Final Project Report

Grant AOARD-10-4163

Neuro-ergonomic Research for Online Assessment of Cognitive Workload

October 2011

Dr. Fang Chen, National ICT Australia (NICTA)
fang.chen@nicta.com.au
Level 5, 13 Garden Street, Eveleigh, NSW 2015, Australia
Tel: +61 2 9376 2101, Fax: +61 2 9376 2027

Period of Performance: 24th August, 2010 ~ 23rd August, 2011

Executive Summary

This report summarizes the research activities, results of the user studies and research accomplishments out of the online assessment of cognitive workload project in the past year. At the start of the project, we carried out a literature review on electroencephalography (EEG) based cognitive load measurement (CLM) to understand the technology landscape. We then investigated some statistical features of EEG signals for online measurement. After that, we explored further on various spectral features, including entropy, weighted mean frequency and bandwidth. Finally we investigated the feasibility of using a low-cost EEG headset for CLM. All together, we had carried out 3 sets of user experiments to validate the research outcomes based on a common set of *n*-letter word identification tasks while a subject is reading silently.

In terms of concrete research outcomes, the following report and papers were published:

- Zarjam, P., *Cognitive Load Measurement Based on EEG Signal Processing*, NICTA Technical Report, October 2010.

- Zarjam, P., Epps, J. and Chen F., "Evaluation of Working Memory Load using EEG Signals", *Proc. APSIPA Annual Summit and Conference (APSIPA'10)*, Biopolis, Singapore, December 2010, pp. 715-719.

- Zarjam, P., Epps, J. and Chen, F., "Spectral EEG Features for Evaluating Cognitive Load", *Proc. 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'11)*, Boston, USA, August 2011, pp. 3841-3844.

- Knoll, A., Wang, Y., Chen, F., Xu, J., Ruiz, N., Epps, J. and Zarjam, P., "Measuring Cognitive Workload with Low-Cost Electroencephalograph," *Proc. IFIP International Conference on Human-Computer Interaction (INTERACT'11)*, Lisbon, Portugal, September 2011, pp. 568-571.

1. REPORT DATE **26 OCT 2011**	2. REPORT TYPE **Final**	3. DATES COVERED **31-08-2010 to 30-08-2011**
4. TITLE AND SUBTITLE **Neuroergonomic research for online assessment of cognitive workload**		5a. CONTRACT NUMBER **FA23861014163**
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) **Fang Chen**		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **National ICT Australia (NICTA),Level 5, 13 Garden Street,Eveleigh, NSW Sydney 2015,Australia,NA,NA**		8. PERFORMING ORGANIZATION REPORT NUMBER **N/A**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) **AOARD, UNIT 45002, APO, AP, 96338-5002**		10. SPONSOR/MONITOR'S ACRONYM(S) **AOARD**
		11. SPONSOR/MONITOR'S REPORT NUMBER(S) **AOARD-104163**

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This is the final report of a project to use electroencephalography based cognitive load measurement to understand its place in the technology landscape.

15. SUBJECT TERMS
Brain Science and Engineering, Computer and User Interface, Congnitive Modeling, Cognitive Modeling

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**91**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

1   Introduction

It is well known that high cognitive workload often leads to human errors and omissions, and decreased performance. Cognitive load (CL) refers to the amount of mental demand imposed by a particular task on a person, and has been associated with the limited capacity of working memory. It is crucial to maintain the cognitive load within an optimal range, not too high or too low, to achieve the highest productivity. Therefore, measuring individual cognitive load in real-time, unobtrusively and automatically is very important for developing better interfaces, delivering better information flow and appropriate content, and even dynamically adjusting system behaviors to maximize individual performance. Automatic and online cognitive load measurement (CLM) can play a crucial role in various application areas involving human-computer interface, such as air traffic control, in-car safety and electronic games.

This project has investigated some important research issues related to robust feature extraction of electroencephalography (EEG) signals for online assessment of cognitive load. In this report, we will first summarize the findings of a literature review on EEG based cognitive load measurement. Followed by that will be a presentation of the statistical features and spectral features that we have experimented with. Finally an experiment on using a low-cost EEG headset for cognitive load measurement will be presented.

2   Literature Review - EEG based CLM

Since the brain is the source of cognitive activity, with appropriate measurement tools such as electroencephalography, it is expected to be able to monitor and assess the brain states more directly. Recently, electroencephalography has been found to be an effective non-intrusive technique for estimating cognitive states of a person in research areas such as educational psychology, neuroscience, and cognitive psychology. However, the automatic measurement of cognitive load using EEG signals is still at its early stage and there is a clear need to validate the applicability of the approach across different tasks and persons, and also to establish the precision of the measurement. Furthermore, most of the previous studies use traditional analysis of EEG signals such as power spectral density. There is a need to explore other more effective signal features for EEG based CLM.

To understand the research landscape, a literature review on EEG based cognitive load measurement has thus been carried out and it can be found in Appendix A. This review covers an introduction to cognitive load measurement in general, clinical and research use of EEG, key characteristics of EEG signals, recent literature on EEG based CLM, and the various time-frequency and time-scale signal processing methods for EEG based CLM.

3   Statistical Features

This study investigates online cognitive workload assessment using statistical features derived from EEG signals. The features investigated include mean, root mean squared, and correlation-based features. Results reveal that for the given task, these features derived from the EEG signals consistently exhibit a very high degree of discrimination between the induced load levels, confirming EEG as an important method for the real time, objective determination of cognitive load level. More details about this study can be found in Appendix B.

3.1 Experimental design and procedure

Five healthy male volunteers, 24-30 years of age, engaged in post-graduate study participated in the experiment. The experiment was defined as a silent reading task, displayed and controlled on a lap-top PC with a viewing distance of 70 cm to the participant. The reading task was chosen to be semantically neutral and comprehension-independent to avoid any expertise effect and the reading levels of all the participants were expected to be relatively similar.

The task was split into three levels; each level consisted of four pages displayed in a Power Point presentation. Each page lasted 30 seconds automatically, thus each reading task lasted 2 minutes in each level. In the first level, defined as low difficulty, the participants were asked to read the displayed pages silently and pick up three letter words by pressing a mouse left button. In the second level, defined as medium difficulty, they were required to read on and pick up three and four letter words by pressing the left and middle buttons, respectively. In the third level, defined as high difficulty, they picked up three, four, and five letter words from the remainder of the presentation, by pressing the left, middle, and right buttons, accordingly. To minimize any muscle movement artifact, the participant's hand and mouse were kept fixed in a certain position. Each load level task was conducted over two minutes and the experiment was repeated two times per subject. In the baseline condition, conducted after the experiment, the subjects were asked to sit relaxed and keep their eyes open.

The EEG signals were recorded by means of a BioSemi Active Two system, at the ATP Laboratory of NICTA in Sydney. Each recording contained 32 EEG channels, according to the International 10-20 system. The data were recorded in digital form, at a 256 Hz sampling rate and a sensitivity of 100 µV. The EEG signals were passed through a low-pass filter with a cut-off frequency of 100 Hz.

3.2 Analysis and results

After examining different statistics of the EEG segments, the following features were found to be more suitable for discriminating between the different levels of loads: mean, root mean squared, and maximum cross-correlation between the given segment and a segment of the same size of the baseline EEG, denoted by F_1, F_2, and F_3, respectively.

All those three features were extracted from the EEG segments of 5 seconds length. The performance of each feature for all the 32 EEG channels was examined and it was found that those two frontal EEG channels; namely: Fp1, and Fp2, were the most sensitive to the load on working memory across the participants. This result is supported by previous studies indicating that the brain frontal lobes often play a critical role in working memory related tasks associated with attention and mental efforts. For illustration purpose, the mean value (F_1) extracted from the EEG signals representing 3 different task difficulties of all the five participants in one trial is shown in Figure 1.
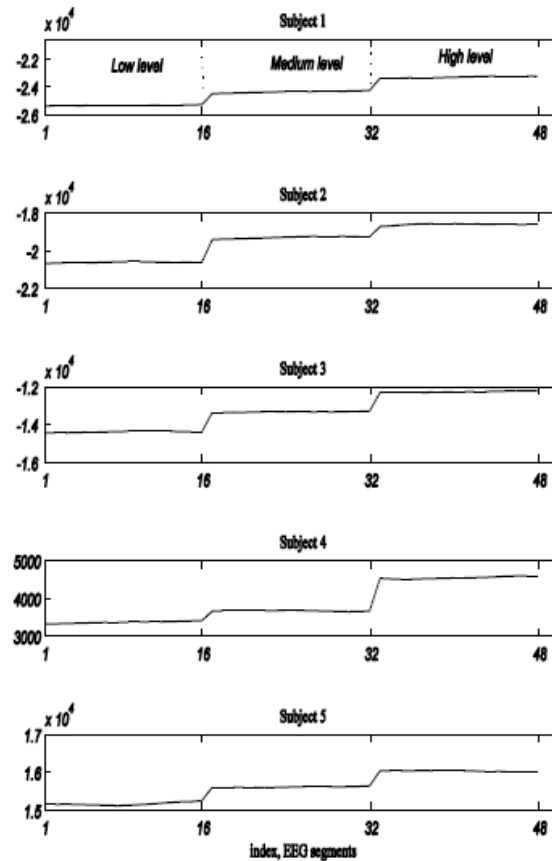
Fig. 1. Values for the mean feature (F_1) extracted from the segmented EEG data for five participants. The first, second, and third 16 segments belong to the EEG signals recorded during low, medium, and high levels of task difficulty, respectively.

Figure 1 shows that as the task difficulty increases, the mean of the EEG signal tends to increase, and therefore, each level of difficulty is clearly distinguishable from other levels by different mean.

We also statistically analyzed the extracted features for all five subjects and found the results to be almost the same for all subjects. Here, we present the results of the statistical analysis for subject 1 only. Figure 2 plots the values of the three features extracted from the EEG signals acquired from the frontal channels during different load levels for subject 1, together with their 95% confidence intervals. The feature F_1 shows a consistent increasing trend as cognitive load is increased. The features feature F_2 and F_3, on the other hand, show decreasing trends. The features extracted from the channel Fp1 show larger confidence intervals, indicating that the trends observed here are less reliable.
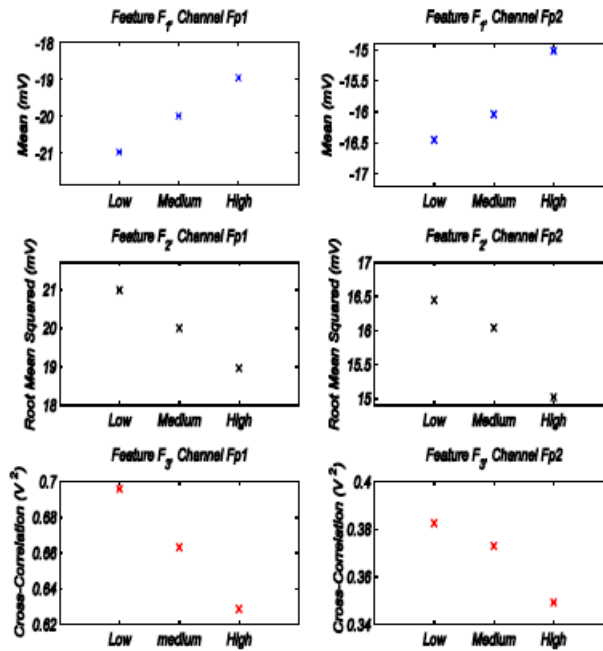
Fig. 2. Extracted features from the EEG signals acquired from the frontal channels, Fp1 and Fp2, during different cognitive load conditions for the first subject.

4   Spectral Features

This study was undertaken to investigate spectral features derived from EEG signals for measuring cognitive load. Based on EEG recordings for a reading task in which three different levels of cognitive load were induced, it is shown that a set of spectral features, including spectral entropy, weighted mean frequency and its bandwidth, and spectral edge frequency, are all able to discriminate the load levels effectively. An interesting result is that spectral entropy, which reflects the distribution of spectral energy rather than its magnitude, provides very good discrimination between cognitive load levels. In addition, the effect of frequency bands on the spectral features is also investigated here. A copy of the published paper is shown in Appendix C.

The dataset as described in Section 3.1 is reused for this study. Therefore the experimental conditions and the participants are identical to those presented before.

4.1   Analysis and results

Five spectral features were extracted from each 5-second segment of the EEG signals. The feature set includes:
- Spectral entropy (SpEn),
- Subband energy (Enrg),
- Intensity weighted mean frequency (IwMf),
- Intensity weighted bandwidth (IwBw), and
- Spectral edge frequency (EdFr).

Since the features of interest are spectral, we initially investigated the spectral components of the recorded EEG signals. It appears that 90% of the energy of the spectral components resides in the 0-3.8 Hz region, computed by the extracted EdFr. Clearly, the delta subband practically provides the most separation between the three load levels induced in the experiment. Therefore, the performance of all the features were examined in the delta sub-

band and compared for all the 32 EEG channels for all participants. We initially calculated the median of all the features for each EEG channel recorded, and then compared the effectiveness of each feature using a Kruskal-Wallis test.

Table 1 lists the EEG channels in which the medians of the features calculated across all participants have shown consistent trends. It displays that both the SpEn and Enrg exhibit a consistent decreasing trend as load level increases, in selected channels. However, the IwMf, IwBw, and EdFr exhibit an increasing trend as load level increases.

Table 1. Variations of the extracted medians of the EEG features in the delta sub-band, indicating the channels that follow a consistent trend associated with the 3 load levels induced, across all participants.

Feature	EEG channels	Trend with increasing load
SpEn	Fp1, AF3, F7, CP5,P7,Pz, P8, CP6,CP2, T8, F4, F8, AF4	decreasing
Enrg	Fp1, AF3, F7, T7, CP3, P7, Pz, P8, CP6, CP2, T8, F4, F8, AF4	decreasing
IwMf	F7, FC5, T7, C3, P7, P3, Pz, P8, CP2, FC6, FC2, Fp2, Fz	increasing
IwBw	Fp1, AF3, F7, T7, C3, CP5, P3, Pz, PO3, P4, CP2, FC6, FC2, F4, Fp2, Fz	increasing
EdFr	Fp1, AF3, F7, P3, Pz, CP2, Fp2, Fz	increasing

In order to examine how effective the extracted features are in differentiating the different load levels from the EEG channels, we used the Kruskal-Wallis test. The largest calculated p-values (indicating the worst case scenario) across the selected channels for each feature for all participants are displayed in Table 2. As the p-values suggest, the Enrg feature shows a great statistical significance in differentiating the cognitive load levels in all the EEG channels. It is followed by the SpEn feature, with the second lowest set of p-values. As seen, most of the selected channels in Table 1 are confirmed in Table 2, statistically.

Table 2. The EEG channels with a p-values<0.01 for all participants in the delta sub-band. The biggest calculated p-values display the worst case scenario.

Feature	EEG channels	Maximum p-value
SpEn	Fp1, AF3, T7, CP5,O1, P8, CP2,F8, Fp2	0.00357
Enrg	All 32 channels	5.0005372e-05
IwMf	P3, Pz, Oz, CP2, Fp2, Fz	0.0076619
IwBw	P3, Pz, CP2, F4, AF4, Fz	0.0091394
EdFr	P3, Pz, CP2, F4, AF4, Fp2, Fz	0.00847211

Following our investigation of the delta frequency subband, we divided this subband further into finer frequency subbands. Therefore, the delta subband (0-4Hz) was split into three subbands of δ_0 (0-1 Hz), δ_1 (1-2 Hz), and δ_2 (2-4 Hz) using wavelet decomposition. The feature medians were recomputed for the δ_0, δ_1, and δ_2 subbands. It was observed that for the SpEn and Enrg, δ_0, and δ_1 sub-bands provided the same results as those in Table 1, showing that the lower sub-bands seem to underlie most of the load level variations. For

the IwMf, IwBw and EdFr, the medians in the finer sub-bands did not provide any significant results, indicating the load level information extracted by these features is distributed all over the delta sub-band. This suggests that the performances of the extracted spectral features are highly dependent on the frequency subdivisions.

To measure the classification accuracy of the extracted features, we used a multi-class support vector machine (SVM) as a classifier. The spectral features were classified with three SVMs in a pair wise strategy. The classification was performed on 5 channels with the smallest p-values. The classification results for all features, on the selected channels which are averaged across all participants are shown in Table 3. As shown, 3 channels, namely, Pz, Cp2 and Fp2, present a high classification accuracy of over 95%.

Table 3. Accuracy of the 3 load level classification for all features by a SVM with linear kernel, averaged over all participants.

EEG channel	P3	Pz	CP2	Fp2	Fz
Classification accuracy %	86.62	95.55	95.55	97.77	82.21

5   Measuring Workload with Low-Cost EEG Headset

Although EEG is a promising tool for continuous measurement of cognitive workload, most previous research has employed high-end EEG systems costing more than $15,000 (e.g. see www.biosemi.com), which limits their widespread usage in human-computer interfaces. On the other hand, low-cost (under $1000) EEG headsets have become accessible for HCI research in recent years. This work takes an initial step in exploring the feasibility of cognitive workload evaluation using a low-cost multi-channel EEG system. The published paper can be found in Appendix D.

5.1   Experimental design and procedure

Sixteen students and NICTA employees (16-46 years old, 4 females) were invited to perform silent reading tasks. Brain waves from each subject were recorded with a low-cost EEG device originally designed for gaming interfaces (Emotiv EPOC, a 14-channel 128Hz neuro-signal acquisition and processing wireless neuro-headset, see Figure 3).
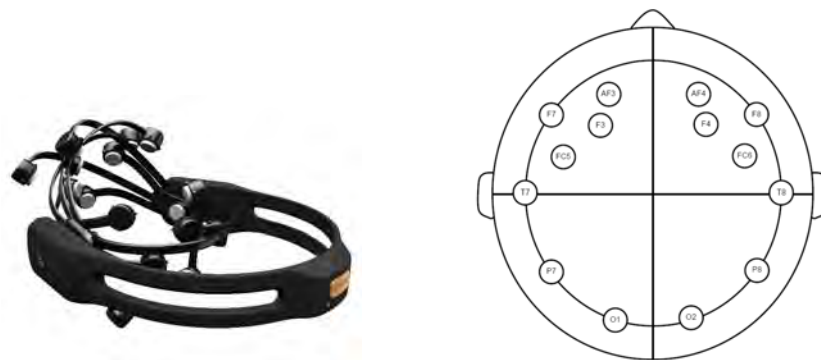

Fig. 3. A low-cost EEG device (Eomtiv EPOC neuroheadset) and layout of EEG channels.

