Maritime Cognitive Workload Assessment

Daniel Miklody[1(✉)], Wendie M. Uitterhoeve[2], Dimitri van Heel[2], Kerstin Klinkenberg[3], and Benjamin Blankertz[1]

[1] Neurotechnology Group, Technische Universität Berlin, Sekr. MAR 4-3, Marchstr 23, 10587 Berlin, Germany
miklody@tu-berlin.de
[2] MARIN's Nautical Center, Wageningen, The Netherlands
[3] K+S Projects, Rangsdorf, Germany
http://www.neuro.tu-berlin.de/

Abstract. The human factor plays the key role for safety in many industrial and civil every-day operations in our technologized world. Human failure is more likely to cause accidents than technical failure, e.g. in the challenging job of tugboat captains. Here, cognitive workload is crucial, as its excess is a main cause of dangerous situations and accidents while being highly participant and situation dependent. However, knowing the captain's level of workload can help to improve man-machine interaction. The main contributions of this paper is a successful workload indication and a transfer of cognitive workload knowledge from laboratory to realistic settings.

Keywords: Workload · BCI · EEG

1 Introduction

In the maritime world, as in many other workplaces, working memory, the ability to process information and to take decision is crucial. The quantification of *cognitive workload* is a measure to study these aspects. The insight can shed light onto the limitations posed by the *human factor* and point out how to improve equipment, conditions or training. While there exists no generally accepted definition of cognitive workload, there is a large agreement that it encompasses the two concepts of *activation* and *resources* or capacity [15].

In the present investigation, we designed a *BCI system* based on spectral decompositions of electroencephalographic data [17] that is trained to detect states of high/low *cognitive workload*. The workload was manipulated by the main task itself within one of the conditions while in the other we employed the 2-back task [8] as a secondary task which is a common tool for the measurement of workload [14]. The aim of this study was to investigate cognitive workload in a more realistic environment and set the results into context with those from experiments obtained in clean laboratory settings. Therefore, professional tugboat captains were observed in a realistic training simulator study, where we investigated whether and how laboratory based cognitive workload studies are

L. Gamberini et al. (Eds.): Symbiotic 2016, LNCS 9961, pp. 102–114, 2017.
DOI: 10.1007/978-3-319-57753-1_9

transferable to more realistic settings. In particular, we investigated if and how alpha & theta oscillations are modulated by cognitive workload.

1.1 Neurophysilogical Correlates of Cognitive Workload

Cognitive workload is reflected in different components of brain activity. In view of the present target application, modulations of event-related potentials due to workload [11,12,18] are not relevant, since there are no controlled and continuously repeated stimuli. Therefore, we concentrate on workload-induced modulations of spontaneous brain activity.

The power of oscillatory brain activity in the theta frequency range (4 to 7 Hz) in frontal brain regions have been found to positively correlate with the level of workload, see e.g. [5,7,20].

With respect to the more prominent alpha frequency band, most studies report a negative correlation of cognitive workload and alpha power at parito-occipital scalp locations, see e.g. [4,7]. However, these studies used tasks in the visual modality to induce workload, such that one can only derive the implication of alpha reduction for workload in visual resources. In general, the functional role of alpha band oscillations is not yet conclusive. For a memory task in the auditory domain, [6] reports a modulation of theta oscillations only, but no modulations of the alpha rhythm. Some studies using auditory stimulation even found an increase of alpha activity with increasing workload [2,10,13,16]. A possible interpretation is provided by the hypothesis of functional inhibition, which postulates that strong alpha activity reflects active inhibition of task-irrelevant processes [9]: when the critical processing load is in the non-visual, the visual areas are actively deactivated. The idea to build EEG-based workload monitoring systems was presented, e.g. in [3,4,19].

2 Experimental Design

In a 10-participant simulator study, we recorded electroencephalographic data from a realistic tugboat scenario with professional captains (participant 8 excl.: sickness). The participants were recruited along the training network of MARIN and were compensated for their voluntary participation. They were all male with ages ranging from 30 to 65 years and different levels of experience. The experiment consisted of 3 different scenarios (approx. 40 min each), where scenario 1&3 were identical, see Fig. 1.

