Xixie Zhang, Laurens R. Krol, Thorsten O. Zander

Towards task-independent workload classification: Shifting from binary to continuous classification

WISSEN IM ZENTRUM
UNIVERSITÄTSBIBLIOTHEK

Technische
Universität
Berlin

Towards Task-Independent Workload Classification: Shifting from Binary to Continuous Classification

Xixie Zhang
Biological Psychology and Neuroergonomics
Technical University of Berlin
Berlin, Germany
xixie.zhang@gmail.com

Laurens R. Krol
Biological Psychology and Neuroergonomics
Technical University of Berlin
Berlin, Germany
lrkrol@gmail.com

Thorsten O. Zander
Zander Laboratories B.V.
Amsterdam, Netherlands
thorsten@zanderlaboratories.com

Abstract—**Passive Brain-Computer-Interfaces provide a promising approach to the continuous measurement of mental workload in realistic scenarios. Typically, a BCI is calibrated to discriminate between different levels of workload induced by a specific task. However, workload in realistic scenarios is typically a result of a mixture of different tasks. Here, we present a study on investigating the possibility of a task-independent classifier, which can be applied to classify mental workload induced by various tasks (including n-back, backward span, addition, word recovery and mental rotation). Furthermore, our approach is not limited to binary classification of workload but can discriminate it on a continuous metric.**

Keywords—passive Brain-Computer-Interfaces, neuroadaptive technology, mental workload, task-independent classifier, continuous metric

I. INTRODUCTION

Passive brain-computer interfaces (pBCI) provide a tool for Human-Computer Interaction that allows the assessment of cognitive and affective states in real-time [1]. This allows a technical system to adapt itself to changes in user state in order to support the ongoing interaction, leading to neuroadaptive technology [2][3]. For example, during driver-vehicle-interaction, passive BCI could enable the real-time detection of the driver's mental state like fatigue, workload, and degree of relaxation [4], and could provide this essential information regarding driver's state to the car. Combining with other sensor data, the car could adapt to individual aspects of the driver in their context and change its behavior accordingly.

Mental workload is an important part of a users' state, specifically in safety-critical environments. It has been identified as having a great influence on performance in many tasks, such as in learning, driving, or monitoring (e.g. [5], [6]). Currently, the user's actual mental workload is difficult to assess during human-computer interaction, making mental workload a *covert aspect of user state* [7]. Regardless of this difficulty, a reliable estimation of the actual mental workload during the execution of the task is highly valuable [8], [9]. Currently, the most used method for assessing a subjects' mental workload are subjective measures, e.g. questionnaires. Such methods are typically intrusive and prone to social expectance and are hence unsuitable for continuous workload assessment. Alternatively, psychophysiological measures can be used to measure correlates of user state nonintrusively and continuously. Human brain activity reflects an increase of mental workload by an increase in frontal theta power and a decrease in parietal alpha power [10]. These changes can be detected reliably by a passive Brain-Computer Interface during an ongoing Human-Computer Interaction [11].

Current approaches to workload detection are typically limited to the investigation of a specific task. As long as it remains unclear how well an approach can be transferred between tasks, a classifier for each task needs to be calibrated which is a cumbersome and time-consuming procedure. Therefore, we are motivated to investigate whether we could calibrate a classifier that is transferable between different tasks. However, if we investigate different tasks it is unclear how we can compare the different difficulty levels of these tasks and how these affect the actual mental workload: an "easy" exercise in task A may still be more difficult than a "hard" one from task B. From this perspective, it becomes clear that binary classification as typically used in a BCI setting, does not make sense here. In the calibration phase, the decision threshold is chosen specifically for the calibration task, discriminating optimally between low and high workload as defined for that task. In another task the optimal threshold is likely to be very different which makes direct transfer of the classifier unfeasible. This is specifically a problem for realistic settings where we cannot foresee which kind of workload will be induced in the human user, as they might face a summation of different kinds of tasks, partially even simultaneously. Hence, it is necessary to omit binary classification when setting up a task-independent classifier.

Here we present a study investigating the task-independency of a classifier to indicate actual workload level using EEG, which is based on the preliminary study in [12]. We extend the application of this task-independent classifier to tasks that are commonly used in workload research and to application-oriented tasks. We show the unreliability of binary classification and instead define a continuous workload level by interpreting the classifier output directly.