Channel names based on the International 10-20 locations are: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4. During the experiment, each subject was asked to silently read the text displayed on-screen, with a viewing distance of 70cm (see Figure 4). Similarly, different task difficulty levels, as described in Section 3.1, were employed to manipulate cognitive workload during the experiment.
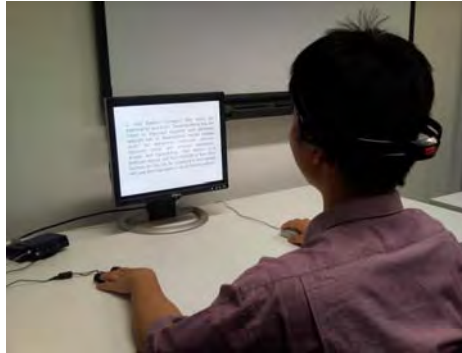

Fig. 4. Experiment setup for using low-cost EEG device.

5.2   *Analysis and results*

The EEG signals were first divided into segments of 1.5 seconds in length. Statistical features including mean, variance, root mean square (RMS), spectral powers of theta (3-7 Hz), alpha (8-12 Hz), beta (13-29 Hz), and gamma (30+ Hz) frequency bands were then calculated for each data segment.

Among the features obtained from different EEG channels, RMS from nodes F3 and F4 exhibit significant correlation with task. This finding is consistent with previous research indicating that the brain frontal lobes play an important role in cognitive tasks associated with attention and mental effort. Moreover, the spectral power of gamma frequency band at each of nodes AF3 and AF4 shows a statistically significant difference between the baseline condition and the task condition, which is consistent with previous study on gamma activation of EEG during cognitive tasks. There is an increase in average gamma power with each rise in task difficulty. However, the difference between task levels is not statistically significant. Details can be found in the published paper.

6   Conclusion

We have explored various statistical features and spectral features for EEG based online assessment of cognitive load, and the feasibility of using low-cost EEG headset to do that. Discrimination between the different workload induced has been found to be highly significant, and the changes in feature values between load levels are consistent for mean features and mostly consistent for root mean squared and correlation features. This discrimination has also been found to be statistically significant using the Enrg, and SpEn as the best performing spectral features, followed by IwMf, EdFr, and IwBw, respectively. Combination of all the features into the SVM classifier has resulted in superior performance compared to one feature taken alone, suggesting that each feature captures different characteristics of the underlying cognitive states. Finally, it is demonstrated that cognitive workload can be effectively measured even with low-cost EEG headset. We hope that this outcome will promote the application of EEG-based physiological measures in various HCI areas involving cognitive workload evaluation.

There is extensive scope for future work, including validating the findings on more participants, inducing a finer scale of task load levels to test the cognitive load measurement precision, using evoked potential methodology to assess cognitive load, and exploring the use of other possible features.

Contents

Cognitive Load Measurement Based on EEG Signal Processing

Pega Zarjam

Signal Processing Research Lab

University of New South Wales

NICTA Lab, ATP

Research Proposal

October 26, 2010

1 Introduction

It is quite well known that the brain has limited cognitive processing in both capacity and duration. These limitations will, under some conditions, hold back people's learning, and performance. Underloading or overloading of cognitive processing will lead to degrading performance or failures of learning and performing complex cognitive tasks. Hence, when applying the task demand on cognitive processing the working memory architecture and its limitations should be an important consideration.

Cognitive load stands for the load on working memory during instructions. Instructions may include teaching learners, problem solving skills, thinking and reasoning skills (i.e., perception, memory, language, etc). Cognitive load theory deals with the instructional methods development that effectively deploy people's limited cognitive processing capacity. In deed, this is designed to optimise intellectual performance and avoid sources of errors.

In recent years, there has been an increased focus on maintaining an optimal level of cognitive load to avoid either underloading or overloading the working memory, especially in critical decision making fields such as air traffic control or military operations, where the load requirements may be heavy and there is no room for error. As a result, there is a significant need for tools to assess and predict cognitive load, and its measurement plays a considerable role in cognitive load theory.

Since, the brain is the source of cognitive activity, with appropriate measurement tools such as Electroencephalography; it is expected to be able to monitor and assess the brain states. Recently, Electroencephalography has been successfully applied in cognitive load processing and proven to be an effective non-intrusive technique for monitoring memory load continuously, and highly sensitive to different cognitive states in research areas such as; educational psychology, neuroscience, and cognitive psychology.

2 Significance and Motivation of this Research

Previously, researchers have used analytical and empirical methods to measure cognitive load [1, 2]. Analytical methods use techniques rely on expert opinion, mathematical models, and task analysis, and are task-specific or conducted post-hoc. Empirical methods gather behavioural [3, 4, 5] or physiological signals [2, 6]. Physiological based techniques are justified by the presumption that changes in cognitive functioning are reflected by physiological variables (e.g., heart rate variability [2]), task-evoked brain activities [6, 7], eye movement [8, 9], and skin conductance [10]. Among the empirical methods, physiological measures are the most suitable ones for practical applications allowing online assessment of cognitive load at all levels. Since, they are able to present the detailed trend and pattern of cognitive load such as instantaneous, peak, accumulated, and overall load [11].

But, some physiological methods have practical limitations in measuring cognitive load. For instance, heart rate variability has been proven to be insensitive to fluctuations of instantaneous cognitive load. Eye movement (blink rate/latency/duration) have shown to have a weak relation to cognitive load. However, the brain activity has proven to be one of the best measures which reflects varying responses to changing levels of cognitive stimuli [12]. Therefore, this research attempts to target cognitive load assessment by means of Electroencephalography (EEG) signals. EEG is a neuroimaging technique that monitors and measures electrical activity produced by the brain, in response to cognitive stimuli presented.

So far, many previous researchers have discovered some of potential features associated with cognitive load and some EEG-based systems for measuring various mental states have been proposed to date. But, this measurement using EEG signals is still in its early stages, and there is a clear need: to validate the applicability of the approach across different tasks and par-

ticipants, to detect different cognitive load levels automatically, to establish the precision of cognitive load measurement. Furthermore, most of the previous studies use traditional analysis of EEG signals such as power spectral density or ERP based methods [6, 13, 14, 15, 16]. Since, EEG signals are non-stationary and change abruptly; the chances are these methods elude the sudden changes in cognitive states associated with the different load levels. Additionally, some previous works rely on costly EEG equipment and techniques which make it difficult for non-EEG-experts to replicate and use their works for further research purposes.

This research aims to evaluate various cognitive load states, and apply advanced signal processing techniques to characterise the associated EEG signals in depth, fulfil the above mentioned needs, and aid the clinician and researchers in their interpretations of cognitive load EEG signals.

3   Cognitive Load Theory

Cognitive load **(CL)** stands for **the load on working memory** during instructions. Instructions may include teaching learners, problem solving skills, thinking and reasoning skills (including perception, memory, language, etc) [17].

Cognitive load theory **(CLT)** investigates **the instructional methods development** that effectively use people's limited cognitive processing capacity. This is because of the fact that researchers in the area of CLT have been concerned with analysing the effects of CL on learning and inventing new techniques and tools to assist learners and users, specially in the field of educational psychology. Indeed, this has been designed to **optimise intellectual performance** [17].

CLT relies on a cognitive architecture consisting of a limited working memory in capacity and time. This consideration is of more importance

when holding or processing new information and a long-term memory with virtually unlimited capacity [18].

One aspect of CLT is based on the fact that only 7 elements of information can be retained in working memory. This number decreases when information has to be not only remembered but also processed. As an example, learning the grammar of a foreign language is more intrinsically complex than learning individual words. Since grammar involves the interaction of several information units (e.g., subject, predicate, and object), while vocabulary can be learned as a sequence of single information elements. In another words, the higher the number of interacting elements in a task, the more difficult it is and the higher the intrinsic load it applies on working memory [19].

Another aspect of CLT is based on the fact that most people learn better when they can build on what they already know. In another word, information that has already been learned, namely, stored in long-term memory, reduces working memory load. Because a schema can be handled in working memory as a single information element. Hence, having prior knowledge or expertise on a task lowers the cognitive load imposed by that task. Moreover, when a task or aspects of a task are repeatedly practiced, cognitive schemata become automated, and no longer require controlled processing, which sets free working memory resources [19].

In recent years there has been an increased focus on the **effectiveness** and efficiency of instructional design strategies in education and training too. According to Pass *et. al*, effectiveness is an essential concept to the design of instruction, which represents a match between results achieved and those desired. In an applied sense, efficiency in instruction can suggest that resources should be used to perform results to maximise learning and minimise the amount of CL effort invested [20].

In [21], CLT is defined as an instructional model fashioned from the field

4

of cognitive science research. It describes learning in terms of an information processing system made up of *long-term memory*, which stores knowledge and skills on a more or less permanent basis, and *working memory*, which performs the intellectual tasks related to learning. Information may only be stored in long-term memory after first being dealt with by working memory. As a matter of fact, working memory is limited in both capacity and duration, and these limitations will, under some circumstances, hold back learning [22]. That is why CLT is designed to optimise learning and avoid underloading or overloading of cognitive processing.

However, the structure of human cognitive architecture, is not fully known yet, but it can be perceived through the results of experimental research. For instance, by the early work of Miller, in 1956, it was showed that short-term memory is limited in the number of elements it can contain simultaneously. Proposed Swellers theory, in 1986, treats schemas, or combinations of learning elements, as the cognitive structures that make up an individual's knowledge base and presents more insight towards it [17].

According to CLT, the quality of instructional design will be greater if attention is paid to the role and limitations of working memory. The total amount of mental activity placed on working memory in an instance of time, known as CL, has been proposed to have three distinct categories [23]:

1. *Intrinsic load* consists of the inherent complexity of the subject matter and manifests the level of difficulty of the material to be learned. As a simple instance, the mental calculation of $2+2$ has lower intrinsic load than solving an advanced algebraic equation, due to a higher number of elements that must be handled simultaneously (element interactivity) in working memory.

2. *Extraneous load* is the load applied by the elements of the instructional design itself. For example, an audiovisual presentation will usually have lower extraneous load than a visual-only format, since the audio form is also being used to carry information to the learner and ease the load experienced. It is

desirable to decrease extraneous load.

3. *Germane load* relates to the effort involved in processing and automating new information. Automation helps overcome working memory limitations and decreases CL. As an instance, knowledge and skills that are used frequently, such as reading, may be accessed automatically without high levels of conscious effort even though the associated task may be complex. It is also preferable to advance germane load.

CLT highlights several practices that can be applied to training and performance improvement. The most fundamental of these include methodologies for reducing the effects of the extraneous CL of instructional materials to ensure optimal leaning. These effects include split attention, redundancy, and modality [24].

It is known that performance improvement comes from the expansion of cognitive schemas (plans) and a high level of automation in the retrieval of information [25]. Because CLT views the limitations of working memory to be the primary hinderer to efficient and effective learning, reducing the total CL in instruction increases the portion of working memory available to succeed the learning process. This is could be attained by engineering reduced levels of CL through better instructional design [24].

4 The Measurement of Cognitive Load

As aforementioned, working memory is limited in both capacity and duration. Therefore, there is a clear need to asses and predict CL, and avoid underloading or overloading it. Since it may result in degrading performance or failures of learning and performing complex cognitive tasks, especially in critical decision making fields such as air traffic or military operations [1].

Due to the multidimensional construct of CL, it's assessment has been proven hard to do for researchers in the area [1]. It was known after Paas

et al. model in 1994, that CL can be assessed by measuring mental load, mental effort, and performance.

The available methods for classifying cognitive load measurement can be divided into two main classes. The first class comprises of *analytical methods* which are directed at estimating the mental load and collect subjective data. These methods rely on expert opinion or mathematical models and task analysis. The second class consists of *empirical techniques* which intend to estimate the mental effort and the performance. These methods can collect subjective data, performance data, and psycho-physiological data. Subjective data can apply rating scales based on the presumption that people can estimate their own mental load experienced and mirror their cognitive processes. It has been also proven that people are quite capable of scoring their perceived mental load. Usually, subjective techniques include a questionnaire with one or multiple semantic differential scales on which the subject can define the CL level experienced. Most subjective measures have multidimensional aspects so that they can also determine other related parameters, like mental effort, fatigue, and frustration, which are highly correlated [1].

According to Paas *et al.* model [1], most of subjective rating techniques deploy the psychologically oriented concept of overall load experienced due to the high reliability.

Physiological methods are based on the hypothesis that variations in cognitive functioning are manifested by physiological variables. These techniques could include measures of heart activity (e.g., heart rate variability), brain activity (e.g., task-evoked brain potentials), eye activity (e.g., pupillary dilation, and blink rate), and hormone levels (e.g., Catecholamine, Adrenaline, and Noradrenaline). Physiological measures can best be used to visualise the detailed trend and pattern of load (i.e., instantaneous, peak, average, over all, and accumulated load) [1].

But physiological methods have their own limitations, like any other

methods. Some of these measures provide only a weak link to CL (e.g., blink rate, blink duration); others are too slow for online measurement (e.g., hormone level), or insensitive to fluctuations of instantaneous load (e.g., heart rate variability [2]). Pupil dilation, which has none of these limitations, has been shown to be unsuitable for some tasks that involve continuous reading. There are also indications that the sensitivity of pupillary responses to changes in CL diminishes with participants age [19].

Therefore, among all the physiological methods, brain activity can be best to analyse CL experienced as it is the seat of cognitive activity. Unlike other neuroimaging devices, which require subjects to lie in restricted positions (e.g., fMRI or MEG), or to ingest hazardous materials (PET), Electroencephalograph (EEG) can noninvasively measure brain activity in authentic, real-world settings. EEG is a popular neuroimaging technique that measures electrical activity produced by the brain via electrodes that are placed on the scalp. These measurements vary predictably in response to changing levels of cognitive stimuli. This makes EEG an appropriate choice for measuring CL. Specially by recent advancements in EEG systems, such as wireless EEG, this technique has become more applicable to practical environments. Wireless EEG solutions like the B-Alert [26] system appear especially promising where as they offer better ecological validity by reducing the overall size of the equipment and allowing collection of data from multiple participants at the same time [19].

5 Electroencephalogram

Electroencephalography (EEG) is the recording of electrical oscillations along the scalp produced by the firing of neurons within the brain, in various frequencies [27].

In clinical field, the main application of EEG is epilepsy diagnosis, as

epileptic activity can manifest abnormalities on a routine EEG study clearly [28]. It is also extensively employed to manifest other neurological disorders such as head injury, stroke, sleep disorders, and tumour [29].

Evoked potentials (EP) is a derivative of the EEG. It relates to averaging the EEG activities which are time-locked to the *presentation of a stimulus/excitement* (i.e., visual, auditory, or somatosensory). Several behavioural disorders are diagnosed by analysis of EPs. This is carried out by averaging a number of successive trails for the same stimuli [30]. *Event-related potentials (ERP)* is averaged EEG responses that are time-locked to more complicated processing of stimuli. In fact, they are voltage fluctuations in the EEG induced within the brain. They appear as a sum of a large number of action potentials that are time locked to sensory, motor, or cognitive events. They are typically produced in response to external stimulations or excitements, and appear as somatosensory, visual, and auditory brain potentials, or as slowly evolving brain activity observed before voluntary movements or during anticipation of conditional stimulation. This technique is normally used in *cognitive science, cognitive psychology*, psychophysiological research, and most recently in brain computer interfacing (BCI) [30].

5.1   Clinical Use of EEG

EEG is widely used for diagnosis of many neurological disorders and abnormalities in the human, such as [29]:

- investigating epilepsy and locating seizure origin;

- monitoring alertness, coma, and brain death;

- locating areas of damage following head injury, stroke, and tumour;

- monitoring cognitive engagement (alpha rhythm);

- controlling anaesthesia depth (servo anaesthesia);

- testing afferent pathways (by evoked potentials);

- producing biofeedback situations;

- examining mental disorders;

- investigating sleep disorders and physiology;

Above applications confirm the high potential for EEG analysis and motivates the need for advanced signal processing techniques to aid the clinician in their interpretation.

5.2 Research Use of EEG

EEG signals offer an inexpensive and easy way of monitoring brain activities compared to other methods in the research field. This is worth repeating that unlike other neuroimaging devices, which require subjects to lie in restricted positions (fMRI), or to ingest hazardous materials (PET), EEG can noninvasively measure brain activity instantly in real-world settings. The EEG measurements vary sensitively in response to changing levels of cognitive stimuli which makes it an appropriate choice for assessing cognitive studies [19].

Some general advantages of EEG can be summarised as follows [30]:

- EEG hardware is significantly low cost;

- EEG is a silent device, allowing better study of the responses to auditory stimuli;

- EEG provides high temporal resolution (ms);

- EEG is tolerant of subject movement;

- EEG device is less bulky and more mobile, specially with the new wireless EEG available; and

- EEG usage needs less expertise.

As is the case with all devices, EEG has its limitations such as [30]:

- The low spatial resolution which is limited to the number of channel recording;

- EEG is subject to motion artefacts (e.g., blinking and movement);

- The electrodes attachment to the scalp which makes it uncomfortable for some individuals with tactile sensitivities, (e.g., some patients with autism).

5.3 Recording Methods

Routine EEG is the recording obtained by placing electrodes on the scalp with a conductive gel or paste. Most systems use caps or nets with electrode holders. The EEG recording electrodes and their proper function are vital for collecting high quality data. For standardisation purpose, the location of electrodes used in recording EEGs are defined by an international agreement known as $10-20$ system. This system was recommended by the International Federation of Societies for Electroencephalography and Clinical Neurophysiology. It is used for most clinical and research applications (except when high-density arrays are used) [31]. The system's placement of the EEG electrodes on the scalp is shown in Figure 1. According to this standard, even number electrodes are placed on the right side of the head and odd are placed on the left. The electrodes in this arrangement are placed along a bisecting line drawn from the nose (nasion) to the back of the head (inion), first at the position of 10% of the distance along the line, then at 20% intervals [32].
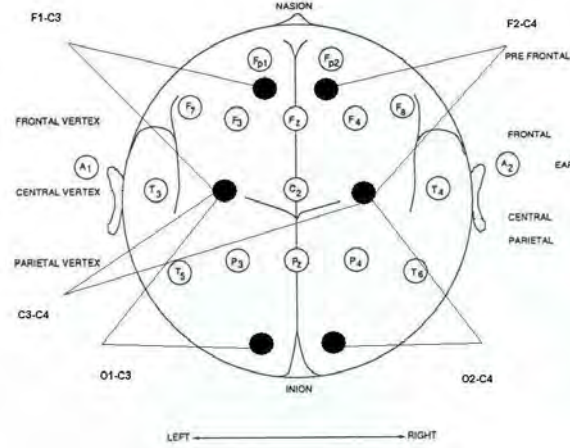
Figure 1: Scalp electrode placement according to the International $10-20$ system [31].