The simulator was a professional ship simulator bridge optimized for tugboat missions which can be observed in Fig. 2. It consisted of a 360° projected screen around a set of ordinary tugboat controls. This included several additional screens for radar and ship-parameters. The study was approved by the committee of the ethical department of Philips, the Netherlands, as we collaborated with them on this project. An informed written consent was obtained from all participants.

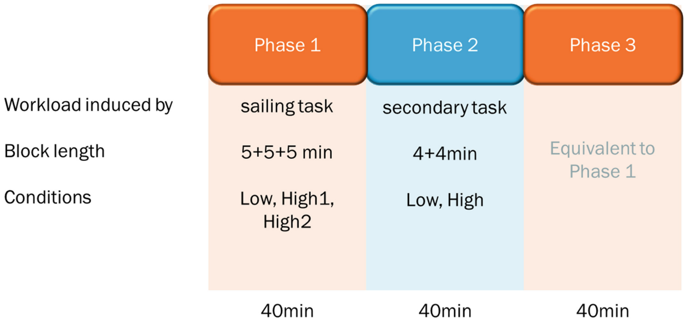

Fig. 1. Experimental design - overview

Fig. 2. Simulator Setting with 360° projections

While in phase 1&3, the cognitive workload was modulated by the sailing task itself in combination with environmental changes (weather, sea), we increased it in phase 2 by an additional task (2-back task [8]) and kept sailing constant.

2.1 Phase 1 and 3: Bow-to-Bow

In this scenario the focus was to keep the experimental conditions as naturalistic as possible while still being able to modulate the workload induced on the captain. Therefore, 3 conditions were generated with different tasks and different weather conditions. An overview about the temporal structure can be found in Fig. 3 on the left. For the later classification, the *high1* and *high2* epochs are combined to a common *high* class.

Condition 1: Free Sailing Condition: *low workload*: The captain was instructed to follow a large container ship astern while the weather was manipulated to have no extra effect on the workload.

Condition 2: Connecting Condition: *high1 workload*: After a transition phase moving to the front of the vessel the tugboat captain got the instruction to get ready for bow-to-bow connection while the weather conditions were changed to

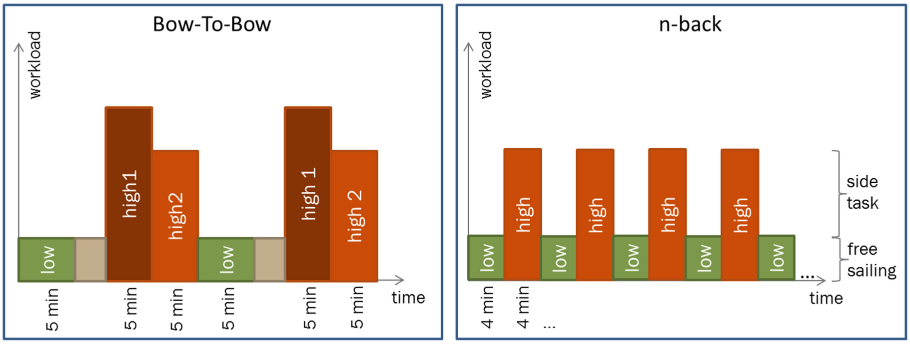

Fig. 3. Experimental design of workload modulation

harsh (wind, waves, fog). Then, the captains were told to wait for connection for 5 min which represents the high1 phase.

Condition 3: Pulling Condition: *high2 workload*: After the rope was connected, a constant tow force and line length was instructed. The weather stayed harsh.

2.2 Phase 2: n-Back

The n-back task is commonly used in neuroscientific research as a manipulation tool for cognitive workload. We used this as a secondary task to have a condition comparable to common research and to see how much our bow-to-bow scenario corresponds to the neural patterns of this commonly known task. We used an auditory 2-back task, where the participant had to follow a stream of spoken numbers. If the last number heard corresponded with the digit 2 back, they had to press a button. The digits 1–9 were used with 3 s interleave randomly (75%) and forced 2-back repetition (25%) to get a reasonable amount of repetitions. The 2-back was played auditorily to keep a realistic behavioral scheme of the captain. There were 2 conditions, 4 min each, which were repeated 5 times, resulting in a total duration of 40 min for the whole phase (see also Fig. 3 on the right):

Condition 1: Free Sailing Condition *low workload*: In this condition, the same low workload task of the bow-to-bow Scenario was induced for comparison.