A. Experiment Setup

This experiment was conducted in a well-controlled laboratory and 15 participants with a mean age of 28.6 ± 4.1 participated in this experiment. The participants were instructed to carry out a few tasks on a 25-inch screen standing on a table in front of them. All participants performed a calibration task, a no-workload baseline session, and five testing tasks of two different workload levels each.

- Calibration task: This task consisted of 40 trials in random order (20 with high, 20 with no workload). During high workload conditions, participants were presented with an equation of the form $a - b$, instructing them to count backwards from a in steps of b (i.e. a modified Brown-Peterson distraction technique). a was any integer between 200 and 1200; b ranged from 6 to 19, excluding 10 and 15. During no-workload phases, a crosshair was shown in the middle of screen and participants were instructed to relax, with eyes open. In 50% random trials of both high and no workload conditions, there were 10 small visual lightly-colored "sparkles" wandering smoothly over the screen in random walks governed by perlin noise, providing visual distraction. This 50% chance of sparkles was presented so as to evenly balance eye movements between classes. A self-paced break was implemented after every 10 trials. Each trial lasted 10 seconds, for a total of 200 seconds of EEG data per workload class.

- In the testing phase there were three conditions of different demand levels. One is the no-workload condition. In this session, participants were instructed to relax themselves and press a button when a certain letter was shown in the middle of the screen. This no-workload session was to serve as a reference. As other testing tasks also involved button presses, we included these in the baseline session as well. Participants were informed that reaction time of button press would not be recorded in order to minimize any pressure they might experience. Following this, the testing phase continued with different low- and high-workload conditions. Five tasks were performed: n-back, backward digit span, addition, scrambled word recovery and mental rotation, each consisting of two difficulty levels to induce different workload. N-back task [13] and digit span task [14] are commonly used for inducing workload and represent two different theories of mental workload [15]. The addition, word recovery, and mental rotation tasks are more practical tasks representing different aspects of cognitive abilities, i.e. mathematical, linguistic and spatial ability, and could be used to investigate the applicability of workload detection in more real-life tasks (e.g. [16], [17], [18]).

- N-back task: Participants were given a sequence of numbers and had to indicate when the currently displayed number was equal to the number displayed n elements previously. Participants performed 24 trials of 25 seconds each, among them 12 trials with high

difficulty, 12 with low difficulty. We used 1-back (low difficulty) and 3-back (high difficulty) varieties of the n-back task to manipulate the task difficulty and the task load. The sequence of high/low workload blocks is randomized. The same applies to the following tasks.

- Backward span task: Participants were given a sequence of numbers which they then had to repeat in reverse order. This task consisted of 2-digit span task (low difficulty) and 6-digit span task (high difficulty) conditions, and had 20 blocks each lasting 30s: randomly 10 with high difficulty and 10 with low difficulty (the same for the following three tasks).

- Addition task: Participants were given two numbers to add. The workload level was varied according to the Q-value of the equation ([16], [19]): for low difficulty 2-digit additions with Q-value from 2 to 2.5, for high difficulty 4-digit additions with Q-value from 4 to 5.

- Word recovery task: Participants were presented scrambled German words and they had to recover the original words. The workload level was differentiated by the use frequency and syllable number of the words [20]. The scrambled words in the low-difficulty condition are frequently-used 2-syllabled words, and those in high-difficulty condition are infrequent 3-syllabled words.

- Mental rotation task: Participants were presented either two figures of 2D objects containing 6 squares (low difficulty) or two figures of 3D objects containing 9 cubes (high difficulty) and were instructed to determine whether they were the same object (dataset from [18]).

The two difficulty levels in each of these five tasks were validated in a pretest with 9 participants reporting mental workload levels using the NASA-TLX and a one-itemed scale on task difficulty. Their subjective ratings were significantly different between the workload levels of each task. In this experiment, after the first and last blocks of each workload level, participants reported their subjective mental workload rating using the one-itemed scale.

B. Data Acquisition

EEG data were recorded continuously using a 64 active Ag/AgCl electrode mounted according to the extended 10–20 system on an elastic cap (ActiCap by Brain Products GmbH, Gilching, Germany). The signal was sampled at 5000 Hz and amplified using BrainAmp DC amplifiers (Brain Products GmbH, Gilching, Germany). All electrodes were referenced to FCz and the ground electrode was placed at position AFz.