The notation F represents frontal lobe, C central scalp, P parietal lobe, and O occipital lobe. P_g is the naso-pharyngeal point (nose) and A is on the ear lobe. Given that the electrodes reside in the positions designated by the $10-20$ system, *montage* refers to the different ways in which the electrodes are connected to the EEG monitoring system [31].

The conventional electrode setting (so called $10-20$) includes 19 recording electrodes plus ground and system reference, totally 21 electrodes (excluding the earlobe electrodes) and is used widely in most clinical/research applications. A smaller number of electrodes are typically used when recording EEG from neonates. A larger number of electrodes might be applied when spatial resolution for a particular area of the brain is required. High-density arrays may include up to 256 electrodes evenly spaced on the scalp. In this case, extra electrodes are placed in between the standard electrodes with equi-distance between them. For example, C1 is placed between C3 and Cz

[30]. Extra electrodes are also used for the measurement of EOG, ECG, and EMG of the eyelid and eye surrounding muscles or ERP analysis and BCI. In such applications, the location of the corresponding electrode has to be determined well [30].

The raw EEG signals have low amplitudes (μV) and contain frequency components of up to 300 Hz. A typical adult human EEG signal is about $10\mu V$ to $100\mu V$ in amplitude when measured from the scalp and is about $10 - 20$ mV when measured from subdural electrodes [33].

To keep the effective information, the signals have to be amplified and filtered, to reduce the noise and make the signals suitable for processing and visualisation. The filters are designed in such a way not to change or distort the signals. Highpass filters have a cut-off frequency of usually less than 0.5 Hz, are used to remove the disturbing very low frequency components (i.e. breathing). High-frequency noise is also removed by using lowpass filters with a cut-off frequency of $50 - 70$ Hz. Notch filters are also applied to reject the strong 50 Hz power supply noise. In this case, the sampling frequency can be as low as twice the bandwidth commonly used by most EEG systems [30].

Each electrode is then connected to one input of a differential amplifier to amplify the voltage between the active electrode and the reference (typically up to 100000 times, or 100 dB of voltage gain) [31].

In the EEG recording, the specific electrode locations used for the creation of each channel (between two electrode inputs) is called *montage*. In another word, EEG voltage signal shows a difference between the voltages at two electrodes, of which set up could be done in one of several ways. The currently available methods are as follows [31]:

1. Bipolar montage: which measures the difference between two adjacent electrodes; i.e."C1-C3" representing the voltage difference between the C1 and C3 electrodes.

2. Referential (mono) montage: which measures the difference between a specific electrode and a designated reference electrode. A common reference is "linked ears" which is takes an average of electrodes attached to both earlobes or mastoids.

3. Average reference montage: in this case outputs of all of the amplifiers are summed and averaged over. Then it used as the common reference for each channel.

4. Laplacian montage: which implies the difference between an electrode and a weighted average of the nearby electrodes. This method is used to improve EEG spatial resolution, when average electrode spacing is less than about $3cm$ [34].

Two commonly used methods are referential and bipolar. The first one, is useful for detecting subtle features on adjacent electrodes, but is susceptible to noise and artefact. The latter, eliminates signals common to the two electrodes including certain artefact, but may also eliminate important EEG information. When recordings, these points should be taken into account. However, in clinical recordings, the montage may change several times in order to provide sufficient information for EEG analysis [32].

6 Characteristics of the EEG Signals

The interpretation of EEG records require knowledge from four related sources. *Firstly*, it is necessary to learn, recognise and classify EEG waveforms which have individual identity. *Secondly*, it is useful to have a theoretical framework for the analysis of the signals as it helps the understanding of both visual and automatic methods. *Thirdly*, spatial or topographical features of the EEG are required as location of the signals give important information

about their origin. *Fourthly*, information derived from empirical and clinical observations [35].

The raw EEG signals have amplitudes of $10-20$ μv, some times reaching amplitudes as high as 200 μv and contain frequency components of up to 300 Hz [33].

The EEG signals are interpreted through assessing three broad classes of features known as background activity (or normal activity), sharp transient (or paroxysmal activity), and ERPs. The last two classes are defined as abnormal activity [30].

6.1 Background Activity

Background activity is normal EEG which contains the stable EEG activity without any major temporal changes (stationary or quasi-stationary). It may be considered as a general indication of the excitability of the Central Nervous System (CNS). It is usually situated in the δ and θ regions for infants and children, but speeds up with age to be situated in the α region of adults. A slowing of the background rhythms normally occurs in sleep, in intense or anxious patients, as well as in patients on certain drugs such as anti-convulsants [36].

The normal EEG is typically represented by (1) *rhythmic activity* and (2) *transients*. The rhythmic activity is split into frequency bands. Frequency bands are usually extracted using spectral methods. Most of the rhythmic signals obtained on the scalp are within the range of 1-20 Hz. The normal EEG varies by age. In childhood it contains slower frequencies than that of the adult EEG. Therefore, the neonatal EEG is quite different from the adult. The normal EEG varies by state too, depending on sleep, awareness, disorder.... Therefore, the EEGs and their interpretation could be pretty different in different states [37].

The rhythms of the background EEG are classified into five main groups, depending on the predominant frequency contents [31]:

1. **Delta rhythm** (δ): describes any activity less than 4 Hz;

2. **Theta rhythm** (θ): resides in the frequency range of $4 - 8$ Hz;

3. **Alpha rhythm** (α): is a bilateral posterior rhythm which has an approximately constant frequency in the range of $8 - 12 Hz$;

4. **Beta rhythm** (β): resides in the frequency range of $12 - 30$ Hz and is not localised to any region of the scalp; and

5. **Gamma rhythm** (γ): describes the frequency activity above 30 Hz.

The rhythmic EEGs can also be described as below [31]:

Delta waves are not found in the EEG of the normal awake adult. They occur during deep sleep and in infancy.

Theta waves are found in the parietal and temporal lobes of children, but not usually in adults. They are found in some adults during emotional stress, particularly during disappointment or frustration.

Alpha waves of amplitudes of $20 - 200$ μv are found in almost all normal persons when awake and in a quiet resting state with eyes closed. However, opening of the eyes, attention to some other mental task or simulation results in the alpha rhythm being replaced by asynchronous waves of much lower amplitude. When asleep, there is no alpha rhythm.

Beta waves behave in a very similar manner to alpha waves, once again being inhibited by mental activity.

Gamma waves tend to appear only during periods of intense concentration or during tension or nervousness.

The background could consist of *continuous* or *discontinuous patterns*. A continuous background refers to a background with no distinct changes in

frequency or amplitude of waves. The above rhythms may last if the state of the subject does not change, therefore they are approximately cyclic in nature [30].

The normal EEG could be *transient* too. Apart from spikes and sharp waves that may represent seizure activity or interictal activity in individuals with epilepsy or towards epilepsy. Other transient features are normal such as: *ERPs* that contain positive occipital sharp transient, *vertex waves* and *sleep spindles* which are observed in normal sleep. It should also be noted that there are types of activity which are not statistically defined but are not related to any dysfunction or disorder. These are called *normal variants* of which the *mu rhythm* is a known example [30].

The EEG is also used to study sleep stages. It is proven that sleep is related to memory and working memory can be affected by sleep deprivation. Working memory is important because it retains information active for further processing and supports higher-level cognitive functions such as decision making, reasoning, and episodic memory [38].

6.2   Abnormal Signals

Variations in the EEG patterns for certain states of the subject indicate abnormality. This may be due to distortion and the disappearance of abnormal patterns, appearance and increase of abnormal patterns, or disappearance of all patterns. There are different classifications of abnormal EEG signals. They can be split into epileptiform and non-epileptiform activity, into focal or diffuse, an into sharp waves and slow waves. *Abnormal sharp waves* are predominantly of two types; *paroxysmal events* and *ERP*s [31, 39].

Paroxysmal events in the EEG pattern are suddenly appearing electrical explosions and refer to patterns of electrical activity occurring with a sudden onset and termination. They are typically short, frequent and stereotyped

symptoms that can be observed in various clinical conditions, such as: head trauma, stroke, asthma, breath-holding spells, epilepsy, malaria, and so forth. Exercise, tactile stimuli, hot water, anxiety and neck flexion may provoke paroxysmal attacks too [31].

Evoked potential events/Evoked related potentials, or ERP, is the brain electrical activity that results from a change in an ongoing neural activity created by the stimulation of a sensory organ or pathway. They are usually measured with EEG. ERP are caused by the higher processes of the brain, that might involve memory, expectation, attention, or changes in the mental state, among others [40].

The following terminology is associated with abnormal sharp waves in EEG signals:

- **Abnormal paroxysmal events**: are short, frequent and stereotyped symptoms that can be observed in various clinical conditions, such as: head trauma, stroke, asthma, breath-holding spells, epilepsy, malaria, and so forth. Exercise, tactile stimuli, hot water, anxiety and neck flexion may provoke paroxysmal attacks. They are typically different from other transient symptoms by their brevity (lasting no more than 2 mins), frequency (from $1 - 2$ times/day up to a few hundred times/day), stereotyped fashion and excellent response to drugs (usually carbamazepine). Withdrawal of symptoms without any residual neurological finding is another key feature in their recognition.

  Adults paroxysmal patterns show higher amplitude than the background which is not always the case in newborns. It frequently consists of distinct morphological transients which are almost ictal or inter-ictal (between seizures) and imply an *epileptic* process even if seizure does not occur. Some normal and abnormal paroxysmal features occurring in the EEG of the newborn carry no message, however, identical pat-

terns are truly ictal [41].

- An **ERP**: is any measured brain response to an internal or external stimulus. They are usually measured with EEG. As the EEG reflects thousands of simultaneously ongoing brain processes, the brain response to a single stimulus or event of interest is not usually visible in the EEG recording of a single trial. Therefore, to observe the brain response to the stimulus, the experimenter must conduct many trials (100 or more) and average [1] the results together, causing random brain activity to be averaged out and the relevant ERP to remain.

  While EPs reflect the processing of the physical stimulus, ERP are caused by the *higher* processes, that might involve memory, expectation, attention, or changes in the mental state, among others [40]. The MEG counterpart of ERP is the ERF, or event-related field.

  Slow waves such as the δ waves in the awaken adults often appear when the brain cells have been damaged. Where there is focal damage of the cortex or white matter, there is an increase in slow frequency rhythms and/or a loss of normal higher frequency rhythms. It may also appear as focal or unilateral decrease in amplitude of the EEG signal. These signals are out of this research scope.

6.3 Event-Related Potential

As mentioned previously, an ERP is any measured brain response to an internal or external stimulus. It is usually a negative or positive-going EP with well defined peaks, latencies, and topography. It can be separated into two classes, as below:

---

[1]The use of averaging improves the signal-to-noise ratio.

Clinical ERP: Physicians and neurologists will sometimes use a flashing visual checker board stimulus to test for any damage or trauma in the visual system. In a healthy person, this stimulus will elicit a strong response over the primary visual cortex located in the occipital lobe in the back of the brain.

Research ERP: An EP or evoked response is an electrical potential recorded from the nervous system of a human following presentation of a stimulus, as distinct from spontaneous potentials as detected by EEG or EMG. EP amplitudes tend to be low, ranging from less than a μV to several μV, compared to tens of μV for EEG, mV for EMG, and often close to a volt for ECG. To resolve these low-amplitude potentials against the background of ongoing EEG, ECG, EMG and other biological signals and ambient noise, signal averaging is usually required. The signal is time-locked to the stimulus and most of the noise occurs randomly, allowing the noise to be averaged out with averaging of repeated responses [42].

Experimental psychologists and neuroscientists have discovered many different stimuli that elicit reliable ERPs from participants. The timing of these responses is thought to provide a measure of the timing of the brain's communication or time of information processing.

The stimulus presented could be visual, tactile (relating to the sense of touch), auditory, olfactory (relating to the sense of smell), gustatory (relating to the sense of taste), etc. Because of this general invariance in regard to stimulus type, this ERP is understood to reflect a higher cognitive response to unexpected and/or cognitively salient stimuli. An example wave with several ERP components (with positive and negative deflections) is shown in Figure 2.
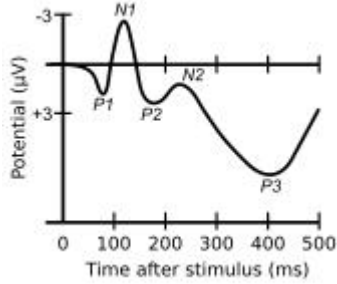
Figure 2: A wave with several ERP components [43].

More common ERPs are the visual evoked potential (VEP) and the auditory evoked potential (AEP) which are used for studying the visual and auditory systems, respectively. The parameters of interest relating to the ERP are usually the latency (or delay of occurrence) and amplitudes of its various components. However, the detection of the ERP is extremely difficult, as it is usually embedded in the ongoing background EEG activity. The ratio of the ERP power to the power of the ongoing background EEG activity typically ranges from SNR of $0 - 60 \ dB$. The detection of ERP parameters is made easier though, by the fact that the response to repetitive stimuli is usually available. In this case, the improvement of the SNR of the ERP is made possible through the use of some form of ensemble average [32].

In Figure 2, the N1 (or N100) stands for a large, negative-going EP that peaks in adults between 80 and 120 ms (approximately 100 ms) after the onset of a stimulus, and distributed mostly over the front to central region of the scalp. The N1 mostly focuses on auditory stimuli, but also occurs for visual, olfactory, heat, pain, balance, respiration blocking, and somatosensory stimuli [44].

In Figure 2, also, the P3 (or P300) stands for a positive deflection

21

in voltage at a latency of roughly 300 ms in the EEG. It is typically measured most strongly by the electrodes covering the parietal lobe. The presence, magnitude, topography and time of this signal are often used as metrics of cognitive function in decision making processes. The P3 itself is thought to be comprised of two wavelets known as P3a and P3b signals. These components respond individually to different stimuli, and it has been suggested that the P3a wave originates from stimulus-driven frontal attention mechanisms during task processing, whereas P3b originates from temporalparietal activity associated with attention and appears related to subsequent memory processing. The two wavelets are sometimes referred to as non-target (P3a) and target (P3b) ERPs. In practice, the P300 waveform must be evoked using a stimulus delivered by one of the sensory modalities. One typical procedure is the *oddball paradigm*, whereby a target stimulus is presented amongst more frequent standard background stimuli. A distracter stimulus may also be used to ensure that the response is due to the target rather than the change from a background pattern. In this case, the task is called three-stimulus task (similar to the oddball with a compelling distracter (D) stimulus that occurs infrequently). In each task, the subject is instructed to respond only to the target and otherwise to refrain from responding. The distracter elicits a P3a, and target elicits a P3b (P300). A schematic illustration of the three stimulus is given in Figure 3.

The classic oddball paradigm has seen many variations, but in the end most protocols used to evoke the P300 involve some form of conscious realisation or decision making. Attention is required for such protocols [15].

It is known that the P300 component indexes brain activities underlying revision of the mental representation induced by incoming stimuli.
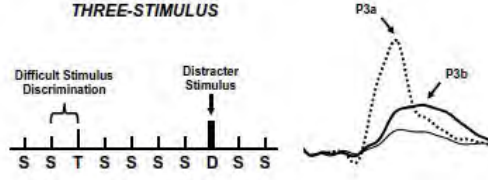
Figure 3: The distracter elicits a P3a, and target a P3b (P300) [15].

After initial sensory processing, an attention-driven comparison process evaluates the representation of the previous event in working memory a process distinct from although related to sensory stimulus feature mismatch detection [15, 45]. The schematic of the theory is shown in figure 4. If the incoming stimulus is the same, the neural model of the stimulus environment is unchanged, and sensory EPs (N100, P200, N200) are obtained after signal averaging. If the incoming stimulus is not the same and the subject allocates attentional resources to the target, the neural representation of the stimulus environment is changed or updated, such that a P300 (P3b) potential is generated in addition to the sensory EPs.

6.3.1   Event-Related Potential Models

Generally, amplitude of the EPs are low when merged with the background brain waves. Therefore, mathematical tools are needed to analysis the time series EEGs to discover the EPs in the background EEGs. In [46], an overview was done on spectral analysis of the EEGs to distinguish the classes, which correlate to the brain activities, with respect to olfactory stimulus. It is a general practice to work out wavelet transform (WT), fast Fourier transform (FFT), and fractal dimension (FD) of the EEGs to predict the EPs. The generally used technique

23

CONTEXT UPDATING THEORY OF P300

WORKING MEMORY
COMPARISON

NEURAL
REPRESENTATION

INPUT

DIFFERENT
STIMULUS?
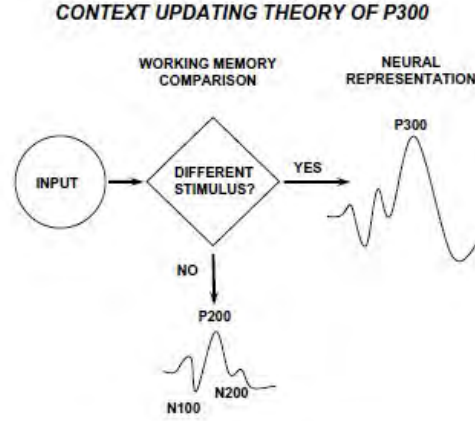
YES

P300

NO

P200

N100

N200

Figure 4: If the incoming stimulus is not the same P300 (P3b) potential is generated [15].

for analysing spontaneous EEG is the Fourier power spectrum (FFT power spectrum). Many researches showed that FFT power spectrum could be reliably used to detect pathological activity of seizure from the EEGs. But, this research reported that FFT power spectra could not predict EPs in the background EEGs stimulated on inhalation of six different odors. Also, in [47], it was demonstrated applying FFT, to find the cognitive ability of the brain dynamics could be misleading. Indeed, where the brain is a highly complex non-linear system, it should be processed through non-linear analysis, while FFT analysis is a linear based approach. It was concluded, discrepancies between the two systems and inaccuracies were possible when EPs were subjected to Fourier spectral analysis, even though, FFT is largely used as a clinical tool. It was also demonstrated that wavelet transform (WT) spectra give more information of the EEG activities than that of FFT. However, the WT is also a linear analysis and contradictions are expected. In fractal dimension (FD) analysis, a quantitative measure

24

of self-similarity [2] is considered. Therefore, fractal is a ever-finer fragmented pieces quantified by a dimension. Correlation dimension and information dimension are methods used to quantify fractals, referred to as FD. It was illustrated that FD provided information on the brain activities, a well-delineated phase with increased FD and slow wave activity with decreased FD. FD and fractal spectra analysis predicted the brain activities more precisely than FFT.