Condition 2: Free Sailing with 2-Back Condition *high workload*: The 2-back task was used additionally to the Free Sailing to induce a higher workload while keeping the primary task constant.

3 EEG-Analysis

3.1 Dealing with Artifacts

A preliminary analysis of the data's spatio-spectral content showed that the EEG of some participants was heavily affected by artifacts. This was expected due to the participants being allowed to act naturally. Head and trunk movements were required for the sailing tasks, as the simulator provided a 360° projection.

First, the automatic artifact removal method MARA [22] had been employed, that gives good results in usual EEG datasets. The method is based on a decomposition of the multivariate EEG by the use of an Independent Component Analysis (ICA). The components were classified into artifacts and neuronal components. Then, the cleaned EEG signals were obtained by projecting only the neuronal components back into the sensor space. The classifier that distinguishes between artifactual and neuronal components was trained on a large data base of EEG datasets for which the ICA decomposition was manually annotated. For datasets that contain artifacts unlike those ones contained in the training data base, some of the artifactual components may go undetected. This seemed to be the case for the dataset at hand.

Therefore, we went the tedious way of annotating all ICA components (ICs) manually. This decision between artifactual and neuronal components was based on the following plots: the propagation pattern that corresponds to the IC, the time series of the IC and its power spectral density. Examples of those plots are given in Fig. 4.

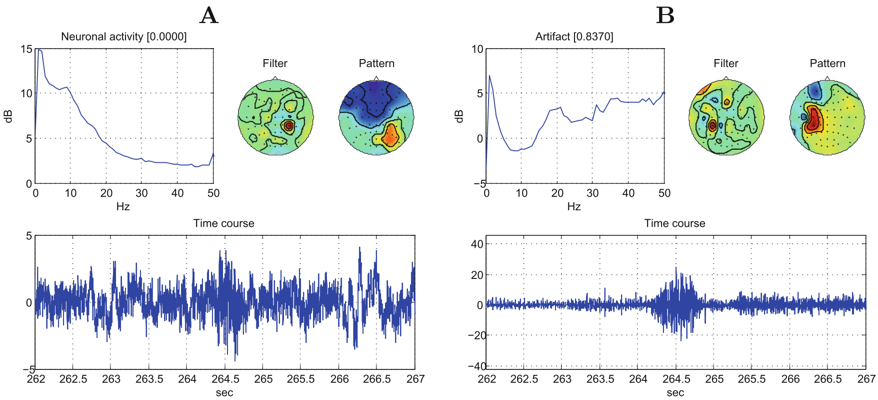

Fig. 4. Inspection of components obtained by ICA: **A** neuronal component: The pattern suggests a left arm motor area origin and shows a smooth dipolar structure. The power spectral density (upper left subplot) has the typical $1/f$ shape with enhanced power around 9.5 Hz, which is the typical frequency of the sensorimotor rhythm. There are no obvious irregularities in the time course. **B** artifactual component: Pattern, spectrum and time course do not look like neuronal activity: the pattern is very focal with no typical spectral $1/f$ shape, has least power in the 10 Hz range and strong power in high frequencies and the time course contains a high frequency burst.

3.2 Results

The grand average (i.e., average across participants) of the spectral analysis of the automatic artifact removal method MARA cleaned signals is shown in

Fig. 5 (2-back task) and Fig. 6 (bow-to-bow task). When comparing the results for individual participants, it became clear that the grand average is only of limited use. The effect of the workload conditions seen in the alpha band varies with respect to the specific frequency range: we observed individual frequencies and topologies at which alpha levels appeared characteristic for the individual participants - also across conditions.

For the *2-back* as a secondary tasks, effects in higher visual alpha can be observed across participants around parietal to occipital electrodes. The effect is contrary to common results in laboratory settings, as the alpha increases with increasing task difficulty. On a single participant level, we find this at an individual frequency of around 10–12 Hz in 6 of 7 participants (subj 1 excluded due to recording error, subj 10 excluded due to obvious task misunderstanding: button press after every stimulus). One participant with seemingly different results shows no effect at these frequencies but contrary signed r^2 at lower alpha (around 9 Hz). A possible explanation to this difference is personal stress we observed on a subjective behavioral level. Unexpectedly, theta is not very relevant except for one participant.