C. Data Analysis

To investigate the task-independence of the workload classifier, individual classifiers were first calibrated based on data from the calibration phase. The data were divided in consecutive 1-second epochs of high versus low condition data. Filter bank common spatial patterns [21] was applied to extract features describing the power in theta (4-7 Hz) and alpha (8-13 Hz) bands using three patterns per band. Linear discriminant analysis (LDA) was used to separate the classes

with a 5-fold nested cross-validation with margins of 5. The calibrated classifiers were then applied to 1-second epochs taken from the other five tasks. Classification was made between two different workload levels (high vs. low workload) as well as between high vs. no and low vs. no workload conditions. Furthermore, the classifier output of different workload conditions in all tasks was calculated. Differences were compared with permutation tests.

Calibration and classification was done using the open-source MATLAB-based toolbox BCILAB (version 1.2) [22].

III. RESULTS

Table I lists the classification accuracies of a classifier trained on data from the calibration task, applied to the data from the five testing tasks. The mean estimated classification accuracy over participants in the calibration task reaches 76% ±11 (significance is reached at 55% or above). Applied to the five testing tasks, however, no meaningful performance can be reported.

Table II lists the classification accuracies of applying the same calibrated classifier on testing data between high and no workload conditions, as well as between low and no workload conditions. In some tasks, we see average classification accuracies above 75%. However, there is a large variability in the results and some tasks, notably n-back and backward span tasks, appear to consistently elude reliable classification.

Table III lists the mean classifier output values produced when applying the calibrated classifier to the data of the five testing tasks as well as the no-workload baseline recording. Values could vary between 0 and 1, and represent the classifier's predictions, with 0.5 representing complete uncertainty between the classes, 0 representing certainty for class 1 (i.e. the classifier predicts a 100% chance of the data belonging to class 1, low workload) and 1 certainty for class 2 (high workload).

TABLE I. CLASSIFICATION ACCURACIES OF HIGH VS. LOW WORKLOAD CONDITIONS

Participant	Calibration	Classification Accuracy High vs. Low Workload Levels (%)				
		N-back	Span	Add	Word	Rotat.
1	64	50	51	56	56	54
2	84	65	60	62	60	53
3	67	60	60	62	61	54
4	60	54	64	55	51	54
5	78	54	49	55	59	52
6	71	55	56	54	58	56
7	92	50	62	64	65	49
8	78	58	65	61	55	49
9	92	59	67	51	53	57
10	82	54	60	57	50	54
11	88	60	68	64	51	50
12	82	59	51	74	62	58
13	80	53	51	60	55	53
14	62	49	58	54	47	49
15	61	57	55	56	56	57
Mean	**76**	**56**	**59**	**59**	**56**	**53**
St.dev.	11	4	6	6	5	3

TABLE II. CLASSIFICATION ACCURACIES OF HIGH VS. NO WORKLOAD CONDITIONS AS WELL AS LOW VS. NO WORKLOAD CONDITIONS

Participant	Calibration	Classification Accuracy High vs. No Workload (%)					Classification Accuracy Low vs. No Workload (%)				
		N-back	Span	Add	Word	Rotat.	N-back	Span	Add	Word	Rotat.
1	64	48	62	86	81	83	47	61	76	73	76
2	84	69	72	83	78	72	51	58	67	65	67
3	67	70	67	89	87	89	60	55	75	74	83
4	60	63	70	64	69	56	58	53	58	70	51
5	78	54	58	88	80	87	50	59	83	69	87
6	71	60	59	65	69	73	54	52	60	61	64
7	92	64	66	75	59	74	64	53	58	41	76
8	78	60	74	86	86	86	52	57	73	83	87
9	92	67	80	88	84	90	56	67	88	82	82
10	82	46	60	72	80	73	42	48	63	82	69
11	88	62	73	92	96	98	50	53	74	96	98
12	82	59	57	92	87	86	49	58	64	76	76
13	80	46	56	85	68	76	43	56	75	62	74
14	62	46	64	57	53	62	47	52	52	57	64
15	61	59	58	69	55	63	51	53	61	47	54
Mean	**76**	**58**	**65**	**79**	**75**	**78**	**52**	**56**	**68**	**69**	**74**
St.dev.	11	8	7	11	13	12	6	4	10	14	13