6.4   EEG Generator

The potential recorded over the cortex electro-corticogram (ECoG) or over the scalp (EEG) derives from the activity of many sources known as *EEG generators*. The recorded amplitude is basically a function of the unitary potential of a generator and the statistical relationship between different EEG generators in the recorded population [48]. Since early days of clinical EEG, it has been well known that the EEG represents summated action potentials of cortical neurons [49]. Further experiments corroborated the fact that extracellularly recorded action potentials of cortical cells often show a statistically significant relationship with EEG sharp waves such as spindle waves or paroxysmal potentials [50]. However, the fact that EEG waves could also be observed in states when no action potentials were present (such as during deep anaesthesia or hypoxia), demonstrated that action potentials could not be the only source of slow surface potentials. The existence of dendritic potentials (for slowly propagating potentials in the superficial layers of the cerebral cortex) suggested a relationship between EEG waves and

---

[2]In fractal geometry, self-similarity is an important concept (nerve cells are statistically self-similar in space. Concurrently, electrical voltage across the cell membrane of a T-lymphocyte cell is an example of statistical self-similar in time)

activity in the apical dendrites of cortical neurons [51, 52].

The close relationship between postsynaptic potentials and surface EEG was later demonstrated by intra- and extracellular recordings from individual cortical cells [53]. Further experimental studies provided an estimate of the size of unitary generators of the EEG and suggested that these generators must have cellular dimensions [54]. The theory that cortical slow waves are constituted primarily by postsynaptic potentials of cortical neurons is widely accepted today.

From an electrical activity point of view, the electrical activity of the brain can be described in spatial scales from the currents within a single dendritic spine to the relatively gross potentials that the EEG records from the scalp. Neurons, (nerve cells), are electrically active cells which are primarily responsible for carrying out the brain's functions. Neurons create action potentials, which are discrete electrical signals that travel down axons and cause the release of chemical neurotransmitters at the synapse, which is an area of near contact between two neurons. This neurotransmitter then fits into a receptor in the dendrite or body of the neuron that is on the other side of the synapse, the post-synaptic neuron. The neurotransmitter, when combined with the receptor, typically causes an electrical current within dendrite or body of the post-synaptic neuron. Thousands of post-synaptic currents from a single neuron's dendrites and body then sum up to cause the neuron to generate an action potential (or not). This neuron then synapses on other neurons, and so on. EEG reflects correlated synaptic activity caused by post-synaptic potentials of many cortical neurons. Figure 5 illustrates the major elements in a synaptic transmission. An electrochemical wave called an action potential travels along the axon of a neuron. When the wave reaches a synapse, it provokes release of a puff of neurotransmitter molecules, which bind to chemical receptor

molecules located in the membrane of another neuron, on the opposite side of the synapse. More specifically, the scalp electrical potentials that produce EEG are generally thought to be caused by the extracellular ionic currents caused by dendritic electrical activity, whereas the field potentials producing MEG signals are associated with intracellular ionic currents [55, 56].

EEG activity therefore always reflects the summation of the synchronous activity of thousands or ons of neurons that have similar spatial orientation, radial to the scalp. Currents that are tangential to the scalp are not picked up by the EEG. The EEG therefore benefits from the parallel, radial arrangement of apical dendrites in the cortex. Because voltage fields fall off with the fourth power of the radius, activity from deep sources is more difficult to detect than currents near the skull [58].

Scalp EEG activity shows oscillations at a variety of frequencies. Several of these oscillations have characteristic frequency ranges, spatial distributions and are associated with different states of brain functioning (e.g., waking and the various sleep stages). These oscillations represent synchronised activity over a network of neurons. The neuronal networks underlying some of these oscillations are understood (e.g., the thalamocortical resonance underlying sleep spindles), while many others are not (e.g., the system that generates the posterior basic rhythm) [58].

The processes of generation of the EEG signal in a large number of neurons forming a neural mass are complex. This led to both analytic and synthetic approaches to such processes. So far, many models have been proposed to describe the generation of the EEG (by Freeman in 1986, Rotterdam in 1982, Lopes d Silva in 1976, Liely 2003 and etc), which have some drawbacks.
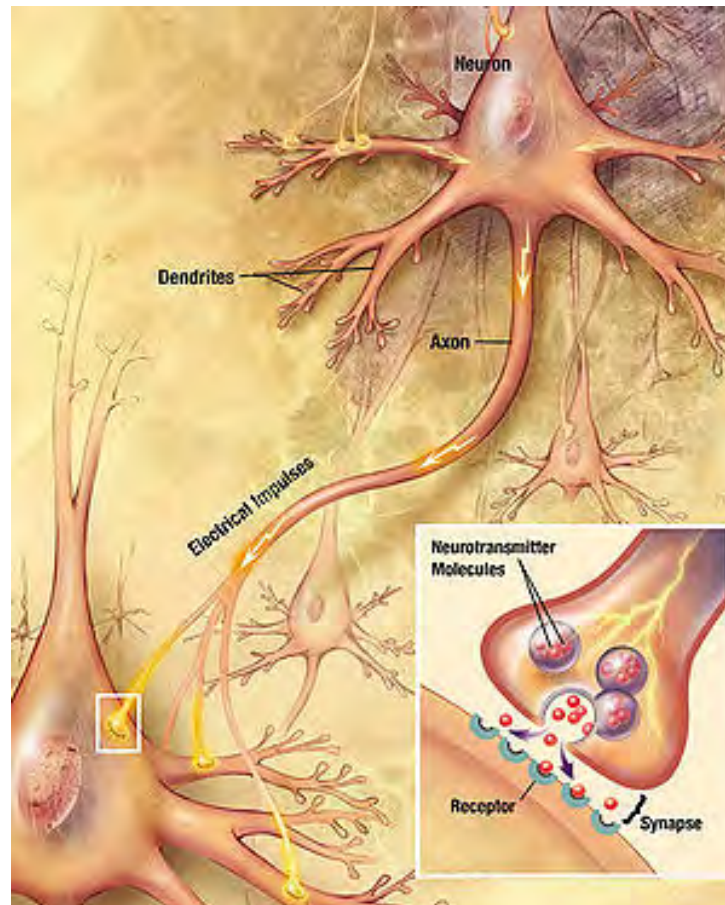
Figure 5: Structure of a post-synaptic neuron [57].

A new definition of a generator that may overcome previous models' difficulties, has been proposed in [48]. It defines a generator as an axon and the cloud of synapses lying along its branches. That is, the unitary activity of the generator is a cloud of synapses lying along a specific axon activated each time an action potential is conducted along the axon. This definition of the axon and its cloud of synapses solves the problem of dependency between generator activity. The triggering event initiating generator activity is the action potential conducted along the axon, and firing trains that are conducted through the axons of two neurons sharing a common source may have a wide range of correlation values. Thus activity in adjacent generators may be independent. In addition, by using this definition of a generator, it is possible to directly estimate the rate of triggering events activating the generators by recording action potentials via extracellular microelectrodes. Furthermore, by analysing extracellular recordings, the correlation coefficients between driving forces and evaluating other activation statistics' parameters of generator activation can be estimated [48]. The definition of a generator as an axon and its cloud of synapses also overcomes the problem of uniformity because each time an action potential is produced, it is conducted along the axon and its processes, and all the synapses lying along this axon and its processes are activated by that specific action potential. Thus the activity of the generator is uniform each time it is activated. However, it should be noted that this generator is not a point in space but rather a volume defined by the axons processes in space. A simplified and idealised compartmental model of a neuron is used in this study [48]. This model is based on the cell membrane potential formulated by shown by formula 6.4.

$$V(\vec{r_0}) \simeq -\frac{\sigma_i}{4\pi\sigma_e} \int_{surf} V_m(\vec{r}) d\Omega(\vec{r} - \vec{r_0})$$

where $V(\vec{r_0})$ is the potential at a point $\vec{r_0}$ in the volume conductor due to the injected current densities, where σ_i is the intracellular and σ_e is the extracellular conductivity and $d\Omega(\vec{r} - \vec{r_0})$ is the solid angle subtended by an infinitesimal surface on the membrane surface and seen from the extracellular point $\vec{r_0}$.

The neuron consists of a cylindrical soma, long cylindrical apical dendrite, and three equally spaced cylinders representing basal dendrites extending from the bottom of the soma cylinder. Figure 6 (Part A) shows the schematic morphology of the neuron where each compartment (cylinder) is represented by a line that is the main axis of the cylinder it represents. The schematic morphology consists of a long apical dendrite, soma, and three basal dendrites forming a symmetric structure (each cylinder is represented by a line that is the cylinders main axis. The 3 lower lines represent 3 basal dendrites, the long vertical line is the main axis of the soma and the apical dendrite. The electrodes are placed in circles with a gradually increasing radius, where the centre of each circle is located just above the main axis of the apical cylinder. Figure 6 (Part B) shows the structure of three basal dendrites (spaced 120 from each other) as though we are looking from a surface point on top of the apical dendrite. Every segment of the apical or the basal dendrites was defined as a potential location for termination of a synapse. A single synapse was activated each time on a different segment of the modelled neuron. For each activation of a synapse, all the segments membrane potentials were as a function of time. The postsynaptic potential amplitude depends on the position of the acti-
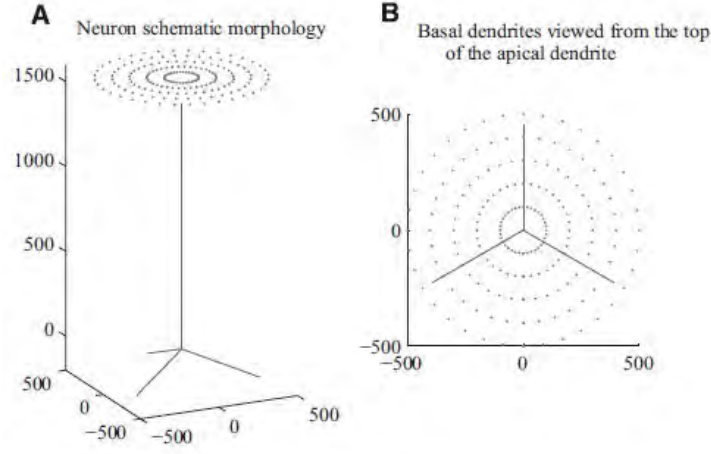
Figure 6: Neuron schematic morphology [48].

vated synapse along the neuron and on the position of the point we are measuring this amplitude.

By activating any synapse on the modelled neuron and having its effect on the trans-membrane potential, we can assess the potential each activated synapse contributes to the surface potential. The computation of the surface potentials consists of the following phases: triangulation of the neuron's membrane surface area, and contribution of additional patches of membrane surface. The surface potential was computed due to activation of a single synapse lying on top of the apical dendrite at several surface points. The potentials shown in Figure 7 (Part A) are the potentials computed from a set of surface electrode points with gradually increasing distances from the soma of the activated neuron. As expected, a negative wave due to activation of the top apical synapse was obtained. The peak amplitude of the waves decreased as the electrode was moved farther away from the neuron. Similarly, the surface potential due to activation of a basal synapse was computed

31

using the same array of electrode points as before (Figure 7 (Part B)). The potentials showed positive waves. As seen in the potentials, due to activation of the apical synapse, the decrease in amplitude as the surface electrode moved farther from the neuron can again be seen.

In the presented model, the contribution of a synapse activation terminating on a large pyramidal neuron to the surface potential was calculated. Then an extension of this model from a surface electrode listening to the activation of a single synapse terminating on a single cell to an electrode listening to a population of synapses terminating on a population of neurons lying in a certain brain volume (a cylinder) was presented. This extension assumed that the volume of the treated cylinder only consists of large pyramidal cortical neurons. Clearly, this would lead to a larger RMS than if we replaced some of the large pyramidal neurons by smaller ones representing layer 2 and 3 pyramidal neurons. It is shown that low RMS values are obtained because of completely desynchronised activity.

To sum up, in [48] a new definition for the term EEG generator has been presented. This definition makes it easy to find the activity rate of a generator and the correlation between given generators simply by using multiple parallel recordings of single-unit activity. Relying on this new definition, we can examine the relationship between generators activity, generators statistics, and the EEG/ECoG signal recorded in the preceding text. In this study, the contribution of a single generator to the variance of the recorded signal has been calculated and showed that under the condition of completely unsynchronised activity, low values of RMS as expected was obtained. This calculation of the RMS value can be made for any statistical organisation of generators, from completely unsynchronised to fully synchronised populations.
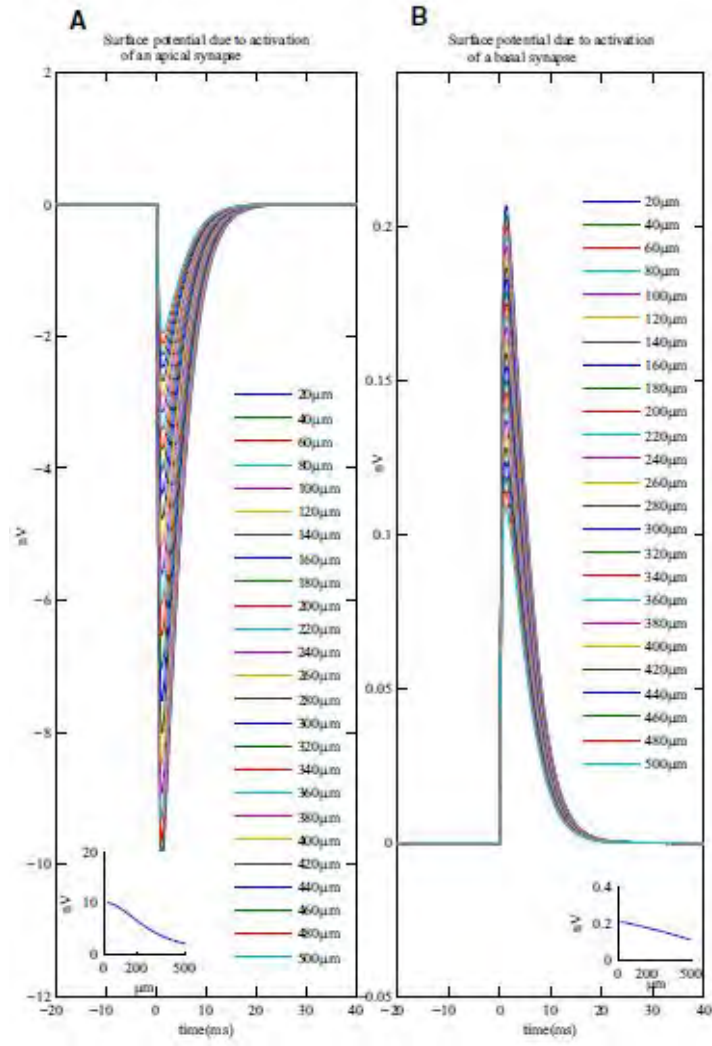
Figure 7: Surface potential recorded by an array of electrodes [48].

7 Cognitive Load Measurement Using EEG

Application of physiological methods, in particular EEG, offers new and promising approaches to CL measurement. EEG is identified as a physiological index that can serve as an online, continuous measure of CL detecting subtle fluctuations in instantaneous load. This can help explain effects of instructional interventions when measures of overall cognitive load fail to reflect such differences in cognitive processing in educational psychology field [19].

A comprehensive review of EEG oscillations reflecting cognitive and memory performance was conducted in [59], to correlate the EEG signal oscillations with a focus on α and θ activity. It was shown that EEG power is indeed related to cognitive and memory performance, but in a complex and partly non-linear way. Within the α frequency range EEG power is positively related to cognitive performance and brain maturity, whereas the opposite holds true for the θ frequency range. α and θ reactivity as well as event-related changes in their band power show yet another pattern of results. According to the results, during actual task demands the extent of α power suppression is positively correlated with cognitive performance and memory performance, in particular. Whereas again the opposite holds true for the θ band. In this research, the extent of θ synchronisation is related to good performance. The increase in lower α power may reflect an attempt to increase cognitive performance too. This research provides a good and reliable insight towards frequency band changes associated with CL.

In another related study [60], it was revealed that there are two types of changes in the α and θ frequency range. One is called tonic changes, which is a response to circadian rhythms, fatigue, distress, neurological disorders, etc. These changes are less under volitional control and occur

at a much slower rate. On the contrary, phasic changes are task or stimulus related or (event-related) changes in the EEG. These are more or less under volitional control and occur at a rapid rate. A phasic change is measured as an increase or decrease in band power during task performance as compared to a reference or resting period. The results of spectral analyses show that compared to closed eyes, α power decreases but θ power increases. At this point, there is an interesting paradox. During actual cognitive performance (as compared to a resting state), the EEG is characterised by increased θ but decreased α power and thus resembles the EEG during a tonic change that reflects decreased cognitive performance. It was shown that large α power which is correlated with a pronounced decrease in event-related band power and small θ power which is correlated with a pronounced increase in band power indicate good cognitive.

In [6], the importance of minimising working memory load and the need to apply a proper load monitoring system was investigated during human-computer interaction. In this study, it is emphasised that applied contexts require that a candidate method for monitoring working memory load that: *firstly*, must not interfere with operator performance, *secondly*, must be employable across many contexts, *thirdly*, must have reasonably good time resolution and *fourthly*, must be robust enough to be reliably measured under relatively unstructured task conditions yet sensitive enough to consistently vary with some dimension of interest. Fortunately, EEG appear to satisfy all these requirements. In fact, EEG measures have been shown to be highly sensitive to variations in task difficulty. In addition, multivariate combinations of EEG variables can be used to accurately discriminate between specific cognitive states. Therefore, an experiment was designed here to identify three working memory levels, and to examine how classifica-

35

tion networks could be generalised across tasks and participants. In this regard eight healthy volunteers underwent a matching task, in two different versions of verbal and spatial, at each of three levels of difficulty. In a verbal version, they were required to remember the identity of the visual stimulus presented. In the low level of difficulty (LL), they were asked to compare the current stimulus with the previous stimulus, in the medium level (ML) to compare with the one presented two trials ago, in the high level (HL) to compare with the one presented three trials ago. In a spatial version, they were required to remember the position of the visual stimulus at different levels, with respect to the above procedure. This study demonstrated a NN based classification method was capable of discriminating task-imposed CL by high percentage correct rate of %91 for LL, %94 for ML, and %88 for HL. It also showed α, θ, and β were quite sensitive to CL difficulty levels.

