

# Prediction of Difficulty Levels in Video Games from Ongoing EEG

Laura Naumann<sup>1,2(✉)</sup>, Matthias Schultze-Kraft<sup>3,4</sup>, Sven Dähne<sup>2,5</sup>,  
and Benjamin Blankertz<sup>3,4</sup>

<sup>1</sup> Bernstein Center for Computational Neuroscience Berlin, Berlin, Germany  
`laura-bella.naumann@bccn-berlin.de`

<sup>2</sup> Berlin Big Data Center, Berlin, Germany

<sup>3</sup> Bernstein Focus: Neurotechnology, Berlin, Germany

<sup>4</sup> Neurotechnology Group, Technische Universität Berlin, Berlin, Germany

<sup>5</sup> Machine Learning Group, Technische Universität Berlin, Berlin, Germany

**Abstract.** Real-time assessment of mental workload from EEG plays an important role in enhancing symbiotic interaction of human operators in immersive environments. In this study we thus aimed at predicting the difficulty level of a video game a person is playing at a particular moment from the ongoing EEG activity. Therefore, we made use of power modulations in the theta (4–7 Hz) and alpha (8–13 Hz) frequency bands of the EEG which are known to reflect cognitive workload. Since the goal was to predict from multiple difficulty levels, established binary classification approaches are futile. Here, we employ a novel spatial filtering method (SPoC) that finds spatial filters such that their corresponding bandpower dynamics maximally covary with a given target variable, in this case the difficulty level. EEG was recorded from 6 participants playing a modified Tetris game at 10 different difficulty levels. We found that our approach predicted the levels with high accuracy, yielding a mean prediction error of less than one level.

**Keywords:** BCI · Cognitive workload · Video games · EEG · Machine learning · Spatial filtering

## 1 Introduction

While the original and predominant goal of brain-computer interfaces (BCIs) has been to provide a channel for communication and control [1], in recent years BCI research has expanded towards applications that aim for the detection of covert mental states [2], with one focus lying on the assessment of *cognitive workload*. A reliable assessment of cognitive workload from EEG has been suggested as a means for enhancing human-machine interaction in everyday environments [3] and is thus an endeavor that fosters the development of future symbiotic systems.

Typical EEG estimators of workload are based on the fact that changes in workload are associated with modulations in the power of oscillatory activity in particular frequency bands of the EEG, the most prominent ones being theta

(4–7 Hz) and alpha (8–13 Hz). Theta power has been shown to be positively correlated with workload, most notably over frontal regions, whereas alpha power is typically found to be negatively correlated with workload, in particular over parietal regions [4, 5]. In typical BCI studies that aim for workload assessment, the experimental paradigm introduces two conditions which induce two levels of workload, e.g. low and high [6, 7]. This represents a classical classification setting where common practice is to train a linear classifier on EEG features using the binary labels. A well-established approach for the extraction of EEG features in this setting is the Common Spatial Patterns (CSP) analysis [8] which finds spatial filters that maximize the power contrast between the two classes and has become one of the corner stones of sensorimotor rhythm based BCIs.

In this study, we aimed at predicting the difficulty level of a video game that participants were playing from the ongoing EEG. In a typical video game that requires continuous mental and visuomotor effort, an increase of difficulty is expected to increase the player’s cognitive workload. Therefore, it seems obvious to employ a BCI that uses workload induced power modulations of the theta and alpha frequency band as neurophysiological markers for prediction. However, if the goal is to predict from multiple difficulty levels, we leave the regime of binary classification, thus foiling the use of CSP as a method for extracting workload induced power modulations. Here, we employ a novel method called the Source Power Co-Modulation (SPoC) analysis that finds spatial filters such that their corresponding bandpower dynamics maximally covary with a given target variable [9], in this case the difficulty level. A recent study that employed a computer game-like experimental task, demonstrated that SPoC can be used to detect workload states from the EEG with high accuracy [10]. In that study, the task induced two levels of workload and the participant’s error rate was considered an indirect measure of workload and thus used as a continuous target variable for SPoC. In this study, we decided to directly use multiple difficulty levels in a video game from a predefined continuum as target. A suitable candidate game for the proposed experiment is the classic Tetris, because the game’s difficulty can be easily adjusted by changing the falling velocity of the items. In the conducted experiment, participants played the game several times while EEG data were recorded. In a subsequent analysis, we first selected two crucial model parameters for SPoC and finally evaluated our approach for the prediction of difficulty levels in a cross-validation.