For the *bow-to-bow* condition, the differences are much weaker (note the different scale) and results are more variable in general: the spectra look very noisy and variable across participants. In the grand-average, we find no clearly peaked differences in the spectrum. The lower alpha range shows an increase in power for high workload. For single participants, visual alpha peaks show an effect mostly at lower levels around 7–9 Hz in 6 of 9 while around 10 Hz for 2 participants. These visual alpha peaks are mostly at a frequency 2–3 Hz lower than those of the *2-back* condition and mostly with opposite sign (5 participants). Strong frontal theta is found to be significant in one participant (same as in phase 2), less in others.

4 Classification Analysis

4.1 Aim and Approach

The spectral analysis showed strong artifacts in the data. Apart from noise of the technical devices, there are artifacts from muscle activity (seen in high frequencies, mostly at outer temporal and occipital electrodes) as well as from eye movements (seen in low frequencies at very frontal channels). Furthermore, there may be motion artifacts due to the motion of the electrode cables induced by head and trunk movements.

The muscular and ocular artifacts are indicative of the workload condition for a number of participants and could in principle be used for the workload classifier. However, the goal of this analysis was to estimate the contribution of genuine brain activity to the workload level. Still, we evaluated an approach that works on uncorrected data (R), which can be expected to exploit workload-specific artifacts to some degree, and methods including artifact corrected data which work presumably on brain activity only.