A fundamental research with respect to event-related synchronisation and desynchronisation in EEG/MEG was carried out in [61]. According to this study, an internally or externally paced event results not only in the generation of an ERP but also in a change in the ongoing EEG/MEG in form of an event-related desynchronisation (ERD) or event-related synchronisation (ERS). Since, ERPs represent frequency specific changes of the ongoing EEG activity and may consist either of decreases or of increases of power in given frequency bands. This may be considered to be due to a decrease or an increase in synchrony of the underlying neuronal populations, respectively. The ERP on the one side and the ERD/ERS on the other side are different responses of neuronal structures in the brain. While the former is phase-locked, the latter is not phase-locked to the event. The most important difference between both phenomena is that the ERD/ERS is highly frequency band specific, whereby either the same or different locations on the

scalp can display ERD and ERS simultaneously. Traditionally, ERPs that can be considered as a series of transient post-synaptic responses of main pyramidal neurons triggered by a specific stimulus and represent the responses of cortical neurons due to changes in afferent activity. ERD/ERS phenomena can be viewed as generated by changes in one or more parameters that control oscillations in NNs. They reflect changes in the activity of local interactions between main neurons and inter-neurons that control the frequency components of the ongoing EEG.

A central advantage of EEG over behavioural indices and other neu-roimaging methods lies in its high temporal resolution, which also allows for the assessment of the time-course of cognitive load during execution of the learning task (instantaneous cognitive load). However, it needs to be kept in mind that ERD/ERS values require averaging over time or trials to yield a satisfactory reliability. Finally, it needs to be emphasised that EEG, like any other method, does not allow to distinguish whether the load measured is evoked by processes that are germane or extraneous to learning. Therefore, it remains important to analyse the results of EEG measures of cognitive load in the context of the associated learning performance [19].

8 Signal Processing Methods for Cognitive Load Measurement

The EEG signals, like most biological signals, are inherently difficult to quantify. These signals can be characterised as non-stationary, with highly complex time-frequency characteristics, and often with a low signal-to-noise ratio (SNR). They are irregular in nature and never replicate themselves exactly. This type of signals therefore can not

reasonably described as deterministic signals and should be studied through statistical methods [32].

The available methods for EEG signal processing in the literature can be divided into two main classes. The first class consists of parametric methods which assume a model for the EEG signals and then estimate the parameters of the model.

The second class includes non-parametric methods which do not require a model of the signal. This class may be further classified into three sub-classes as follows:

1. **time-domain** methods,

2. **frequency-domain** methods, and

3. **time-scale** (TS) and **time-frequency** (TF) methods.

The first and second methods use the characteristics of EEG signals in time or frequency-domain, respectively, as the distinction criteria. These methods are based on the assumption that the EEG signals are quasi-stationary [3]. However, many works show that the EEG signals exhibit non-stationary features [62, 63, 64]. A close examination of these signals clearly shows that they exhibit significant non-stationary and multicomponent features, as observed in Figure 8. In this Figure a TFD plot of an analysed EEG segment is depicted to show that the frequency content of the analysed EEG segment changes with time. This fact supports the idea of using TS and TF methods to be the most proper methods for processing of EEG signals. These methods take into account the non-stationary nature of the EEG signals and use its characteristics in time and frequency-domain simultaneously.

---

[3]The frequency content of the EEG signals are assumed to be constant during the short analysis window, as short as 2 s window length.
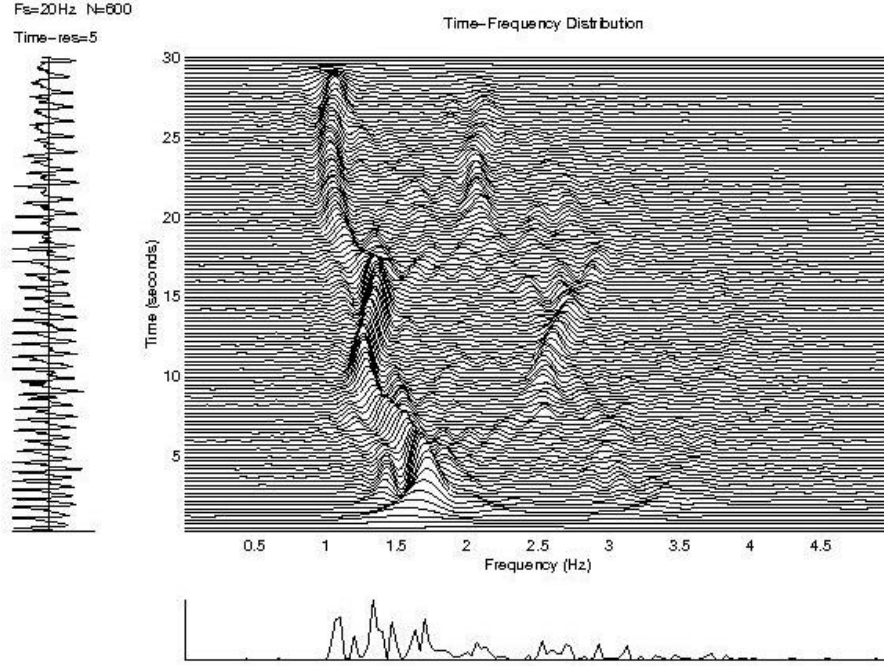
Figure 8: A TFD plot of an EEG segment.

In the next subsection, existing TF and TS methods in the literature for EEG processing with regard to CL/mental load are outlined and their relative merits are discussed.

8.1 Time-Frequency Methods

Processing and classification of the EEG signals has become an active topic of research in the area of signal processing as well as machine learning. It has also found its applications widely in brain computer interface (BCI) and medical diagnoses areas.

In [65], Kullback-Leibler (KL) divergence was used in the classification

of raw EEG signals. It was shown that k-nearest neighbour (k-NN) algorithm with KL divergence as the distance measure, when using the feature vectors, outperformed the more commonly used Euclidean k-NN. In this study, the EEG data was recorded while the subject performed 5 different mental activities such as; math problem solving, letter composing, 3-D block rotation, counting, and resting (baseline). The data was recorded from 6 EEG electrodes, at 250 Hz sampling freq. and for 10 s during each mental task. Each task was repeated 5 times per session and the subjects attended two such sessions resulting in 10 trials per task per subject. The EEG of each task was then windowed and feature vectors were derived from the power spectrum of these 2 s windows. Each feature vector consisted of normalised power spectral density (PSD) in the frequency range of $8 - 32$ Hz of each of the 6 channels concatenated together. Two classification algorithms for classifying the mental tasks, namely, k-NN and support vector machine (SVM) was used here. The underlying distance of choice for both methods was the KL divergence. Classification was done by the two distinct classifiers (the Euclidean distance 5-NN classifier and the KL-distance 5-NN classifier), using different sets of activities. The feature vector and distance measures were tested in pairwise classification (only two classes were used at a time for the classification). The highest distinction rate was achieved between the baseline and rotating task at %98, followed by the baseline and math task at %95, when using KL-distance based 5-NN classifier. The given results proved distance measure outperformed the Euclidean distance in k-NN and the SVM using a KL distance based kernel outperformed linear kernels.

The research carried out in [7] investigated the CL correlation with dimensional complexity of the EEG signals. The difference in neurones firing modes can be expressed in the correlation dimension of

the EEGs. In this study, 12 subjects performed 3 different tasks with different difficulty levels. In the baseline task, they were asked to sit relaxed and shut their eyes for 5 min. In the estimation task (as a medium level task), a demonstration of 20 s was presented and they were asked to reproduce the interval by pressing a button. The calculation task, consisted of adding up two numbers continuously during 20 s. The tasks were verified by a subjective rating scale. The correlation dimension [4] was estimated using a correlation integral to assess the scaling region. Regression analysis was then applied to this region to estimate the correlation dimension. To assess the subjective mental loads, means and SDs of EEGs were extracted. The results showed that the calculation task had induced a higher mental load than the time estimation task and the baseline. The difference in correlation dimension between calculation and time estimation was also significant, with the highest value for calculation task. It was concluded that cognitive and mental activity was associated with a higher correlation dimension in the EEG. This implied that the correlation dimension was a sensitive parameter in the analysis of the electrical brain activity. It was also confirmed that the correlation dimension of the EEG was higher in conditions that require greater mental effort, as a significantly higher correlation dimension was established during the two tasks than the baseline period.

An automatic augmented cognitive state estimation based on the EEG signals for ambulatory users was discussed in [66]. The experiment involved a user outfitted with wearable monitoring, communication, and mobile computing equipment walking outside. The monitoring equipment was an ActiveTwo EEG system, recording EEGs at sampling freq.

---

[4]The correlation dimension seems to reflect the number and the complexity of cognitive process that take place in the brain.

of 256 Hz from seven channels (CZ, P3, P4, PZ, O2, P4, F7) while the subject was walking about and performing various prescribed tasks. A bandpass filter between $2 - 50$Hz was employed, as this interval is generally associated with cognitive activity. The PSD of the EEG signals, estimated using the Welch method with 1 s windows was integrated over 5 frequency bands of θ, α, lower β, upper β, and γ. To effectively reduce this high dimensional vector of the PSD features, mutual information filtering was applied. The reduced feature vectors were then used as inputs to a committee of classifiers: Gaussian-Mixture-Model (GMM) , KNN , and non-parametric Kernel Density Estimate (KDE) to obtain estimates of the cognitive state. The committee decision was the majority vote and was offered in real-time at a rate of 10Hz. To eliminate fast fluctuations in the cognitive state assessment, a median filter was employed to smoothen the final decision over a sliding interval of 2 s with the assumption that the CL does not vary faster than the corresponding rate and that the integer class labels were assigned to cognitive tasks (the classes) in correlation with their actual corresponding CLs. Cognitive states were associated with three tasks in two modalities: when the subject was stationary or mobile. The tasks in the stationary case were labelled relaxed (waiting for orders), communicate (getting orders from base via radio communication), and count (starting from 100 and decreasing by 7). The tasks in the mobile case were labelled navigate (walking to a designated target), navigate and visual search (walking while looking for snipers), and navigate and communicate (receiving and giving mission status reports). The classification results demonstrated that this method was effective in determining a lower dimensional projection (3 dimensional input) that achieved at least the same performance as the original high dimensional feature vector (35 dimensional input). The described cognitive state es-

timator was employed in a real-time closed-loop adaptive performance enhancement scheme that scheduled the communication traffic to the subject during a mission. It also showed that the assistance offered by this interface improved task-related performance greatly based on the reduced and robust features selected. For instance, the scheduling of communication based on the CL assessment resulted in %100 improvement in message comprehension and %125 improvement in situation awareness. (This paper was reviewd due to the methodolody and similar EEG data acqusuion device!).

Three spectral estimation techniques for classification of mental tasks have been investigated in [67]. These techniques are parametric Burg method, and non-parametric standard periodogram, and Welch periodogram methods. The EEG signals were collected from six healthy subjects using six electrodes; F3, F4, C3, C4, P3, and P4 according to the 10-20 positioning system, at a sampling rate of 150 Hz. A Butterworth bi-directional bandpass filter, of order 10, was set at $6-40$ Hz. The subjects were asked to perform three mental tasks as follows; relax task (to close the eyes, relax and think of nothing), math task (to make a regressive count from 30 with steps of 3), imagine task (to imagine an incandescent lamp that is on). The utilised freq. bands included upper θ, α, β, and γ. The γ was used here due to its contribution to improve results. In order to estimate the PSD of the signals, Standard periodogram, Welch periodogram, and Burgs method were used. For each bands, two parameters (power of the signal and the RMS) were computed over every 5 s segment. The classification was conducted by a Linear Discriminate Analysis (LDA) implemented in Matlab. The classification was computed in two ways, in one, all the band parameters were utilised, and in the other, a previous selection of parameters was performed. This selection was based in the minimisa-

tion of the Wilks Lambda Function (WLF) with values F of entry=3.84 and removal=2.71. It was concluded that the best performances were obtained based on the RMS feature for most subjects with any spectral estimation technique. Among the spectral techniques, the Welch periodogram and Burg method were preferable than the standard periodogram with higher classification accuracy (about %82).

8.2   Time-Scale Methods

The Wavelet Transform (WT) is a transform that provides the TS representation of a given signal. This transform was developed to overcome some resolution related problems of the time-frequency transforms like short-time Fourier transform (STFT). The WT analysis uses a set of functions to represent the signal. A TS or WT representation of the signal can be obtained using digital filter techniques. In this case, filters of different cut-off frequencies are applied to analyse a signal at different scales (1/frequencies)[5]. This procedure offers a good time resolution at high frequencies and a good frequency resolution at low frequencies [68].

Due to the following three significant features of TS, it is a suitable tool for analysing the EEG signals [69]:

1. **multi-resolution** (to overcome the problem of time and frequency resolution),

2. **constant relative bandwidth** (the time widths of the wavelets is adapted to their frequencies), and

3. abilities to present signal whether is **localised in time-domain or frequency-domain**.

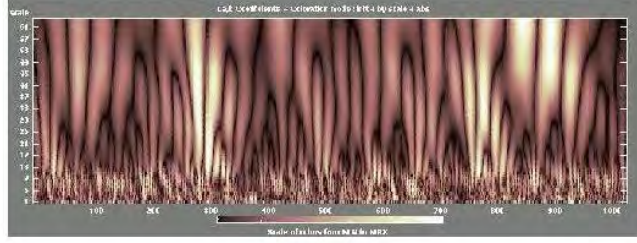[5]The scale concept is related to inverse of frequency in TS.

Figure 9: A WT plot of an EEG signal [70].

For illustration, a TS plot of an EEG signal is shown in Figure 9 displaying non-stationary nature of the signal. As seen, the scale representation of the signal varies in time.

In [69], using WT, an analytic study was done on adult EEGs to discriminate epileptic-form signals from non-epileptic ones, in 31 patients. The obtained results showed that the features of the EEG signal on each scale reflected the states of the signal. In this method, some regular rhythms of the EEG signals (θ, α, and δ rhythms) were extracted. The α rhythm was extracted on scales 2^1, 2^2, and 2^3, θ on scales 2^2, 2^3, and 2^4, while δ on scales 2^3, 2^4, and 2^5. When the subject were asked to shut-open their eyes, the amplitude of the waveform on scales 2^1, 2^2, and 2^3 obviously changed. The features on epileptic-form discharge also appeared on scales 2^1 and 2^5 with higher amplitude. As a result, using these information, features on epileptic-form in the EEGs of epileptic patients can be discriminated well at different levels.

Mental workload was evaluated using WT of the EEG signals in [71]. In this research, it was emphasised that EEG signals are non-stationary and change abruptly in a short period, which leads to the change of frequency characteristics. It is therefore impossible to detect such a

change of frequency characteristics by means of traditional approaches using power spectral analysis (by the FFT) of the EEG or the P300 component of ERP. In this case, the WT is promising because it changes the window size adaptively and uses short windows at high frequencies and long windows at low frequencies. Consequently, it is expected that the WT is applicable to non-stationary signals best. This adaptive window size, offers better resolution in time at higher frequencies (rapid changes in time) and better resolution in frequencies at low frequencies. Most suitable property for a non-stationary signal. To put this fact to the test and evaluate mental workload induced during human-computer interactions, 8 healthy participants were chosen to undertake a continuous matching task with three levels of task difficulty. For the matching task, they were asked to match the current stimulus with the stimulus of one trial ago or two trials ago in terms of identity and its spatial position (according to Gevins *et al.* task in [6] to compare the results). The EEG and also performance (reaction time and errors trials, if any) were recorded from Fz, Cz, and Pz positions, continuously. The EEG data for 4.5 s for each task at the three levels of difficulty were analysed using a WT (db with octave divisions and the final order set to 3 and 12) to classify the freq bands into θ, α, and β. At each difficulty level, the magnitude of the power extracted from the frequency band of $2 - 32$ Hz depicted that the time when a maximum power appears (appearance time) was delayed with increased task difficulty (the higher the difficulty, the later the maximum power appeared).

The proposed method in this study, can differentiate cognitive task difficulty definitely, and in a shorter period, based on the appearance time and total power obtained. With regard to performance factors, reaction time (RT) increased as task difficulty increased. The percentage correct decreased with increased task difficulty.

In [72] paper, filtering and classification techniques of EEG for BCI applications was compared. These techniques were used for feature extraction and classification of mu features from EEG readings of subjects engaged in motor tasks. The mu rhythm is an arch-shaped oscillation that is strongest in the $8 - 13$ Hz range (α), but is also present in β and γ bands. Mu rhythm is attenuated by motor activity, a phenomenon known as ERD. The analysed EEG data was collected during BCI therapy sessions from two healthy right-arm dominant subjects. In these sessions, the subjects were seated in front of a computer screen with their right hand gripping the manipulandum, or end-effector, of an IMT Inmotion2, shoulder-elbow robot. The EEG data was obtained using 58-channel recording, at a sampling rate of 250 Hz, and bandpass filtered from $0.1 - 40$ Hz. The data were preprocessed by using a common-average reference (CAR) spatial filter to reduce noise. Three methods of temporal filtering were then performed on de-noised data: AR filtering, mu-matched filtering, and wavelet decomposition. The extracted features here were mean and STD. The first classification technique used, was a simple linear one by BCI2000 in the online task. The second technique was SVM as a non-linear classifier. Most applied methods were statistically better than the baseline (AR feature extraction /simple linear classification). With regard to SVM classifier, the method with the highest mean classification rate was wavelet one (db25) with polynomial SVM classification, which has classification accuracy of %63, nearly %10 higher than the baseline. Then it was followed by combination of matched filtering and polynomial SVM, with accuracy of %61. Interestingly, a large improvement was achieved by using wavelets and the simple linear classifier with classification accuracy of nearly %80 (for db2, then db8, followed by db25). According to the achieved results, non-linear SVMs generally showed a

small advantage over linear SVMs. SVMs with all kernels statistically outperformed the simple linear classifier. Also, Daubechies wavelets were indeed a good method for extracting movement-related information from EEG signals.

9 Research Proposal

As aforementioned, this research mainly aims to develop measurement techniques for CL using EEG signal processing. The utilised apparatus is the Active Two recording system [73] for research applications which can be used in cognitive science and psychology research for studying mental activity and work load, comfortably. Some of the main objectives of this research at the beginning stage are as follows:

- To assess the discrimination capability between low, moderate, and high loads imposed to the working memory during different tasks,

- To gauge the mental effort, performance, and the time factor involved,

- To minimise working memory load to avoid performance errors,

- To decide the temporal resolution of EEG for changes in CL.

To start off with the practical part of the research, an experiment was designed and followed through, which is mentioned in the next subsection.

9.1 Materials

9.1.1 Participants

After granting the ethic approval to this research, many healthy volunteers from the UNSW students within a specific age group (24-30 years), engaged in post-graduate study were recruited for the first experiment. They were wearing the EEG head-cap which places electrodes on the scalp with a conductive gel or paste, comfortably. The participants were then instructed how to conduct the experiment.