## 2 Methods

### 2.1 The Experiment

Six healthy participants (4 female) were engaged in the experiment. Recruited participants were lab members or friends, who gave their informed oral and written consent and did not receive monetary compensation. The average age of the participants was 24 years with a standard deviation of 1.5 years, all of them had normal or corrected to normal vision and previous experience with EEG recordings. Participants were seated in front of a computer monitor and

asked to play a Tetris clone<sup>1</sup> with their right hand using the arrow keys of a keyboard (Fig. 1a). The game was implemented in the Python package `pygame` and modified as follows:

1. The current score, the options to restart, pause or quit and the preview of the next item were removed
2. If the blocks reached the top of the field, all blocks from the top half of the field were removed, as shown in Fig. 1b
3. The fast-drop function of the classic Tetris was disabled

Modification 1 aimed to minimize eye movements caused by distracting objects on the screen. The purpose of modification 2 was to prevent the premature end of the game that normally occurs when the top of the field is reached, and modification 3 was implemented to compensate differences in the players performance and strategies, thus keeping the difficulty within one level constant across participants.

The experiment was structured as follows: Each participant played a total of 9 games with 10 min each. Every game consisted of 10 consecutive 60-second blocks, of which each was assigned one of 10 predefined difficulty levels. In games 1, 4 and 7 the levels increased gradually with every block, while in the remaining games the assignment was randomly permuted. The difficulty of a level was defined by the items' falling velocity: In level 1 the velocity was such that it required falling objects 10s to reach the bottom of the game screen, while in level 10 this time was 2.5s. The speed increase across levels was linear. A scoring system was implemented where the points for canceling a line are proportional to the level number and getting the field reduced takes points away. The score was presented to the player at the end of each game (Fig. 1c).

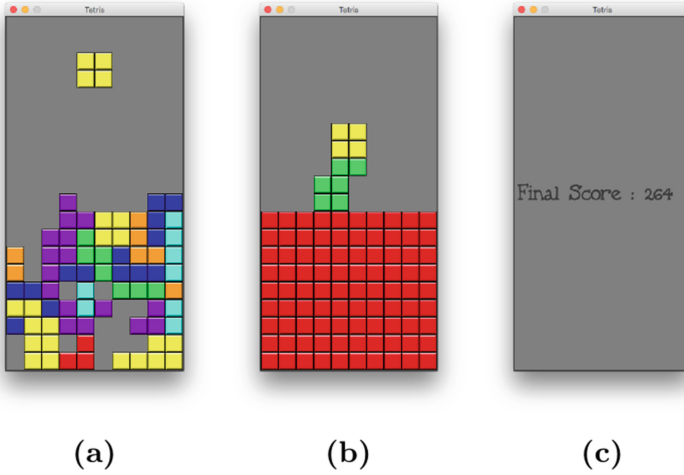
To motivate an active participation in the game, the player with the highest score was rewarded a small price. Apart from that, participation in the experiment was voluntary and participants received no further compensation.

## 2.2 Data Acquisition and Preprocessing

EEG data was recorded at 1000 Hz using BrainAmp amplifiers and a 32 electrode Fast'n'Easy Cap (Brain Products GmbH, Gilching, Germany) and referenced to an electrode positioned at the *FCz* electrode location. Furthermore, an electrooculogram (EOG) electrode was placed under the right eye. During recording, markers were sent via a parallel port to record the change to a different level of difficulty.

Before the analysis, the continuous EEG data was high-pass filtered at 1 Hz and subsequently segmented into epochs of 50s length, starting 10s after the onset of a new level indicated by the level markers. By discarding the first 10s of each difficulty level, we aimed at mitigating possible transition effects. The experimental tasks involved persistent movement of the eyes and thus the recorded data was expected to contain ocular artifacts. Therefore, horizontal and vertical

<sup>1</sup> <http://pygame.org/project-yayatc-1647-.html>.