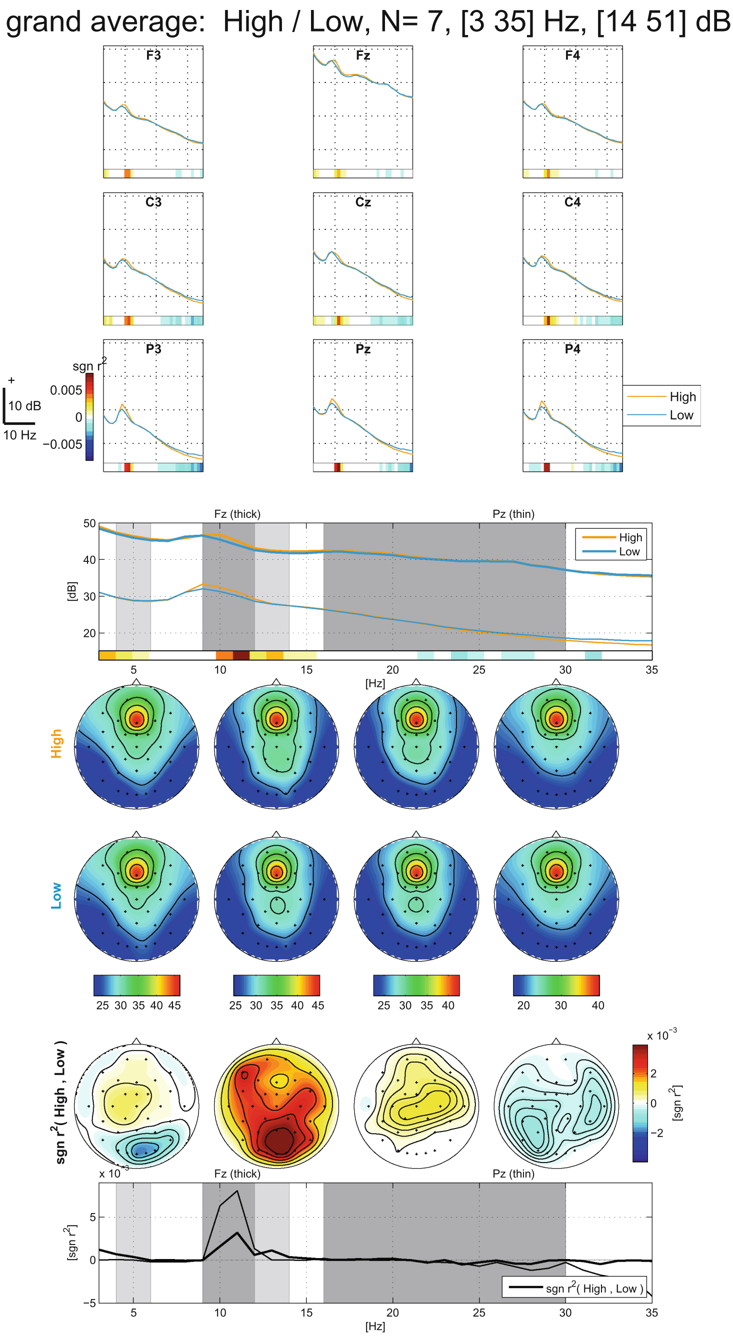

Fig. 5. Grand average of the spectral analysis of the 2-back task.

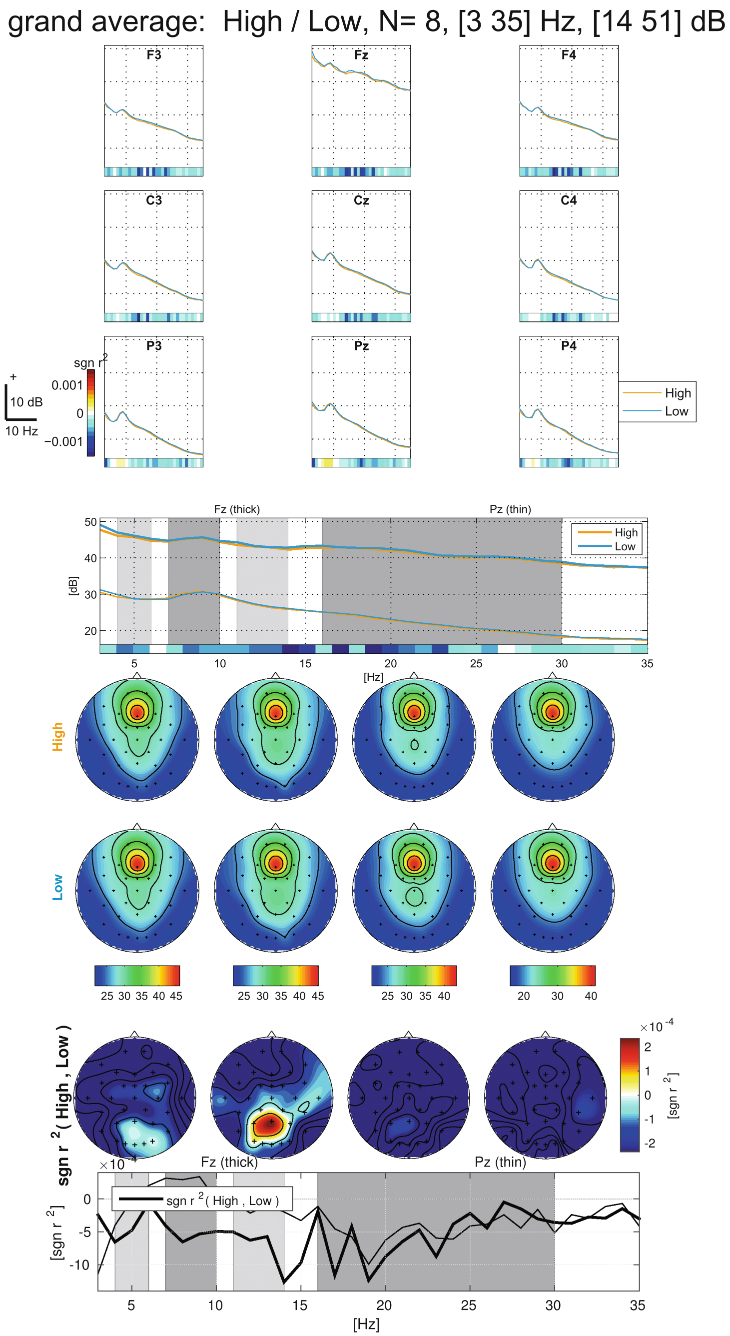

Fig. 6. Grand average of the spectral analysis of the bow-to-bow task.