9.1.2 Experiment

The experiment was defined as a silent reading task, displayed and controlled on a lap-top PC with a viewing distance of 70 cm to the participant. The reading task; was chosen to be semantically neutral and comprehension-independent to avoid any expertise effect and expected the level of all the participants reading to be relatively similar. The Lexile framework for reading [74] was used to rate the task readability/complexity and ensure it induces three different difficulty loads to the participants working memory. The text complexity ratings according to the Lexile measures were 1020, 1090, and 1150 for the low, medium and high levels respectively. The reading task was chosen to be silent to avoid any muscle movement artifact (EMG) resulting from speech production. The task was split into three levels; each level consisted of four pages displayed in a Power Point presentation. Each page lasted 30 seconds automatically, thus each reading task lasted 2 minutes, in each level. In the first level, defined as low difficulty, the participants were asked to read the displayed pages silently and pick up three letter words by pressing a mouse left button. In the second level,

defined as medium difficulty, they were required to read on and pick up three and four letter words by pressing the left and middle buttons, respectively. In the third level, defined as high difficulty, they picked up three, four, and five letter words from the remainder of the presentation, by pressing the left, middle, and right buttons, accordingly. To minimise any muscle movement artifact, the participant's hand and mouse were kept fixed in a certain position. Each load level task was conducted over two minute duration and the experiment was repeated two times per subject. In the baseline condition, conducted after the experiment, the subjects were asked to sit relaxed and keep their eyes open.

9.1.3   EEG Signal Recording

In this research, we applied the Active Two [73] EEG acquisition system, to record signals originating from the brain activities. It has a sensing technology that uses electrodes placed on the scalp to measure electrical potentials related to brain activity. Each electrode typically comprises of a wire leading to a conductive ring that is electrically connected to the scalp using conductive gel. The EEG device records the voltage at each of these electrodes relative to another electrode as a reference point. Since the EEG signals are of low amplitude (typically about 10 -100 μV for an adult), they should be amplified. The amplified signals are then converted from analog to digital format using the A/D box. The A/D box uses one 8 channel module, for operation with 32 EEG channels. All digital data is multiplexed into a serial data stream, and sent to the signal processing PC via an optical fiber. The experiment was conducted in an electrically isolated lab, with a minimum distance of 5 meters from power sources to the experiment desk, and
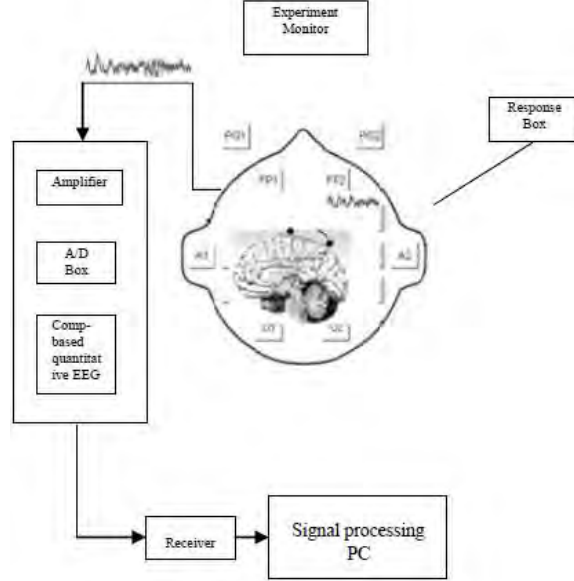
Figure 10: The experiment set up used in the study.

under low level of illumination. The EEG signals were recorded at the ATP Laboratory of National ICT Australia in Sydney. Each recording contained 32 EEG channels, according to the international 10 - 20 system. The data were recorded in digital form, at a 256 Hz sampling rate and a sensitivity of 100 μV . The EEG signals were passed through a band-pass filter with cut-off frequencies of 0.1 and 100 Hz. The used experiment configuration is shown in Figure 10.

9.2   Analysis Method

After collecting the data and making a database, the data analysis was conducted to assess the CL perceived during the experiment. The results achieved was then submitted to the Asia Pacific Signal Process-

51

ing Conference (APSIPA'10), and was accepted, recently. The analysis procedure can be summarised as follows:

1. *Windowing*: The EEG signals were first segmented for feature extraction using a rectangular window length of T=5 seconds. With the sampling frequency of 256 Hz, each EEG segment contains N=1280 samples. There is no overlap between the successive EEG segments.

2. *Feature calculation*: After examining different statistics of the EEG segments, the following features were found to be more suitable for discriminating between the different levels of loads: mean, root mean squared, and maximum cross-correlation between the given segment and a segment of the same size of the baseline EEG.

9.3   Results and Discussions

For each subject in each trial, 320 seconds of EEG recording containing 3 task levels and the baseline (with the length of 80 seconds each) were analysed. All the features explained above were extracted from the EEG segments of 5 seconds length. The performance of each feature for all the 32 EEG channels was examined and it was found that the two frontal EEG channels; namely: Fp1, and Fp2, were the most sensitive to the load on working memory across the participants. This result is supported by previous studies indicating that the brain frontal lobes often play a critical role in working memory related tasks associated with attention and mental efforts [71, 13]. For illustration purposes, the mean value extracted from the EEG signals representing 3 different task difficulties of all the five participants in one trial is shown in Figure 11. Note that each task was split into second duration windows

with no overlap between adjacent windows, giving 16 segments of 5 seconds. This shows that as the task difficulty increases, the mean of the EEG signal tends to increase, and therefore, each level of difficulty is clearly distinguishable from other levels by different mean. We also statistically analysed the extracted features for all five subjects and found the results to be almost the same for all subjects. Here, we present the results of the statistical analysis for subject 1 only. We also statistically analysed the extracted features for all five subjects and found the results to be almost the same for all subjects. Here, we present the results of the statistical analysis for subject 1 only. Fig. 12 plots the values of the three features extracted from the EEG signals acquired from the frontal channels during different load levels for subject 1, together with their 95% confidence intervals. The feature shows a consistent increasing trend as cognitive load is increased. The features feature and , on the other hand, show decreasing trends. The features extracted from the channel Fp1 show larger confidence intervals, indicating that the trends observed here are less reliable.

We also investigated the effect of the window length on the standard deviation of the extracted features and found that among all subjects, the features extracted from the EEG signals acquired from subject 1 were the most sensitive to the window length. The behaviour of the features for both frontal channels was almost the same. Figure 13 plots the standard deviation of the features , and as a function of the window length for different load levels. It shows that the standard deviation of the features 'mean' and 'RMS' for all the three load levels remains steady while the window length increases. The standard deviation of the feature 'correlation' on the other hand, slightly increases, which might be expected since a longer cross-correlation interval could produce a wider range of maximum cross-correlation values. This figure
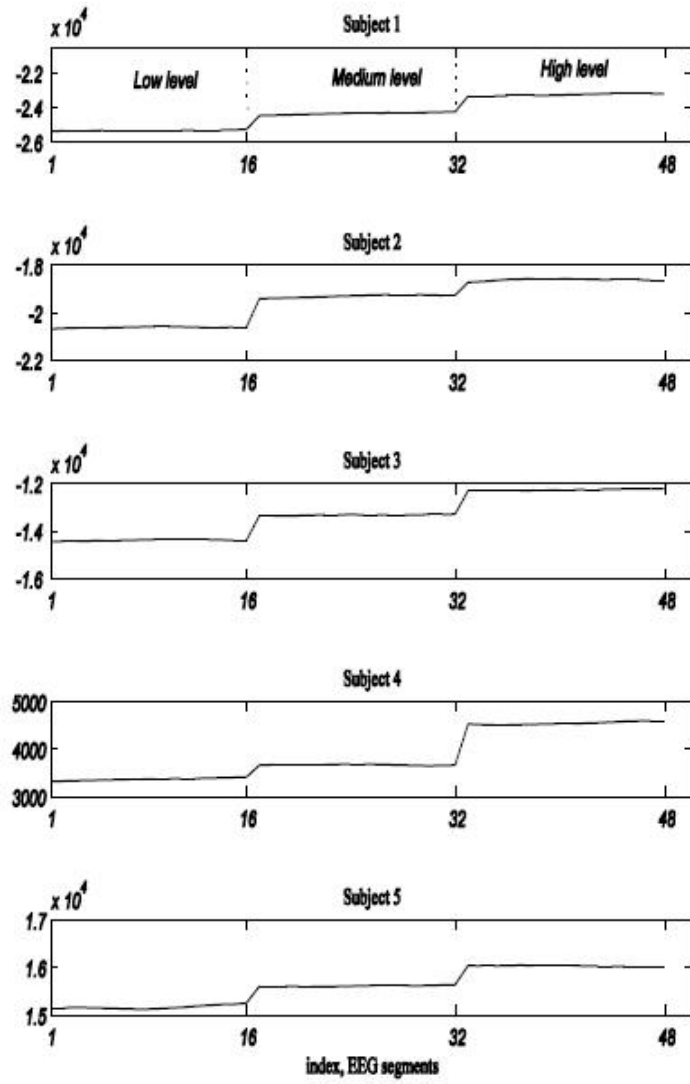
Figure 11: Values for the mean feature extracted from the segmented EEG data of five participants.
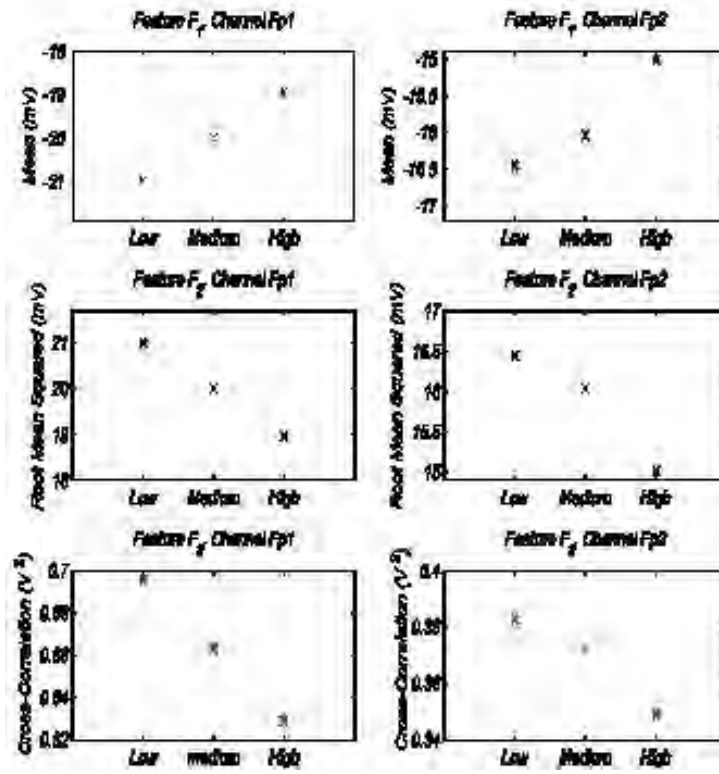
Figure 12: Extracted features from the EEG signals acquired from the frontal channels, during different cognitive load conditions for one subject.
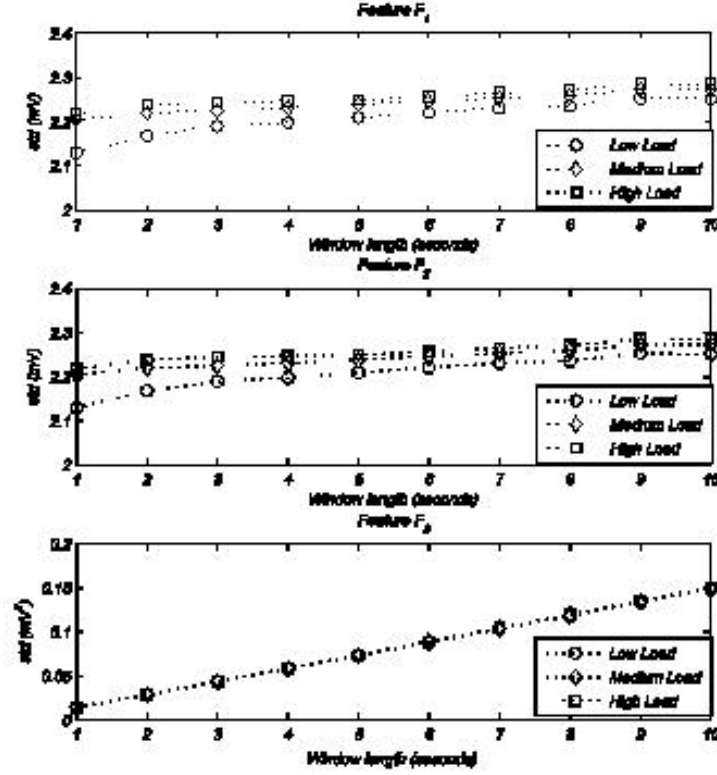
Figure 13: The standard deviation of the features extracted from the EEG signals acquired from channel Fp1 of Subject 1.

suggests that a window length of as short as second could be used for the purpose of cognitive load measurement.

Analysis of the effect of the window length used during feature extraction from the EEG signals suggest that features extracted from EEG segments as short as 1 second exhibit an acceptable amount of standard deviation, suggesting that EEG-based measurement of fairly rapid changes in cognitive load may be feasible.

56

9.4 Practical Problems

Despite the promising results showing the feasibility and usefulness of applying EEG for CL measurement, it was experienced that the EEG data recording is very sensitive to artefacts even to blinking, talking, and subtle body movements. The effect of each of the above could be misleading in the data analysis and interpretation. Therefore, the experiment should be conducted in highly controlled experimental settings and so many parameters should be under controls. Clearly, noise is also easily added into the EEG signal from electrical interference and even the participants' breathing and heartbeat.

9.5 Conclusion

Our first study has investigated the discrimination of mental workloads using statistical features calculated from frontal EEG signals. Discrimination between the different workloads induced was found to be highly significant, and the changes in feature values between load levels were consistent for mean features and mostly consistent for energy and correlation features. As the task difficulty increased, the mean for all participants tended to increase, indicating that the central nervous system is more highly activated when more cognitive load is imposed on a participant. These results support the feasibility and usefulness of employing EEG-based methods using less EEG channels for monitoring and evaluating cognitive load in real time.

This study results will appear in the proceedings of the Asia Pacific Signal Processing Conference (APSIPA'10).

There is extensive scope for future work, including validating the findings on more subjects, inducing a finer scale of task load levels to

test the cognitive load measurement precision, using evoked potential methodology to assess cognitive load, and exploring the use of other possible features. Although classification of cognitive load is certainly a future goal for this research, at present the focus is on establishing the precision and applicability for future measurement systems.

This research has just begun and there is extensive scope for future work. This may including validating the findings on more subjects, inducing a finer scale of task load levels to test the cognitive load measurement precision, using evoked potential methodology to assess cognitive load, and exploring the use of other possible features. Although classification of cognitive load is certainly a future goal for this research, at present the focus is on establishing the precision and applicability for future measurement systems.

10 Timetable for Completion

Task	2010				2011				2012			
Literature review and update	♦	♦	♦	◊	◊	◊	◊	◊	◊	◊	◊	◊
Coursework	♦	♦	♦									
Setting EEG data acquisition system up and calibration study			♦									
Preparation for confirmation of candidature				◊								
Designing and running experiments and making the database		♦	◊	◊	◊	◊	◊	◊				
Analysis of the data (and publishing results in APSIPA'2010)			♦	◊	◊	◊	◊	◊	◊	◊		
Finding spatial distribution of CL					◊	◊						
Modelling the effect of CL on EEG signals							◊	◊				
Finding correlation across CL measurement methods									◊	◊		
Publications of research results								◊	◊	◊	◊	
Thesis writing											◊	◊

References

[1] F. Paas, J. E. Tuovinen, H. Tabbers, and P. W. M. Van Gerven. Cognitive load measurement as means to advance cognitive load theory. In *Educational Psychologist*, volume 38, pages 63–71, 2003.

[2] F. Paas and J. G. Van Merrinboer. Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. In *Journal of Educational Psychology*, volume 86, pages 122–133, 1994.

[3] B. Yin, F. Chen, N. Ruiz, and E. Ambikairajah. Speech-based cognitive load monitoring system. In *ICASSP'08*, pages 5234–5237, 2008.

[4] B. Yin, N. Ruiz, F. Chen, and E. Ambikairajah. Automatic cognitive load detection from speech features. In *OzCHI'2007*, pages 249–255, 2007.

[5] P. Mermelstein. Distance measures for speech recognition, psychological and instrumental. In *Pattern Recognition and Artificial Intelligence*, pages 374–388. New York: Academic, 1976.

[6] A. Gevins, M. E. Smith, H. Leong, L. Mcevoy, S. Whitfield, R. Du, and G. Rush. Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. In *Human Factors, ProQuest Health and Medical*, volume 40, page 79, March 1998.

[7] J. Lamberts, V.D. Broek, P.L. Bener, L. Egmond, J. Dirksen, and A.M. Coenen. Correlation dimension of the human EEG corresponds with cognitive load. In *Neuropsychology*, volume 41, pages 149–153, 2000.

[8] R. W. Backs and L.C. Walrath. Eye movement and pupillary response indices of mental workload during visual search for symbolic displays. In *Applied Ergonomics*, volume 23, pages 243–254, 1996.

[9] Sh. T. Iqbal, X. S. Zheng, and B. P. Bailey. Task-evoked pupillary response to mental workload in human-computer interaction. In *proceedings of CHI'2004*, pages 1477–1480, 2004.

[10] Y. Shi, N. Zheng, R. Taib, E. H. C. Choi, and F. Chen. Galvanic skin response GSR as an index of cognitive load. In *proceedings of CHI'2007*, pages 2651–2655, 2007.

[11] P. Antonenko, F. Paas, R. Grabner, and T. Van Gog. Using electroencephalography to measure cognitive load. In *Published online, DOI 10.1007/s10648-010-9130-y*. Springer Science, 29 April 2010.

[12] N. V. Thakor and S. Tong. Advances in quantitative electroencephalogram analysis methods. In *Annual Review of Biomedical Engineering*, volume 6, August 2004.

[13] A. Gevins, M. E. Smith, L. Mcevoy, and D. Yu. High resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty. In *Cerebral Cortex*, volume 7, pages 374–385, 1997.

[14] W. Klimesch, M. Doppelmayr, T. Pachinger, and B. Ripper. Event-related desynchronisation in the alpha band and the processing of semantic information. In *Cognitive Brain Res.*, volume 6, pages 83–94, 1997.

[15] J. Polich. Updating P300: An integrative theory of P3a and P3b. In *Clin Neurophysiol*, volume 118, pages 2128–2148, 2007.