**Fig. 1.** Snapshots from the modified Tetris game used as a task. (a) Ongoing game with random items falling from the top. (b) Reduction of the field after reaching the top. (c) The score is displayed at the end of every game.

eye movements were estimated by means of the difference of electrodes  $F7$  and  $F8$  and the difference between electrode  $Fp2$  and the EOG electrode, respectively. The estimated contribution of these combined signals to all remaining channels was removed using linear regression. The four electrodes were excluded from all further analyses. For symmetry reasons, electrode  $Fp1$  was also removed since it showed very high synchronization with  $Fp2$ . Remaining artifacts were eliminated by rejecting single epochs based on an excess variance criterion. Finally, the epochs were band-pass filtered in different frequency bands within the theta and alpha frequency range (see Fig. 2).

### 2.3 Source Power Co-modulation (SPoC) Analysis

The preprocessed data was analysed using the SPoC analysis proposed by Dähne et al. [9]. SPoC optimizes a set of spatial filters  $\mathbf{w}$  based on a presumed covariation between band power dynamics and an external target signal, which in the present study corresponds to the level of difficulty. Let the variable  $z$  denote the mean-free target signal and let  $\mathbf{x}_f(t)$  denote the EEG signal band-pass filtered in the frequency  $f$ . Then the time-resolved component bandpower can be expressed in terms of the spatial filter  $\mathbf{w}$  by computing the variance in consecutive time epochs, indexed by  $e$ , as

$$\phi_f(e) \approx \text{Var}[\mathbf{w}^\top \mathbf{x}_f(t)] = \mathbf{w}^\top \mathbf{C}_f(e) \mathbf{w} , \quad (1)$$

where  $\mathbf{C}_f(e)$  denotes the covariance matrix of the bandpass filtered data within the  $e$ th epoch. SPoC then extracts the  $K$  best spatial filters solving the optimization problem

$$\min_{\mathbf{w}} \text{Cov}[\phi_f(e), z(e)] = \min_{\mathbf{w}} \mathbf{w}^\top \langle \mathbf{C}_f(e) z(e) \rangle \mathbf{w} \quad (2)$$

under the constraint that extracted time courses have unit variance and are mutually decorrelated. The optimization problem can be expressed as a generalized eigenvalue problem that can be solved efficiently using standard linear algebra libraries<sup>2</sup>.

## 2.4 Level Prediction via Regression

Having extracted the spatial filters with SPoC, the log-var features corresponding to filters  $i \in [1, \dots, K]$  from frequency bands  $f \in [\theta, \alpha]$  are eventually combined by means of *linear regression* into a single predictive variable. Both the specific  $\theta$  and  $\alpha$  frequency bands and the number  $K$  of SPoC components used for the prediction are important parameters of the method that have to be determined by means of validation. In order to prevent overfitting, we determined the prediction performance in a cross-validation, where only the training set is used to derive and combine the SPoC components and the correlation to the true levels is calculated subsequently on the test dataset alone. To account for potential non-stationarities in the EEG data, the *leave-one-block-out* procedure was used for validation with every full Tetris game being one block. All preprocessing and analysis steps were conducted using the BBCI toolbox<sup>3</sup> for MATLAB.

