., ... & Yang, X. (2018). A review of emotion recognition using physiological signals. *Sensors*, *18*(7), 2074.

Shockley, K. M., Ispas, D., Rossi, M. E., & Levine, E. L. (2012). A meta-analytic investigation of the relationship between state affect, discrete emotions, and job performance. *Human Performance, 25*(5), 377-411. https://doi.org/10.1080/08959285.2012.721832.

Smith, E. E., Reznik, S. J., Stewart, J. L., & Allen, J. J. (2017). Assessing and conceptualizing frontal EEG asymmetry: An updated primer on recording, processing, analyzing, and interpreting frontal alpha asymmetry. *International Journal of Psychophysiology*, *111*, 98-114.

Toet, A., & van Erp, J. B. (2019). The EmojiGrid as a tool to assess experienced and perceived emotions. *Psych*, *1*(1), 469-481.

Verma, G. K., & Tiwary, U. S. (2014). Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiological signals. *NeuroImage*, *102*, 162-172.

Vukelić, M., Lingelbach, K., & Piechnik, D. (2021). *Feinfühlige Technik: Wie Neuroergonomie und Brain-Computer-Interfaces in der Praxis eingesetzt werden können.* Fraunhofer Gesellschaft zur Förderung der angewandten Forschung.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.*, *113*, 767–791.

World Economic Forum. Centre for the New Economy and Society Boston Consulting Group (BCG). (2019). *Towards a reskilling revolution: industry-led action for the future of work*. World Economic Forum, Geneva, Switzerland. ISBN: 978-2-940631-08-7

Yu, B., Funk, M., Hu, J., Wang, Q., & Feijs, L. (2018). Biofeedback for everyday stress management: A systematic review. *Frontiers in ICT*, *5,* 23. https://doi.org/10.3389/fict.2018.00023.

Zander, T. O., & Kothe, C. (2011). Towards passive brain–computer interfaces: Applying brain–computer interface technology to human–machine systems in general. *Journal of neural engineering*, *8*(2), 025005.

Zinchenko, A., Kotz, S. A., Schröger, E., & Kanske, P. (2020). Moving towards dynamics: Emotional modulation of cognitive and emotional control. *Int J Psychophysiol* 147, 193–201. doi: 10.1016/j.ijpsycho.2019.10.018

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: mental workload in ergonomics. *Ergonomics*, *58*(1), 1-17.

Young, M. S., & Stanton, N. A. (2002). Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. *Human Factors, 44,* 365–375.