[16] M.E. Smith, A. Gevins, H. Brown, A. Karnik, and R. Du. Monitoring task loading with multivariate EEG measures during com-

plex forms of human-computer interaction. In *Human Factors*, volume 43, 2001.

[17] J. Sweller. Cognitive load during problem solving: Effects on learning. In *Cognitive Science*, volume 12, pages 257–285, 1988.

[18] J. Sweller, J. Van Merrinboer, and F. Paas. Cognitive architecture and instructional design. In *Educational Psychology*, volume 10, pages 251–296, 1998.

[19] P. Antonenko, F. Paas, R. Grabner, and T. Van Gog. Using electroencephalography to measure cognitive load. In *Springer Science, Published online*. DOI 10.1007/s10648-010-9130-y, 29 April 2010.

[20] F. Paas and J. J. G. Van Merrinboer. The efficiency of instructional conditions: An approach to combine mental-effort and performance measures. In *Human Factors*, volume 35, pages 737–743, 1993.

[21] G. Cooper. Research into cognitive load theory and instructional design at UNSW., 1998. http://www.arts.unsw.edu.au/education/CLT_NET_Aug_97.HTML.

[22] R. Mayer, J. Heiser, and S. Lonn. Cognitive constraints on multimedia learning: When presenting more material results in less understanding. In *Educational Psychology*, volume 93, pages 187–198, 2001. San Diego, CA: Pfeiffer.

[23] J. Sweller. Cognitive load theory, learning difficulty and instructional design. In *Learning and Instruction*, volume 4, pages 295–312, 1994. San Diego, CA: Pfeiffer.

[24] R. C. Clark, F. Nguyen, and J. Sweller. Efficiency in learning: Evidence-based guidelines to manage cognitive load. In *Technical Report*, volume 7, 2005. San Diego, CA: Pfeiffer.

[25] F. Paas, A. Renkl, and J. Sweller. Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. In *Instructional Science*, volume volume 32, pages 1–8, 2004.

[26] C. Berka, D. J. Levendowski, M. M. Cvetinovic, and et al. Real-time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset. In *Int'. Journal of Human Computer Interaction*, volume 17. Lawrence Erlbaum Associates, Inc., 2004.

[27] E. Niedermeyer and L. Da Silva. In *Electroencephalography, 4th Ed.*, page 1258. Williams and Wilkins, Baltimore, 1999.

[28] B. Abou-Khalil and K.E. Musilus. Atlas of EEG and seizure semiology. Elsevier, 2006.

[29] M. Teplan. Fundamentals of EEG measurements. volume 79. Measmt Science, 2002.

[30] S. Sanei and J. A. Chambers. EEG signal processing. volume 79, page 127. John Wiley, 2007.

[31] J. M Stern and J. Engel. Atlas of EEG patterns. pages 27–55. Williams and Wilkins, 2004.

[32] M. A. Roessgen. *Analysis and modelling of EEG data with application to seizure detection in the newborn.* PhD thesis, SPRC, Oct. 1997.

[33] H. Aurlien, I.O. Gjerde, J. H. Aarseth, B. Karlsen, H. Skeidsvoll, and N. E. Gilhus. EEG background activity described by a large computerised database. In *Clinical Neurophysiology*, volume 115, pages 665–673, 2004.

[34] PL. Nunez and KL. Pilgreen. The spline-laplacian in clinical neurophysiology: a method to improve EEG- spatial resolution. volume 8, pages 397–413, Oct 1991.

[35] R. Cooper, J. W. Osselton, and J. C. Shaw. *EEG Technology.*, chapter 6, pages 123–129. Butterworth, 1980.

[36] A. Isaksson and A. Wennberg. Computer analysis of EEG signals with parametric models. In *Proceedings of the IEEE*, volume 69, pages 451–461, 1981.

[37] E. Niedermeyer and F. Lopes da Silva. In *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields.* Lippincot Williams and Wilkins, 2004.

[38] T.H. Turner, S.P.A. Drummond, J.S. Salamat, and G.G. Brown. Effects of 42 hr sleep deprivation on component processes of verbal working memory. In *Neuropsychology*, volume 21, pages 787–795, 2007.

[39] J. R. G. CARRIE. A technique for analysing transient EEG abnormality. In *Electroce. and Clinical Neurophysiology*, pages 199–201. Elsevier, 1971.

[40] M. G. H. Coles and M. D. Rugg. *Event-related brain potentials: An introduction.*, pages 1–27. Electrophysiology of Mind, Oxford Scholarship Online Monographs, 1996.

[41] P. Barklem, B. Boashash, P. Coldtiz, and S. Saliu. Analysis and modelling of EEG signals for detection of seizure in newborns. Technical report, SPRC and Prenatal Research Centre (PRC), Nov. 1998.

[42] K. M. Misulis and T. Fakhoury. Spehlmann's evoked potential primer. In *Butterworth-heinemann*, 2001.

[43] http://en.wikipedia.org/wiki/File:ComponentsofERP.

[44] G. Zouridakis, P.G. Simos, and Papanicolaou A.C. Multiple bilaterally asymmetric cortical sources account for the auditory N1m component. In *Brain Topography*, volume 10, pages 183–190, 2007.

[45] E. Donchin and M. G. Coles. Is the p300 component a manifestation of context updating? In *Behavioural and Brain Sciences.*, volume 11, pages 357–427, 1988.

[46] S. Murali and V. V. Kulish. Modeling of evoked potentials of electroencephalograms: An overview. In *Digital Signal Processing, Elsevier*, volume 17, pages 665–674, 2007.

[47] A. V. Kramarenko and U. Tan. Validity of spectral analysis of evoked potentials in brain research. In *Int'. Jour. Neuroscience*, volume 112, pages 489–499, 2002.

[48] L. Avitan, M. Teicher, and M. Abeles. Eeg generatora model of potentials in a volume conductor. In *Journal of Neurophysiology*, number 102, pages 3046–3059, 2009.

[49] ED Adrian and G. Moruzzi. Impulses in the pyramidal tract. In *Journal of Physiology*, number 97, pages 153–199, 1939.

[50] A. Marsan. Electrical activity of the brain: slow waves and neuronal activity. In *Journal of Medical Science*, number 104, page 117, 1965.

[51] E.D. Adrian. The spread of activity in the cerebral cortex from peripheral sense organs. In *Journal of Physiology*, number 88, pages 127–161, 1936.

[52] H.T. Chang. Dendritic potentials of cortical neurons produced by direct electrical stimulation of the cerebral cortex. In *Journal of Neurophysiology*, number 14, pages 1–22, 1951.

[53] O. Creutzfeldt and J.J. Meisch. Changes in cortical neuron activity and EEG during hypoglycemia. In *Electroencephalogr Clin Neurophysiol*, number 24, pages 158–172, 1963.

[54] R. Elul. Dipoles of spontaneous activity in the cerebral cortex. In *Expert Neurology*, number 6, pages 285–299, 1962.

[55] P.L. Nunez and R. Srinivasan. Electric fields of the brain: The neurophysics of EEG. In *Oxford University Press*, 1981.

[56] O.D. Creutzfeldt, S. Watanabe, and H.D. Lux. Relations between EEG phenomena and potentials of single cortical cells. In *Electroencephalogr Clin Neurophysiol*, number 20, pages 1–18, 1966.

[57] http://en.wikipedia.org/wiki/Chemical$_synapse$.

[58] S. Klein and B.M. Thorne. Biological psychology. In *New York Worth Press*, 2007.

[59] W. Klimesch. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. In *Brain Research Reviews*, volume 29, pages 169–195. Elsevier, 1999.

[60] W. Klimesch, H. Schimke, and J. Schwaiger. Episodic and semantic memory: an analysis in the EEG-theta and alpha band. In *Electroencephalogr. Clin. Neurophysiol*, volume 91, pages 428–441, 1994.

[61] G. Pfurtschellera and F.H. Lopes da Silva. Event-related EEG/MEG synchronisation and desynchronisation: basic principles. In *Clinical Neurophysiology*, volume 110, 1999.

[62] B. Boashash, M. Mesbah, and P. Colditz. Newborn EEG seizure pattern characterisation using time-frequency analysis. In *ICASSP'2001*, page on CDROM, 2001.

[63] B. Boashash, H. Carson, and M. Mesbah. Detection of seizures in newborns using time-frequency analysis of EEG signals. *The 10th IEEE workshop on*

Statistical Signal and Array Processing (SSAP), Pocono Manor, Pennsylvania, USA, pages 564–568, Aug. 2000.

[64] A. J. Olievera, M. L. Nunes, and J. C. Costa. Polysomnography in neonatal seizures. *Clinical Neurophsiology*, 111(2), 2000.

[65] A. Gupta, SH. Parameswaran, and C.H. Lee. Classification of EEG signals for different mental activities using KL divergence. In *ICASSP 2009*, pages 1697–1700, 2009.

[66] D. Erdogmus, A. Adami, M. Pavel, T. Lan, S. Mathan, S. Whitlow, and M. Dorneich. Cognitive state estimation based on EEG for augmented cognition. In *The 2nd Int. IEEE EMBS Conference on Neural Eng.*, pages 4–7, 2005.

[67] P.F. Diez, E. Laciar, V. Mut, E. Avila, and A. Torres. A comparative study of the performance of different spectral estimation methods for classification of mental tasks. In *the 30th Int' IEEE EMBS Conference*, pages 1155–1158, 2008.

[68] R. Polikar. The wavelet tutorial. *The engineering ultimate guide to wavelet analysis*, Jul. 2000.

[69] L. Yong and Z. Shengxun. Apply wavelet transform to analyse EEG signals. *In the Proc. of the 18th EMBS Conference*, pages 1007–1008, Jan. 1996.

[70] P. Zarjam. *EEG data acquisition and automatic seizure detection using WT in the newborns.* PhD thesis, Queensland University of Technology, 2004.

[71] A. Murata. An attempt to evaluate mental workload using wavelet transform of EEG. In *Human Factors*, volume 47, pages 498–508, 2005.

[72] M. Renfrew, R. Cheng, J.J. Daly, and M.C. Cavusogl. Comparison of filtering and classification techniques of EEG for brain-computer interface. In *the 30th Int' IEEE EMBS Conference*, pages 2634–2637, 2008.

[73] http://www.biosemi.com.

[74] The Lexile Framework For Reading (www.lexile.com).

Evaluation of Working Memory Load using EEG Signals

Pega Zarjam[1,2], Julien Epps[1,2], and Fang Chen[2]

[1]School of EE&T, the University of New South Wales, Sydney, NSW 2052, Australia

[2]ATP Research Lab, National ICT Australia, Eveleigh, NSW 2015, Australia

p.zarjam@student.unsw.edu.au, j.epps@unsw.edu.au, Fang.Chen@nicta.com.au

Abstract— **This paper investigates mental workload assessment using statistical features derived from electroencephalography (EEG) signals. Mean, root mean squared, and correlation-based features are extracted from data including EEG signal recordings of five participants performing a reading task with three difficulty levels of low, medium, and high and a baseline condition. Results reveal that for the given task, features derived from the EEG signals consistently exhibit a very high degree of discrimination between the induced load levels, confirming EEG as an important method for the real time, objective determination of cognitive load level. Also, the frontal EEG channels appear to be sensitive to the working memory load than the other channels. Analysis of the effect of the window length used during feature extraction from the EEG signals suggest that features extracted from EEG segments as short as 1 second exhibit an acceptable amount of standard deviation, suggesting that EEG-based measurement of fairly rapid changes in cognitive load may be feasible.**

I. INTRODUCTION

Cognitive load is referred to the load on working memory when cognitive capacity accommodates the demands imposed by the task. In recent years, there has been an increased focus on maintaining the right level of cognitive load applied on the brain. Since, it is quite well known that the brain has limited cognitive processing in both capacity and duration [1], any underloading or overloading of cognitive processing will result in undesirable consequences such as degrading performance or failures of learning and performing complex cognitive tasks. These must be avoided especially in critical decision making fields such as air traffic control or military operations, where the load requirements may be heavy and there is no room for error. Therefore, the working memory architecture and its limitations should be a major consideration when applying the task on cognitive processing [1]. As a result, there a clear need to monitor and measure cognitive load experienced.

Previously, researchers have used analytical and empirical methods to measure cognitive load [1, 2]. Analytical methods use techniques rely on expert opinion, mathematical models, and task analysis, and are task-specific or conducted post-hoc. Empirical methods gather behavioural [3, 4, 5] or physiological signals [2, 6]. Physiological techniques are justified by the assumption that changes in cognitive functioning are reflected by physiological variables (e.g., heart rate variability [2], task-evoked brain activities [6, 7],

and skin conductance [8]). Changes to behavioural signals (e.g. speech [3, 4], eye movement [9, 10] or pen input [11, 12]) may be effected either through the changes in physiology or in different mental processing strategies. These measures can best be used to visualise the detailed trend and pattern of cognitive load such as instantaneous, peak, and accumulated load [13].

Since, the brain is the seat of cognitive activity, with appropriate measurement tools such as electroencephalography (EEG); it should be possible to monitor the brain states. EEG is the recording of electrical activity along the scalp produced by the firing of neurons within the brain [14]. It has been successfully applied in monitoring and measuring different types of mental activities and workloads in cognitive science (e.g. relaxation, attention, and fatigue) and psychology (e.g. sleep, arousal, and anesthesia) [13].

EEG has also been shown to be an effective non-intrusive method of monitoring memory load, and highly sensitive to different cognitive states [6, 15, 16]. Notwithstanding the plethora of EEG-based systems for measuring various mental states and pathologies proposed to date, measurement of cognitive load using EEG signals is still in its early stages, and there is still a need to validate the applicability of the approach across different tasks and to establish the precision of cognitive load measurement.

The present study was designed to discriminate different cognitive states associated with a varying task difficulty from EEG signals. Therefore, an experiment comprising a baseline condition and a cognitive task condition with three different levels of difficulty was created. Features capable of discriminating working memory load/cognitive load levels were empirically compared across different participants. It was found that the statistics of the EEG signals recorded from the frontal channels reflect the level of working load well. Further, an investigation of the temporal precision, in terms of the feature extraction window length, was undertaken, revealing that EEG is a promising real time tool for measuring fairly rapid changes in cognitive load.

The structure of the paper is as follows: in Section II, the materials of this study are presented. In Section III, the analysis method and obtained results are presented and discussed. The effect of the window length used in the analysis of the EEG signals is investigated in Section IV.

Finally, the paper is concluded in Section V.

II. Materials

A. Participants

Five healthy male volunteers, 24-30 years of age, engaged in post-graduate study participated in the experiment.

B. Experiment

The experiment was defined as a silent reading task, displayed and controlled on a lap-top PC with a viewing distance of 70 cm to the participant. The reading task; was chosen to be semantically neutral and comprehension-independent to avoid any expertise effect and expected the level of all the participants reading to be relatively similar.

The Lexile framework for reading [17] was used to rate the task readability/complexity and ensure it induces three different difficulty loads to the participants working memory. The text complexity ratings according to the Lexile measures were 1020, 1090, and 1150 for the low, medium and high levels respectively. The reading task was chosen to be silent to avoid any muscle movement artifact (EMG) resulting from speech production.

The task was split into three levels; each level consisted of four pages displayed in a Power Point presentation. Each page lasted 30 seconds automatically, thus each reading task lasted 2 minutes, in each level. In the first level, defined as low difficulty, the participants were asked to read the displayed pages silently and pick up three letter words by pressing a mouse left button. In the second level, defined as medium difficulty, they were required to read on and pick up three and four letter words by pressing the left and middle buttons, respectively. In the third level, defined as high difficulty, they picked up three, four, and five letter words from the remainder of the presentation, by pressing the left, middle, and right buttons, accordingly. To minimise any muscle movement artifact, the participant's hand and mouse were kept fixed in a certain position.

Each load level task was conducted over two minute duration and the experiment was repeated two times per subject. In the baseline condition, conducted after the experiment, the subjects were asked to sit relaxed and keep their eyes open.

C. EEG Recording

The experiment was conducted in an electrically isolated lab, with a minimum distance of 5 meters from power sources to the experiment desk, and under low level of illumination.

The EEG signals were recorded by means of a BioSemi Active Two system [18], at the ATP Laboratory of National ICT Australia in Sydney. Each recording contained 32 EEG channels, according to the international 10 - 20 system. The data were recorded in digital form, at a 256 Hz sampling rate and a sensitivity of 100 μV. The EEG signals were passed through a low-pass filter with a cut-off frequency of 100 Hz.

III. Analysis And Results

A. Analysis Method

The analysis procedure can be summarised as follows:

- **Windowing:** The EEG signals were first segmented for feature extraction using a rectangular window length of $T = 5$ seconds. The ith segment is represented by $x[n]$, $N = 0,1,...,N-1$ where $N = T \times f_s$ with $f_s = 256$ Hz being the sampling frequency of the EEG data, i.e. each EEG segment contains $N = 1280$ samples. There is no overlap between the successive EEG segments.

- **Feature calculation:** After examining different statistics of the EEG segments, the following features were found to be more suitable for discriminating between the different levels of loads: mean [7], root mean squared, and maximum cross-correlation between the given segment and a segment of the same size of the baseline EEG, i.e. F_1, F_2, and F_3, respectively. The features are given by the following equations, where $x[n]$ and $y[n]$ are raw time-domain EEG signals.

$$F_1 = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \tag{1}$$

$$F_2 = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} x^2[n]} \tag{2}$$

$$F_3 = \text{Max}\big(R(m-N)\big); m = 1,2,...,2N-1 \tag{3}$$

where

$$R(m) = \begin{cases} \sum_{n=0}^{N-m-1} x[n+m]y[n] & ; m \geq 0 \\ R(-m) & ; m < 0 \end{cases} \tag{4}$$

The recorded baseline was used as a reference for calculating feature F_3 in equation (3).

B. Results and Discussions

For each subject in each trial, 320 seconds of EEG recording containing 3 task levels and the baseline (with the length of 80 seconds each[1]) were analysed.