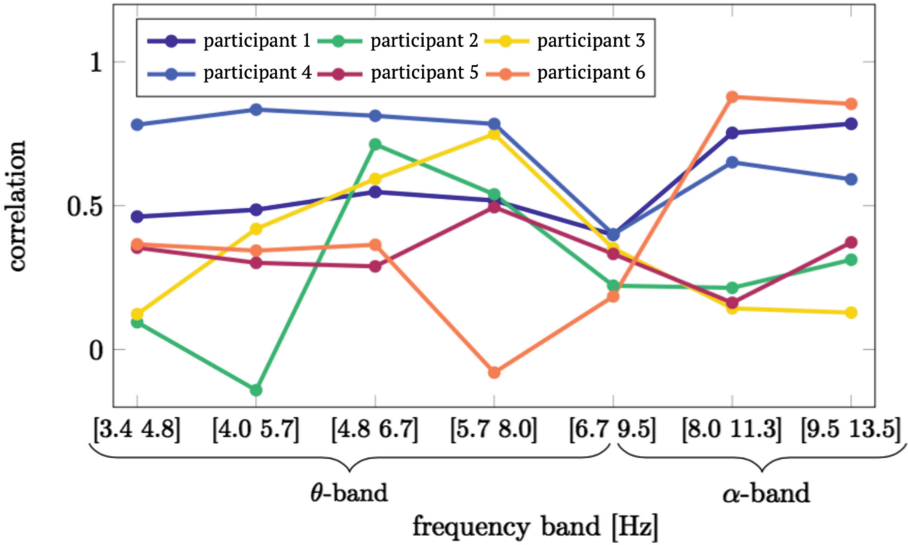
## 3 Results

### 3.1 Optimal Frequency Bands

As a first step, we aimed to determine the optimal theta and alpha frequency band. Since inter-individual differences in the frequency range of the alpha and theta band are expected, the bands rendering the best performance were determined for every participant individually. For that purpose a sweep over frequency bands ranging from around 3 to 14 Hz using only one SPoC component was performed. The width of the bands scaled logarithmically with the frequency to obtain a higher resolution for lower frequencies. Figure 2 shows the correlation of predicted with the true levels for every participant as a function of frequency band. While for participants 1 and 6 the band that yields the best prediction is in the alpha range, for participants 3 to 5 the best prediction is achieved in the theta range. Interestingly, a decline of correlation can be observed for the 6.7–9.5 Hz band, which comprises the transition between the alpha and theta range. For participant 2, a first analysis showed very poor correlations across all frequencies. A closer inspection of the bandpower of the data projected onto the first three SPoC components over the time revealed that three outliers that were not detected by the artifact rejection procedure during the preprocessing

<sup>2</sup> <https://github.com/svendaehne/matlab.SPoC/releases/latest>.

<sup>3</sup> [https://github.com/bbci/bbci\\_public](https://github.com/bbci/bbci_public).



**Fig. 2.** Frequency sweep with SPoC. Shown is, for each participant individually, the correlation between true levels and predictions as a function of frequency band used for SPoC analysis with one output component. (Color figure online)

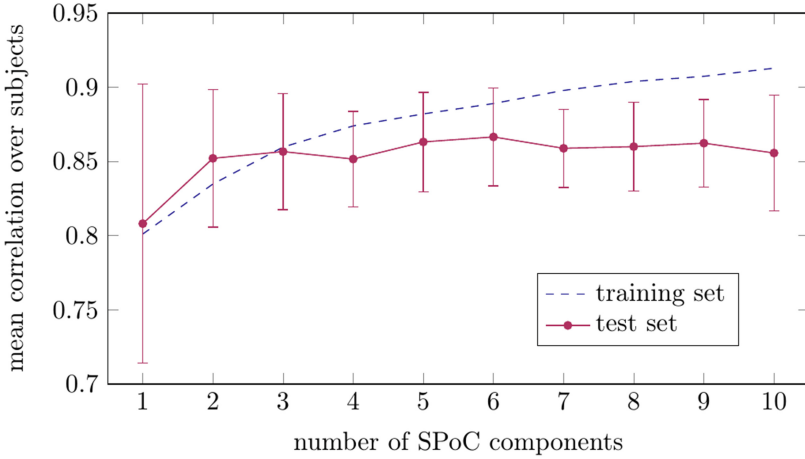
strongly dominated the SPoC filters. After removing the epochs corresponding to these outliers, correlations comparable to those of other participants were obtained (Fig. 2, green line). From the results of every participant, the alpha and theta frequency bands with the highest prediction performance were selected for further analyses.

### 3.2 Number of SPoC Components

For both selected theta and alpha bands, the SPoC method returns a set of components. To determine how many of those are sufficient to yield a high prediction performance for all participants, the cross-validated correlation averaged over all individuals was computed for increasing number of SPoC components used for the prediction. The results on the training and test set are shown in Fig. 3. For one component, the mean correlation is already above 0.8 but shows a high variability. When two components are used, the correlation increases to about 0.85 but stays more or less constant for higher numbers. The mean correlation on the training set indicates overfitting for more than three SPoC features and hence the number of components for the further analysis was set to three.

### 3.3 Interpretation of Components

We next investigated whether the components found by SPoC are of cortical origin or whether they stem from movement or residual ocular artifacts.