4.2 Preprocessing, Artifact Reduction, Feature Extraction and Classification

We used 1 Hz high-pass filtering alone (R), in combination with the automatic ICA artifact reduction MARA [22] (C) as well as manual ICA artifact reduction (CM) (for details see Sect. 3.1). The blocks were subdivided into epochs of 1 min. Then, we built different spectral band power based features. In addition, we performed widely used Common Spatial Pattern analysis (CSP) [1] in different band combinations with the logarithm of the variances as features. We evaluated the different classification designs within phases as block-wise cross-validations (CV) as well as between phases to test for generalization. The classifier itself was based on regularized shrinkage linear discriminant analysis (rsLDA) [21].

4.3 Results

The results show a high variability in performance between participants. Classification works best in the *2-back scenario* (phase 2), but also the intra-phase classification in the more complex *bow-to-bow* scenario (phases 1 and 3) works well with CV-loss below 25% for methods **R** and **C**. The transfer of the classifier between the different tasks (*2-back* and *bow-to-bow*) yielded results around

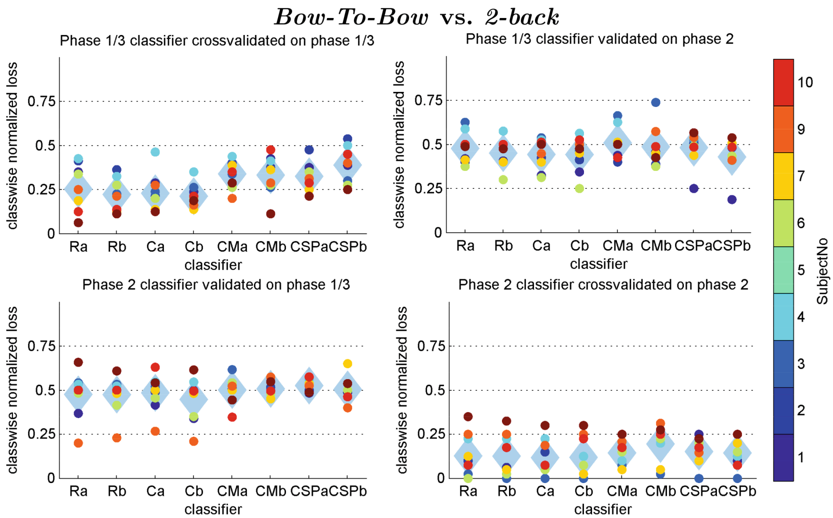

Fig. 7. Classification results: class-wise normalized loss (average indicated by light blue diamond, individual participants colored circles). Methods are labeled by capital letters for preprocessing and lower-case for different combinations of frequency bands: **R** only high-pass (1 Hz), **C** MARA artifact removal, **CM** manual ICA based artifact removal: **a** 1 Hz bins from 1–20 Hz, **b** sum over alpha (8–12 Hz) and theta band (4–7 Hz). CSP common spatial pattern algorithm: **CSPa** alpha and theta band **CSPb** alpha, beta, gamma and theta band. (Color figure online)