Figure 1 presents the mean classifier output across all participants. We see that the mean output for the n-back task in both high and low workload conditions are below the middle line (0.5), and those of the addition task, word recovery task, and mental rotation task in are similarly all above 0.5. Table IV lists the results of permutation tests comparing the output within each participant between high and low workload conditions. For each task, over two-thirds of the participants had significantly different classifier output between conditions. Over all participants, in 76% of all tasks the differences between high and low conditions were significant.
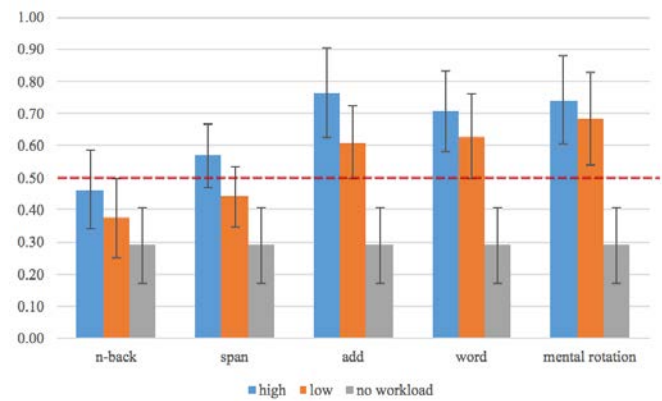

Fig. 1. Averaged classifier outputs across participants

TABLE III. MEAN CLASSIFIER OUTPUTS OF HIGH VS. LOW WORKLOAD CONDITIONS

Participant	Mean Continuous Classifier Output										
	No Workload	N-back		Span		Add		Word		Mental Rotation	
		High	Low	High	Low	High	Low	High	Low	High	Low
1	0.26	0.31	0.32	0.54	0.52	0.82	0.69	0.76	0.65	0.8	0.69
2	0.43	0.71	0.5	0.72	0.56	0.87	0.68	0.77	0.67	0.73	0.67
3	0.23	0.49	0.46	0.62	0.41	0.57	0.49	0.64	0.63	0.45	0.38
4	0.28	0.56	0.41	0.52	0.39	0.95	0.67	0.89	0.69	0.95	0.86
5	0.35	0.41	0.39	0.48	0.52	0.93	0.79	0.75	0.62	0.92	0.92
6	0.29	0.47	0.4	0.49	0.36	0.54	0.48	0.6	0.49	0.63	0.54
7	0.54	0.66	0.69	0.71	0.57	0.77	0.58	0.6	0.46	0.75	0.79
8	0.18	0.44	0.32	0.64	0.43	0.83	0.63	0.83	0.76	0.84	0.84
9	0.23	0.54	0.42	0.71	0.53	0.85	0.79	0.77	0.71	0.84	0.72
10	0.26	0.28	0.22	0.49	0.3	0.64	0.54	0.73	0.76	0.65	0.61
11	0.04	0.35	0.19	0.51	0.23	0.81	0.58	0.87	0.88	0.8	0.79
12	0.2	0.44	0.25	0.43	0.41	0.86	0.48	0.78	0.64	0.79	0.65
13	0.36	0.37	0.29	0.51	0.5	0.86	0.73	0.64	0.57	0.76	0.7
14	0.41	0.37	0.38	0.67	0.47	0.56	0.47	0.49	0.55	0.62	0.64
15	0.34	0.53	0.4	0.49	0.43	0.63	0.54	0.48	0.36	0.56	0.49
Mean	**0.29**	**0.46**	**0.38**	**0.57**	**0.44**	**0.77**	**0.61**	**0.71**	**0.63**	**0.74**	**0.68**
St.dev.	0.12	0.12	0.12	0.10	0.10	0.14	0.11	0.13	0.13	0.14	0.15