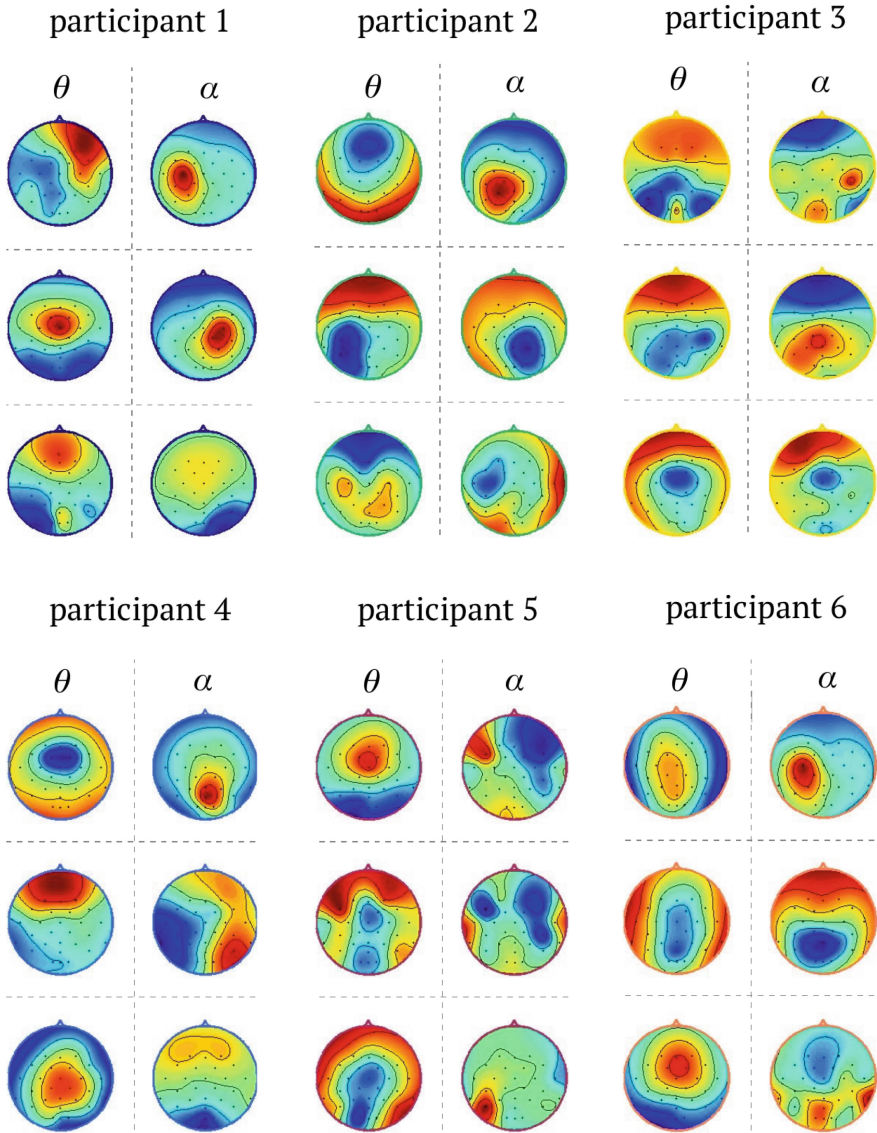
**Fig. 3.** Crossvalidated correlation as function of the number of SPoC components, averaged over all participants. The standard error is shown with bars for the test set.

For this, the activation patterns corresponding to the spatial filters from the SPoC method were examined (Fig. 4). We found that all investigated patterns are physiologically plausible and show none of the characteristics of patterns related to EOG or electromyogram (EMG) activity. When examining SPoC components in the theta range, for each participant at least one component shows a characteristic theta mid-frontal component (e.g. component 2 for participant 1, and component 1 for participant 2). Regarding the patterns in the alpha range, we find that for many participants SPoC found components with a clear lateralization (e.g. components 1 and 2 for participant 1, and component 1 for participant 6), and components with a centro-parietal topology (e.g. component 2 for both participants 3 and 6).

### 3.4 Level Predictions

For each participant individually, we evaluated our approach for difficulty level prediction using the optimal, participant-specific theta and alpha frequency bands and three SPoC components. The single and grand average predicted levels as a function of true levels are illustrated in Fig. 5. This qualitative assessment shows that predictions roughly cover the levels from 1 to 10. While the mean across participants indeed shows a monotonic increase with the true levels, for single participants this does not hold in some cases when comparing neighboring levels. Nonetheless, the relationship between the predicted and true levels appears to be clearly linear.