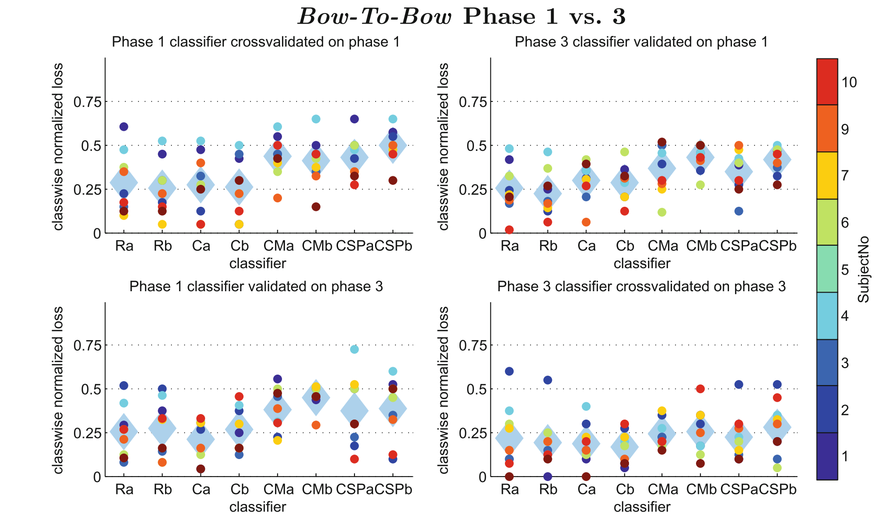

Fig. 8. Classification Results: class-wise normalized loss (average indicated by light blue diamond, individual participants colored circles). Methods are labeled by capital letters for preprocessing and lower-case for different combinations of frequency bands: **R** only high-pass (1 Hz), **C** MARA artifact removal, **CM** manual ICA based artifact removal: **a** 1 Hz bins from 1–20 Hz, **b** sum over alpha (8–12 Hz) and theta band (4–7 Hz). CSP common spatial pattern algorithm: **CSPa** alpha and theta band **CSPb** alpha, beta, gamma and theta band. (Color figure online)

chance level (see Fig. 7), while the transfer of classifiers within the *bow-to-bow* scenario, i.e. between phases 1 and 3, works almost as good as the respective within phase classification (see Fig. 8). On average, the automatically cleaned sums over alpha & theta band (**Cb**) show best results. CSP works comparably in the *2-back* but almost at chance level in the *bow-to-bow* scenario.

5 Discussion

The results of the *2-back* scenario already show the complexity of a physiological index of cognitive workload, when the task is not performed in a constrained laboratory setting, but embedded into a more realistic and complex scenario. The expected increase of the frontal theta oscillation was in general not observed, and the power in the parietal alpha did not decrease as found in most workload studies, but it showed a contrary effect. The alpha effect is consistent with some of the literature on non-visual tasks. The common hypothesis for this ambiguity is that the task-irrelevant visual brain region is actively inhibited in order to focus resources to the relevant non-visual processing. One participant who showed a parieto-occipital alpha decrease might have had a visual strategy to memorize the sequence of numbers - albeit presented auditorily.

Classification of workload levels in the complex *bow-to-bow* scenario is in general successful. Transferring the classifier between the two phases 1 and 3 does not degrade the performance appreciably. The fact that classification in the realistic *bow-to-bow* task worked less well compared to the *2-back* task requires consideration. As found in the literature on electrophysiological correlates of workload, there are two opposing effects concerning the modulation of the alpha rhythm. Tasks in the visual modality mostly decrease alpha activity, while workload in non-visual modalities might increase it (as in our *2-back* paradigm). In the complex *bow-to-bow* scenario, these effects may be in conflict. Retrieving expertise about the maneuvering can be expected to be mainly non-visual. Nevertheless, the control of the boat does not allow a rigorous visual inhibition as it requires synchronized visual processing, in particular as the weather conditions were challenging during the high workload condition. This conflict can be assumed to lead to a much weaker effect on alpha power - an issue well worth deeper investigation.

Another interesting point in the view of applicability is the fact that the complex concept workload encompasses different factors, two of which are *activation* and *resources*. Therefore, a high output of the workload monitor could indicate a strong activation in the sense of an effective focusing of the task at hand (inhibiting task-irrelevant processing). Accordingly, an interpretation of the participant being at the limit of her/his resources might not be adequate.

6 Conclusions

Out-of-the-lab results are not easily transferable to realistic settings, but individual strategy and modality dependent neural activity can be used to adopt the interaction between system and user towards a symbiosis.

Acknowledgments. This research was mainly funded by MARIN, The Netherlands and we want to thank MARIN's Nautical Center, Wageningen for the very kind collaboration. The research leading to these results has also received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 611570 (MindSee), and by the BMBF (contract 01GQ0850, BFNT).