TABLE IV. RESULTS OF PERMUTATION TESTS ON OUTPUTS OF HIGH VS. LOW WORKLOAD CONDITIONS OF EACH PARTICIPANT

Participants	P-value of Permutation Test					Percentage of Significant Tasks (%)
	N-back	Span	Addition	Word	Rotation	
1	$p = 0.801$	$p = 0.491$	$p < 0.001$***[a]	$p < 0.001$***	$p < 0.001$***	60
2	$p < 0.001$***	$p < 0.001$***	$p < 0.001$***	$p < 0.001$***	$p = 0.001$***	100
3	$p = 0.002$**	$p < 0.001$***	$p < 0.001$***	$p < 0.001$***	$p < 0.001$***	100
4	$p = 0.214$	$p < 0.001$***	$p = 0.002$**	$p = 0.649$	$p = 0.004$**	60
5	$p = 0.389$	$p = 0.191$	$p < 0.001$***	$p < 0.001$***	$p = 0.846$	40
6	$p = 0.009$**	$p < 0.001$***	$p = 0.029$*	$p < 0.001$***	$p < 0.001$***	100
7	$p = 0.239$	$p < 0.001$***	$p < 0.001$***	$p < 0.001$***	$p = 0.014$*	80
8	$p < 0.001$***	$p < 0.001$***	$p < 0.001$***	$p = 0.002$**	$p = 0.762$	80
9	$p < 0.001$***	$p < 0.001$***	$p = 0.002$**	$p = 0.016$*	$p < 0.001$***	100
10	$p = 0.012$*	$p < 0.001$***	$p < 0.001$***	$p = 0.313$	$p = 0.102$	60
11	$p < 0.001$***	$p < 0.001$***	$p < 0.001$***	$p = 0.433$	$p = 0.210$	60
12	$p < 0.001$***	$p = 0.347$	$p < 0.001$***	$p < 0.001$***	$p < 0.001$***	80
13	$p = 0.002$**	$p = 0.726$	$p < 0.001$***	$p < 0.001$***	$p = 0.004$**	80
14	$p = 0.634$	$p < 0.001$***	$p = 0.015$*	$p = 0.075$	$p = 0.604$	40
15	$p < 0.001$***	$p = 0.023$*	$p = 0.001$**	$p < 0.001$***	$p = 0.004$**	100
Percent. of particip. w. significance (%)	66.7	73.3	100	73.3	66.7	Average: 76

[a.] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

IV. DISCUSSION

The main aim of this study was to investigate the feasibility of a task-independent classifier for mental workload detection. Such a classifier could potentially greatly simplify the calibration session in more practical environments. To manipulate workload levels, we exposed participants to one no-workload session and five different tasks inducing mental workload, each containing two different workload levels.

From Table I and Table II it can be concluded that a direct transfer between tasks is not sufficient to separate high and low workload levels within new tasks with acceptable accuracy. This, however, was to be expected: different tasks are subjectively more or less demanding, and a direct comparison between "high" and "low" load levels across tasks is thus unrealistic. Instead, we turned to classifier output on a continuous scale.

We can see from Figure 1 that the mean classifier output differs for each of the five testing tasks. The mean classifier outputs of both conditions in the n-back task are below 0.5, whereas those in addition, word recovery, and mental rotation tasks are all above 0.5. As the default separating threshold of high versus low workload conditions in our calibrated classifier is 0.5, the unreliable classification results in these tasks can be explained to some extent. Furthermore, the permutation tests on outputs of each subject reveal that for most subjects in most tasks there are significant differences between conditions. This implies that the calibrated classifier did in fact produce reliable differences between the workload levels, but that these differences were simply not centered on the original threshold.

We conclude that, due to inherent differences between different tasks with respect to the induced workload, it is not advisable to directly transfer a binary classifier, calibrated on one task, to another task. Instead, a more reliable measure may be to use the raw classifier output mapped onto a linear scale.

Alternatively, it is possible to adjust the decision threshold for each task individually. However, this would essentially again require a calibration phase for each task specifically, no longer making the procedure task-independent.

We calibrated a classifier on a continuous subtraction task, and applied it to five other tasks that induced two different levels of workload, as well as to a no-workload condition. We found significant differences in the classifier output between conditions. The tasks we used represent a variety of different types of workload. These initial results imply that it is possible to calibrate a single classifier on a standardized workload-inducing task, and use this same classifier to obtain meaningful information concerning workload levels induced by other tasks. This could greatly simplify the implementation of real-time, continuous workload detection into real-world applications.

In future work, we will analyze how the classifier output correlates to the subjective ratings that participants gave after each block. Different participants could have different abilities and different perceptions of tasks, and this may help explain the variance still present in the results. We will also investigate the generalizability of this classifier between subjects, i.e. we will take further steps towards a *universal* workload classifier—one that is both task- and subject-independent.

REFERENCES

[1] T. O. Zander and C. Kothe, "Towards passive brain–computer interfaces: Applying brain–computer interface technology to human–machine systems in general," J. Neural Eng., vol. 8, no. 2, 2011.

[2] T. O. Zander, L. R. Krol, N. P. Birbaumer, and K. Gramann, "Neuroadaptive technology enables implicit cursor control based on medial prefrontal cortex activity," Proc. Natl. Acad. Sci., vol. 113, no. 52, pp. 14898–14903, 2016.

[3] L. R. Krol, L. M. Andreessen, and T. O. Zander, "Passive Brain-Computer Interfaces: A perspective on increased interactivity," in Brain-Computer Interfaces Handbook: Technological and theoretical advances, C. S. Nam, A. Nijholt, and F. Lotte, Eds. Boca Raton, FL, USA: CRC Press, 2018, pp. 69-86.