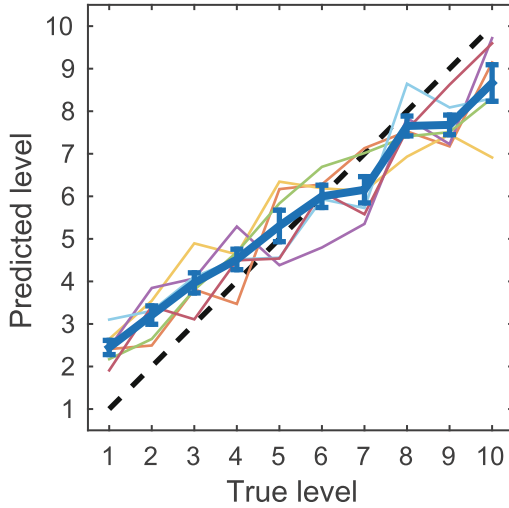
Next, we conducted a quantitative evaluation of prediction performance using three different measures between the predicted and the true levels: (i) The Spearman correlation, (ii) the Pearson correlation, and (iii) the Root Mean Squared Error (RMSE). While the Pearson correlation assumes a linear relation between



**Fig. 4.** Spatial activation patterns corresponding to the spatial filters found by SPoC for all six participants. For each person, the left and right columns show the activation patterns for the optimal theta ( $\theta$ ) and alpha ( $\alpha$ ) band, respectively, and ordered from best to third best, from top to bottom.

the two variables, the Spearman correlation relaxes this assumption and assesses how well their relation is described by any monotonic function. The RMSE allows to estimate how far off a prediction is in units of levels. The mean of the three measures, as well as the grand average are shown in Table 1. Correlations between





**Fig. 5.** Prediction of difficulty levels averaged across all games, for single participants (light colors) and grand average and standard error (blue). The dashed line indicates a perfect prediction. (Color figure online)

predicted and true level are statistically significant for all participants as revealed by a permutation analysis with 250 repetitions at a significance level of 0.01. No considerable difference was found between the two correlation measures, corroborating the linearity of the relationship between predicted and true levels. Using the RMSE as a measure finally shows that across all participants and levels, our approach yields predictions that are, on average, about two thirds of a level off the true level.

**Table 1.** Prediction of difficulty level for single participants, assessed with three measures between predicted and true levels: Pearson correlation, Spearman correlation and RMSE. The last row shows the average across participants and the standard error.

| Participant | Pearson correlation | Spearman correlation | RMSE (in levels) |
|-------------|---------------------|----------------------|------------------|
| 1           | 0.90                | 0.90                 | 0.62             |
| 2           | 0.81                | 0.81                 | 0.78             |
| 3           | 0.84                | 0.85                 | 0.63             |
| 4           | 0.82                | 0.83                 | 0.69             |
| 5           | 0.82                | 0.85                 | 0.75             |
| 6           | 0.89                | 0.91                 | 0.55             |
| Mean        | $0.85 \pm 0.02$     | $0.86 \pm 0.02$      | $0.67 \pm 0.04$  |

## 4 Discussion

The goal of this study was to predict the level of difficulty in a Tetris video game from power modulations in different EEG frequency bands associated with a change in the participants workload. We therefore employed the state-of-the-art spatial filtering method SPoC [9] in order to extract relevant features from the EEG. Whereas established methods like CSP are limited to settings with two classes, SPoC can use any continuous measure or signal as a target variable. We found that our approach predicted difficulty levels with high accuracy, yielding mean correlations of 0.85 and a mean prediction error of less than one level.

An inspection of both the optimal frequency band and the patterns corresponding to the spatial filters found by SPoC allowed us to study which signals contributed to the prediction. Regarding the theta band, we find that the characteristic mid-frontal component is consistently represented in all participants. This is in line with findings which show that this topology is associated both with changes in operator workload [10] and reflects time pressure effects on visuomotor tasks [11]. The results from the frequency sweep analysis further suggest that for participants 1 and 6 alpha power modulations had higher predictive value than in the theta band. Interestingly, for those two participants SPoC found alpha band components that were lateralized over motor regions. Thus, for participants 1 and 2 the contribution to the prediction from the alpha band presumably reflected modulations of the sensorimotor rhythm (SMR) caused by increasing motor demands during high difficulty levels and were thus not directly related to cognitive workload. We therefore repeated our analysis twice, one time using only theta band components and one time using only alpha band components. We found that while in the theta-only analysis a mean correlation of  $0.79 \pm 0.04$  was achieved, which was lower but still comparable to the dual-band case, in the alpha-only analysis the mean correlation dropped considerably to  $0.63 \pm 0.10$  and was not significant for all participants as revealed by a permutation analysis.