References

1. Blanchard, G., Blankertz, B.: BCI competition 2003-data set IIa: spatial patterns of self-controlled brain rhythm modulations. IEEE Trans. Biomed. Eng. **51**(6), 1062–1066 (2004)
2. Galin, D., Johnstone, J., Herron, J.: Effects of task difficulty on EEG measures of cerebral engagement. Neuropsychologia **16**(4), 461–472 (1978)
3. Gevins, A., Leong, H., Du, R., Smith, M.E., Le, J., DuRousseau, D., Zhang, J., Libove, J.: Towards measurement of brain function in operational environments. Biol. Psychol. **40**(1–2), 169–186 (1995)
4. Gevins, A., Smith, M.E.: Neurophysiological measures of cognitive workload during human-computer interaction. Theor. Issues Ergon. Sci. **4**(1–2), 113–131 (2003)

5. Gevins, A., Smith, M.E., Leong, H., McEvoy, L., Whitfield, S., Du, R., Rush, G.: Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. Hum. Factors **40**(1), 79–91 (1998)
6. Gundel, A., Wilson, G.F.: Topographical changes in the ongoing EEG related to the difficulty of mental tasks. Brain Topogr. **5**(1), 17–25 (1992)
7. Holm, A., Lukander, K., Korpela, J., Sallinen, M., Muller, K.M.: Estimating brain load from the EEG. Sci. World J. **9**, 639–651 (2009)
8. Kirchner, W.K.: Age differences in short-term retention of rapidly changing information. J. Exp. Psychol. **55**, 352–358 (1958)
9. Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., Winkler, T.: Paradoxical'alpha synchronization in a memory task. Cogn. Brain Res. **7**(4), 493–501 (1999)
10. Kohlmorgen, J., Dornhege, G., Braun, M., Blankertz, B., Müller, K.-R., Curio, G., Hagemann, K., Bruns, A., Schrauf, M., Kincses, W.: Improving human performance in a real operating environment through real-time mental workload detection. In: Dornhege, G., Millán, J.R., Hinterberger, T., McFarland, D., Müller, K.-R. (eds.) Toward Brain-Computer Interfacing, pp. 409–422. MIT press, Cambridge (2007)
11. Kok, A.: Event-related-potential (ERP) reflections of mental resources: a review and synthesis. Biol. Psychol. **45**(1), 19–56 (1997)
12. Kramer, A.F.: Physiological metrics of mental workload: a review of recent progress. In: Damos, D.L. (ed.) Multiple-Task Performance, pp. 279–328. Taylor & Francis, London (1991)
13. Legewie, H., Simonova, O., Creutzfeldt, O.: EEG changes during performance of various tasks under open-and closed-eyed conditions. Electroencephalogr. Clin. Neurophysiol. **27**(5), 470–479 (1969)
14. Manzey, D.: Determinanten der Aufgabeninterferenz bei Doppeltätigkeiten und ressourcentheoretische Modellvorstellungen in der Kognitiven Psychologie. Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt (DFVLR), Köln (1988)
15. Manzey, D.: Psychophysiologie mentaler Beanspruchung. In: Rösler, F., (ed.) Ergebnisse und Anwendungen der Psychophysiologie. Enzyklopädie der Psychologie, Göttingen. Hogrefe (1997)
16. Markand, O.N.: Alpha rhythms. J. Clin. Neurophysiol. **7**(2), 163–190 (1990)
17. Pfurtscheller, G., Neuper, C.: Motor imagery and direct brain-computer communication. Proc. IEEE **89**(7), 1123–1134 (2001)
18. Polich, J.: Task difficulty, probability, and inter-stimulus interval as determinants of P300 from auditory stimuli. Electroencephalogr. Clin. Neurophysiol. Evoked Potentials Sect. **68**(4), 311–320 (1987)
19. Pope, A.T., Bogart, E.H., Bartolome, D.S.: Biocybernetic system evaluates indices of operator engagement in automated task. Biol. Psychol. **40**(1), 187–195 (1995)
20. Smith, M.E., Gevins, A., Brown, H., Karnik, A., Du, R., Gevins, A.S.: Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. Hum. Factors **43**, 366–380 (2001)
21. Vidaurre, C., Krämer, N., Blankertz, B., Schlögl, A.: Time domain parameters as a feature for EEG-based brain-computer interfaces. Neural Netw. **22**(9), 1313–1319 (2009). Brain-Machine Interface
22. Winkler, I., Haufe, S., Tangermann, M.: Automatic classification of artifactual ICA-components for artifact removal in EEG signals. Behav. Brain Funct. **7**(1), 1–15 (2011)