[4] B. Blankertz, M. Tangermann, C. Vidaurre, S. Fazli, C. Sannelli, … and K. R. Mueller, "The Berlin brain–computer interface: Non-medical uses of BCI technology," Front. Neurosci., vol. 4, no. DEC, pp. 1–17, 2010.

[5] C. Berka, D. J. Levendowski, M. N. Lumicao, A. Yau, G. Davis, ... and P. L. Craven, "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks," Aviat. Space. Environ. Med., vol. 78, no. 5, pp. B231–B244, 2007.

[6] J. Kohlmorgen, G. Dornhege, M. Braun, B. Blankertz, K. R. Müller, ... and W. Kincses, "Improving human performance in a real operating environment through real-time mental workload detection," Towar. Brain-Computer Interfacing, pp. 409–422, 2007.

[7] T. O. Zander and S. Jatzev, "Context-aware brain–computer interfaces: exploring the information space of user, technical system and environment," J. Neural Eng., vol. 9, no. 1, 2011.

[8] G. Borghini, G. Vecchiato, J. Toppi, L. Astolfi, A. Maglione, … and F. Babiloni, "Assessment of mental fatigue during car driving by using high resolution EEG activity and neurophysiologic indices," Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. Annu. Conf., vol. 2012, no. cm, pp. 6442–5, 2012.

[9] C. D. Wickens, "Multiple Resources and Mental Workload," Hum. Factors J. Hum. Factors Ergon. Soc., vol. 50, no. 3, pp. 449–455, 2008.

[10] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis," Brain Res. Rev., vol. 29, no. 2–3, pp. 169–195, 1999.

[11] P. Gerjets, C. Walter, W. Rosenstiel, M. Bogdan, and T. O. Zander, "Cognitive state monitoring and the design of adaptive instruction in digital environments: Lessons learned from cognitive workload assessment using a passive brain-computer interface approach," Front. Neurosci., vol. 8, no. DEC, pp. 1–21, 2014.

[12] L. R. Krol, S. C. Freytag, M. Fleck, K. Gramann, and T. O. Zander, "A task-independent workload classifier for neuroadaptive technology: Preliminary data," 2016 IEEE Int. Conf. Syst. Man, Cybern. SMC 2016 - Conf. Proc., pp. 3171–3174, 2017.

[13] W. K. Kirchner, "Age differences in short-term retention of rapidly changing information," J. Exp. Psychol., vol. 55, no. 4, 1958.

[14] M. L. Turner and R. W. Engle, "Is working memory capacity task dependent?" J. Mem. Lang., vol. 28, no. 2, pp. 127–154, 1989.

[15] T. S. Redick and D. R. B. Lindsey, "Complex span and n-back measures of working memory: A meta-analysis," Psychon. Bull. Rev., vol. 20, no. 6, pp. 1102–1113, 2013.

[16] C. Walter, W. Rosenstiel, M. Bogdan, P. Gerjets, and M. Spüler, "Online EEG-Based Workload Adaptation of an Arithmetic Learning Environment," Front. Hum. Neurosci., vol. 11, no. May, pp. 1–10, 2017.

[17] H. Koshino, P. A. Carpenter, T. A. Keller, and M. A. Just, "Interaction between the dorsal and ventral pathways in mental rotation : An fMRI study," vol. 5, no. January, pp. 54–66, 2005.

[18] W. K. Y. So, S. W. H. Wong, J. N. Mak, and R. H. M. Chan, "An evaluation of mental workload with frontal EEG," PLoS One, vol. 12, no. 4, p. e0174949, 2017.

[19] H. B. G. Thomas, "Communication theory and the constellation hypothesis of calculation," Q. J. Exp. Psychol., vol. 15, pp. 173–191, 1963.

[20] M. Brysbaert, M. Buchmeier, M. Conrad, A. M. Jacobs, J. Bölte, and A. Böhl, "The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German," Exp. Psychol., vol. 58, no. 5, pp. 412–424, 2011.

[21] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface," 2008 IEEE Int. Jt. Conf. Neural Networks (IEEE World Congr. Comput. Intell., no. January 2014, pp. 2390–2397, 2008.

[22] C. Kothe and S. Makeig. "BCILAB: a platform for brain- computer interface development." J. Neural Eng., vol. 10, no. 5, 2013.