A conspicuous observation in the predicted levels in Fig. 5 is that low difficulty levels tend to be overestimated, while high level tend to be underestimated. Thus, the ability of our model to generalize seems to depend on where the levels lie within the continuum of difficulty. Furthermore, it is worth noting that the participants did not practice the Tetris game before the experiment and there was no individual adaptation of the minimal and maximal difficulty level. Such a calibration could ensure a more constant contrast of gaming experience across participants and thus allow for a better comparability.

Our approach allows for a continuous measurement from ongoing EEG activity without intervening with the participants' engagement with the video game. This is in contrast to alternative approaches that rely on the evocation of event-related potentials by means of secondary stimuli [12]. Our approach can furthermore be extended to online application e.g. for adaptive learning systems to adjust progress and presentation of material. With this study we demonstrated the use of novel machine learning techniques for an EEG-based assessment of mental states, thus advancing the endeavor to enhance the symbiotic interaction between human operators and machines.

**Acknowledgements.** The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 611570 (MindSee), and in part by the BMBF (contract 01GQ0850, BFNT).

## References

1. Dornhege, G., del R. Millán, J., Hinterberger, T., McFarland, D.J., Müller, K.-R.: *Toward Brain-Computer Interfacing*. MIT Press, Cambridge (2007)
2. Blankertz, B., Tangermann, M., Vidaurre, C., Fazli, S., Sannelli, C., Haufe, S., Maeder, C., Ramsey, L.E., Sturm, I., Curio, G., Müller, K.-R.: The Berlin brain-computer interface: non-medical uses of BCI technology. *Front. Neurosci.* **4**, 198 (2010)
3. Parasuraman, R., Wilson, G.F.: Putting the brain to work: neuroergonomics past, present, and future. *Hum. Factors* **50**(3), 468–474 (2008)
4. Holm, A., Lukander, K., Korpela, J., Sallinen, M., Müller, K.M.I.: Estimating brain load from the EEG. *Sci. World J.* **9**, 639–651 (2009)
5. Gevins, A., Smith, M.E.: Neurophysiological measures of cognitive workload during human-computer interaction. *Theor. Issues Ergon. Sci.* **4**, 113–131 (2003)
6. Kohlmorgen, J., Dornhege, G., Braun, M., Hinterberger, T., McFarland, D., Müller K.-R.: Improving human performance in a real operating environment through real-time mental workload detection. In: Dornhege, G., del R. Millán, J., Hinterberger, T., McFarland, T.D., Müller, K.-R. (eds.), *Toward Brain-Computer Interfacing*, pp. 409–422. MIT press, Cambridge, MA (2007)
7. Dijksterhuis, C., de Waard, D., Brookhuis, K., Mulder, B., de Jong, R.: Classifying visuomotor workload in a driving simulator using subject specific spatial brain patterns. *Front. Neurosci.* **7** (2013)
8. Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., Müller, K.-R.: Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Proc. Mag.* **25**, 41–56 (2008)
9. Dähne, S., Meinecke, F.C., Haufe, S., Höhne, J., Tangermann, M., Müller, K.-R., Nikulin, V.V.: SPoC: a novel framework for relating the amplitude of neuronal oscillations to behaviorally relevant parameters. *NeuroImage* **86**, 111–122 (2014)
10. Schultze-Kraft, M., Dähne, S., Gugler, M., Curio, G., Blankertz, B.: Unsupervised classification of operator workload from brain signals. *J. Neural Eng.* **13**, 036008 (2016)
11. Slobounov, S.M., Fukada, K., Simon, R., Rearick, M., Ray, W.: Neurophysiological and behavioral indices of time pressure effects on visuomotor task performance. *Cognitive Brain Res.* **9**, 287–298 (2000)
12. Allison, B.Z., Polich, J.: Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. *Biol. Psychol.* **77**, 277–283 (2008)

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