---

[1] Since, the EEG recording was done continuously, to ensure the transition between different load levels has passed, 20 seconds from the beginning and end of the signals were cut off.
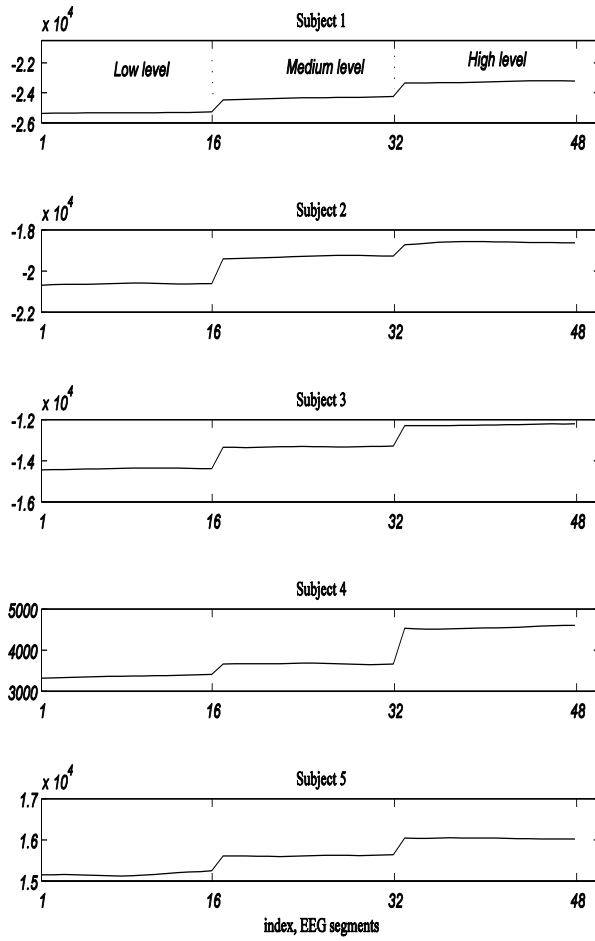
Fig. 1. Values for the mean feature (F_1) extracted from the segmented EEG data for five participants. The first, second, and third 16 segments belong to the EEG signals recorded during low, medium, and high levels of task difficulty, respectively.

All the features explained above were extracted from the EEG segments of 5 seconds length. The performance of each feature for all the 32 EEG channels was examined and it was found that the two frontal EEG channels; namely: Fp1, and Fp2, were the most sensitive to the load on working memory across the participants. This result is supported by previous studies indicating that the brain frontal lobes often play a critical role in working memory related tasks associated with attention and mental efforts [16, 19].

For illustration purposes, the mean value (F_1) extracted from the EEG signals representing 3 different task difficulties of all the five participants in one trial is shown in Fig. 1. Note that each task was split into $T = 5$ second duration windows with no overlap between adjacent windows, giving $\frac{80}{5} = 16$ segments of 5 seconds.

Fig. 1 shows that as the task difficulty increases, the mean of the EEG signal tends to increase, and therefore, each level

of difficulty is clearly distinguishable from other levels by different mean.

We also statistically analysed the extracted features for all five subjects and found the results to be almost the same for all subjects. Here, we present the results of the statistical analysis for subject 1 only.

Fig. 2 plots the values of the three features extracted from the EEG signals acquired from the frontal channels during different load levels for subject 1, together with their 95% confidence intervals. The feature F_1 shows a consistent increasing trend as cognitive load is increased. The features feature F_2 and F_3, on the other hand, show decreasing trends. The features extracted from the channel Fp1 show larger confidence intervals, indicating that the trends observed here are less reliable.

For each of the frontal channels, Table 1 shows the average and standard deviation of all the 3 features during each task level over 2 trials. The results show that the features yield distinctive values depending on the task difficulty. Specifically, it is observed that as the task difficulty increases, the mean value increases for all subjects and the root mean squared and the correlation value decreases for most subjects. The small value of standard deviations in all cases shows that values of the features during each level of task difficulty are very close to the average value.
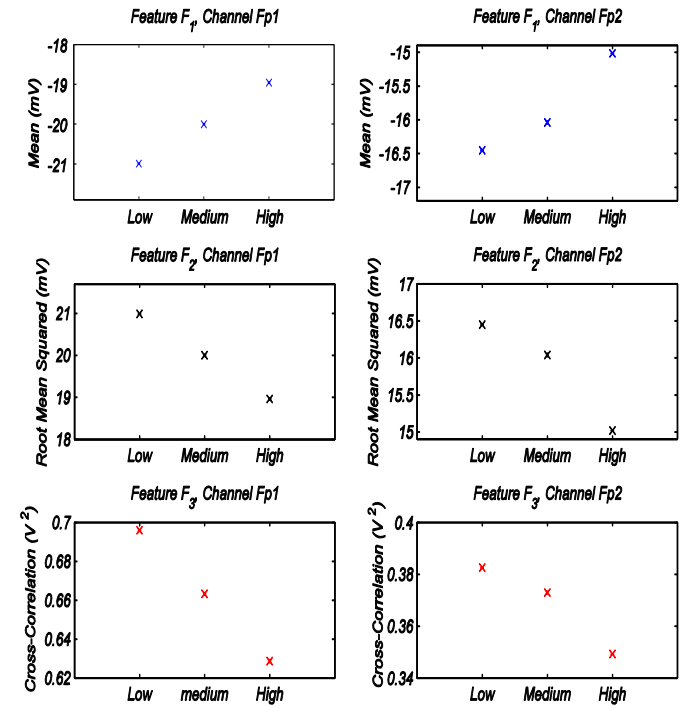

Fig. 2 Extracted features from the EEG signals acquired from the frontal channels, Fp1 and Fp2, during different cognitive load conditions for the first subject.
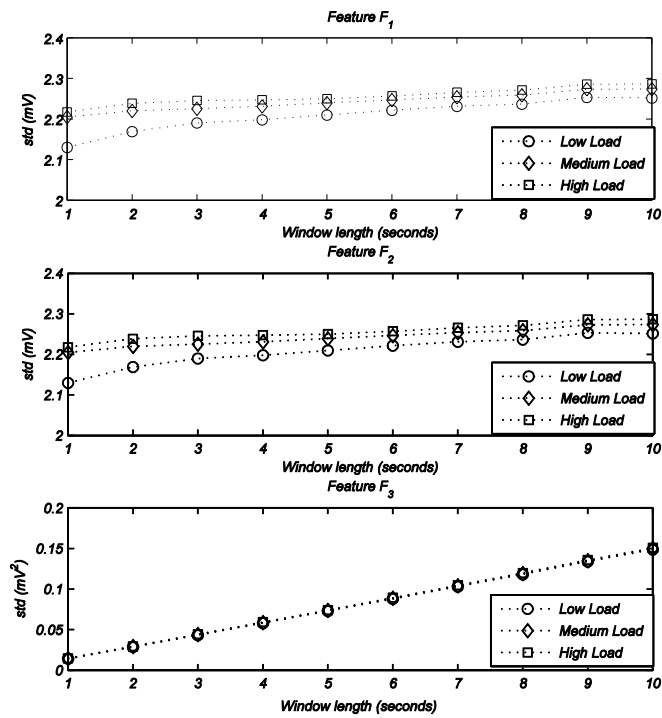
Fig. 3. The standard deviation of the features extracted from the EEG signals acquired from channel Fp1 of Subject 1.

IV. EFFECT OF THE WINDOW LENGTH

As mentioned earlier, the EEG signals during each task were split into T second long windows with no overlap between adjacent windows. We investigated the effect of the window length T on the standard deviation of the extracted features and found that among all subjects, the features extracted from the EEG signals acquired from subject 1 were the most sensitive to the window length. The behaviour of the features for both frontal channels was almost the same.

Fig. 3 plots the standard deviation of the features F_1, F_2 and F_3 as a function of the window length T for different load levels. It shows that the standard deviation of the features F_1 and F_2 for all the three load levels remains steady while the window length increases. The standard deviation of the feature F_3 on the other hand, slightly increases, which might be expected since a longer cross-correlation interval could produce a wider range of maximum cross-correlation values. Fig. 3 suggests that a window length of as short as $T = 1$ second could be used for the purpose of cognitive load measurement.

V. CONCLUSIONS

This paper has investigated the discrimination of mental workloads using mean, root mean-squared and correlation based features calculated from frontal EEG signals. Discrimination between the different workloads induced was

found to be highly significant, and the changes in feature values between load levels were consistent for mean features and mostly consistent for root mean squared and correlation features. As the task difficulty increased, the mean for all participants tended to increase, indicating that the central nervous system is more highly activated when more cognitive load is imposed on a participant. These results support the feasibility and usefulness of employing EEG-based methods for monitoring and evaluating cognitive load in real time.

There is extensive scope for future work, including validating the findings on more subjects, inducing a finer scale of task load levels to test the cognitive load measurement precision, using evoked potential methodology to assess cognitive load, and exploring the use of other possible features. Although classification of cognitive load is certainly a future goal for this research, at present the focus is on establishing the precision and applicability for future measurement systems.

REFERENCES

[1] F. Paas, J. E. Tuovinen, H. Tabbers, and P. W. M. Van Gerven, "Cognitive load measurement as means to advance cognitive load theory". In Educational Psychologist, vol. 38, no.1, pp. 63-71, March 2003.

[2] F. Paas and J. G. Van Merrinboer, "Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive load approach," In Journal of Educational Psychology, vol. 86, pp. 122-133, 1994.

[3] B. Yin, F. Chen, N. Ruiz, and E. Ambikairajah, "Speech-based cognitive load monitoring system." In proceedings of ICASSP'08, pp. 5234-5237, 2008.

[4] B. Yin, N. Ruiz, F. Chen, and M. A. Khawaja, "Automatic cognitive load detection from speech features." In proceedings of OzCHI 2007, pp. 249-255, 2007.

[5] P. Mermelstein, "Distance measures for speech recognition, psychological and instrumental." In Pattern Recognition and Artificial Intelligence, C. H. Chen, Ed. New York: Academic, 1976, pp. 374–388.

[6] A. Gevins, M. E. Smith, H. Leong, L. Mcevoy, S. Whitfield, R. Du, and G. Rush, "Monitoring working memory load during computer based tasks with EEG pattern recognition methods." In Human Factors, ProQuest Health and Medical, vol. 40, pp. 79, March 1998.

[7] J. Lamberts, V.D. Broek, P.L. Bener, L. Egmond, J. Dirksen, and A.M. Coenen, "Correlation dimension of the human EEG corresponds with cognitive load." In Neuropsychology, vol. 41, pp. 149-153, 2000.

[8] Y. Shi, N. Ruiz, R. Taib, E. H. C. Choi, and F. Chen, "Galvanic skin response (GSR) as an Index of cognitive load." In proceedings of CHI 2007, pp. 2651-2655, 2007.

[9] R. W. Backs and L.C. Walrath, "Eye movement and pupillary response indices of mental workload during visual search for symbolic displays." In Applied Ergonomics, vol. 23, no. 4, pp. 243-254, 1996.

[10] Sh. T. Iqbal, X. S. Zheng, and B. P. Bailey, "Task-evoked pupillary response to mental workload in human-computer interaction." In proceedings of CHI 2004, pp. 1477-1480, 2004.

[11] H. Ichikawa, M. Homma, M. Umemura" An experimental evaluation of input devices for pointing work." In Elsevier Science Int. Journal of Production Economics, vol. 60-61, pp. 235-240, 1999.

[12] N. Ruiz, R. Taib, Y. Shi, E. H. C. Choi, and F. Chen, "Using pen input features as indices of cognitive load." In proceedings of ICMI'07, pp. 315-318, 2007.

[13] N. V. Thakor and S. Tong, "Advances in quantitative Electroencephalogram analysis methods." Annual Review of Biomedical Engineering, vol. 6, pp. 453-495, August 2004.

[14] E. Niedermeyer and F. Lopes da Silva, "In Electroencephalography: Basic principles, clinical applications, and related fields." Williams and Wilkins press, 2004.

[15] A. Murata, "An attempt to evaluate mental workload using wavelet transforms of EEG." In Human Factors, vol. 47, no. 3, pp. 495-508, Fall 2005.

[16] A. Gevins, M. E. Smith, L. Mcevoy, and D. Yu, "High resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing ,and practice. " In Cerebral Cortex, vol. 7, pp. 374-385, 1997.

[17] The Lexile Framework For Reading (www.lexile.com)

[18] http://www.biosemi.com

[19] E.E. Smith and J. Jonides, "Working memory in humans: Neuropsychological evidence." In cognitive neurosciences, pp. 1009–1020. Cambridge, MA: MIT Press, 1995.

Channel	Subject	$\overline{F}_1, mV$ $std(F_1), mV$			$\overline{F}_2, mV$ $std(F_2), mV$			$\overline{F}_3, V^2$ $std(F_3), V^2$		
		Low	Medium	High	Low	Medium	High	Low	Medium	High
FP1	1	-23.16 2.21	-22.21 2.24	-21.18 2.25	23.16 2.21	22.21 2.24	21.18 2.25	0.7640 0.0729	0.7325 0.0739	0.6985 0.0742
	2	-19.01 1.65	-18.00 1.35	-17.46 1.20	19.01 1.65	18.00 1.35	17.46 1.20	0.3782 0.0327	0.3580 0.0269	0.3472 0.0239
	3	-12.84 1.58	-11.90 1.47	-10.89 1.45	12.84 1.58	11.90 1.47	10.89 1.45	0.1374 0.0169	0.1273 0.0157	0.1164 0.0155
	4	4.34 1.00	4.68 1.04	5.20 0.68	4.34 1.00	4.68 1.04	5.20 0.68	0.0394 0.0091	0.0424 0.0094	0.0471 0.0061
	5	14.50 0.70	15.68 0.21	15.99 0.04	14.50 0.70	15.68 0.21	15.99 0.04	0.3751 0.0180	0.4055 0.0053	0.4136 0.0009
FP2	1	-17.21 0.78	-16.72 0.70	-15.81 0.80	17.21 0.78	16.72 0.70	15.81 0.80	0.3958 0.0179	0.3845 0.0160	0.3634 0.0185
	2	-13.67 0.11	-13.73 0.21	-13.65 0.27	13.67 0.11	13.73 0.21	13.65 0.27	0.2293 0.0018	0.2302 0.0035	0.2289 0.0046
	3	-18.39 1.90	-17.34 2.03	-15.96 2.16	18.39 1.90	17.34 2.03	15.96 2.16	0.2842 0.0294	0.2680 0.0314	0.2464 0.0334
	4	6.29 0.69	6.42 0.68	6.81 0.40	6.29 0.69	6.42 0.68	6.81 0.40	0.0652 0.0071	0.0666 0.0070	0.0706 0.0041
	5	8.74 3.51	8.71 3.62	7.69 2.98	8.74 3.51	8.71 3.62	7.69 2.98	0.0896 0.0360	0.0893 0.0371	0.0789 0.0306

Table 1. Comparison of the averaged features among levels of task complexity, subjects, and frontal channels. The standard deviations of the features are also shown. For example, the averaged mean ($\overline{F}_1$) of the EEG signal recorded from the channel Fp1 for subject 1 during low level difficulty task is $-23.16\ mV$ and its standard deviation is $2.21\ mV$.

Spectral EEG Featuresfor Evaluating Cognitive Load

Pega Zarjam, Julien Epps, and Fang Chen

Abstract—This study was undertaken to investigate spectral features derived from EEG signals for measuring cognitive load. Measurements of this kind have important commercial and clinical applications for optimizing the performance of users working under high mental load conditions, or as cognitive tests. Based on EEG recordings for a reading task in which three different levels of cognitive load were induced, it is shown that a set of spectral features - the spectral entropy, weighted mean frequency and its bandwidth, and spectral edge frequency - are all able to discriminate the three load levels effectively. An interesting result is that spectral entropy, which reflects the distribution of spectral energy rather than its magnitude, provides very good discrimination between cognitive load levels. We also report those EEG channels for which statistical significance between load levels was achieved. The effect of frequency bands on the spectral features is also investigated here. The results indicate that the choice of optimal frequency band can be dependent on the spectral feature extracted.

I. INTRODUCTION

COGNITIVE load is the amount of task demand applied on working memory when performing a mental task [1]. It is quite well-known that working memory is limited in capacity and time when holding or processing information [1]. Therefore, when a task becomes more difficult, the accessibility of working memory reduces and cognitive load increases. This may lead to decreased performance or even failed task completion, which is undesirable in any circumstance, but particularly in critical decision-making fields, such as air traffic control, fire command and military operations or when designing adaptive human computer interfaces. Thus, measuring the cognitive load experienced is a critical need [1].

Different techniques are available for measuring cognitive load; among them electroencephalography (EEG)-based methods are the most suitable for continuous and on-line assessment of cognitive load at all levels [2]. This is due to the high sensitivity of EEG to variations between different cognitive states, and task difficulty in one hand, and the easy usage of the EEG device and being less costly and bulky on the other hand [2]. Specially, the usage of EEG has become more feasible for real-world applications recently with the availability of wireless EEG systems [3].

Finding features that are good discriminators of different workloads is an important key to successfully measure and classify the cognitive load levels. Previously, a range of spectral features have been deployed for this purpose using EEG signals, indicating that the spectral features discriminate cognitive load best. This includes the use of power spectral density (PSD), the power and maximum\log power spectra [4]-[8]. But to date, other spectral features that could provide more information on the varying cognitive load characteristics have not been investigated.

This study is a continuation of the authors' previous study in which other features were employed to characterize varying behavior of cognitive load, and which identified a few frontal EEG channels as the most effective channels in separation of the three induced loads [9].

In this study, we aim to examine the usefulness of some spectral features (i.e. spectral entropy, weighted mean frequency and its bandwidth, and spectral edge frequency), and to determine in more detail the frequency band in which the maximum discrimination among the three load levels is yielded. These spectral features have been employed in past EEG medical/ brain computer interface (BCI) analysis [10]-[12], but not in the cognitive load context. For the various spectral features, the related EEG channels for which reliable information may be extracted for an EEG-based cognitive load classification system is also investigated.

II. MATERIAL

A. Experiment

EEG signals were acquired from five healthy male participants; postgraduate students aged between 24-30 years. In the experiment, the participants were asked to read text silently, which was displayed and controlled on a laptop PC with a viewing distance of 70 cm to them. The reading task was chosen to be semantically neutral and comprehension-independent to avoid any expertise effect, and assumed the reading ability of all participants to be relatively similar. To rate the task readability/complexity and ensure it induced three different difficulty loads, the Lexile framework for reading [13] was used. The test determined the different difficulty loads as 1020 for low load, 1090 for medium load, and 1150 for high load.

The task was split into three levels; participants were asked to read the displayed pages silently and pick up three (low), three and four (medium) or three, four and five (high) letter words by pressing the mouse left/middle/right button. In the baseline condition, conducted after the experiment, the participants were asked to sit relaxed and keep their eyes open. To minimize any muscle movement artifact (EMG),

P. Zarjam is with the School of EE&T, University of New South Wales, Sydney, NSW 2052, Australia. She is also withATP Research Laboratory, National ICT Australia (phone: +61 2 9385 4803; fax: +61 2 9385 5993; email: p.zarjam@student.unsw.edu.au).

J. Epps is with the School of EE&T, University of New South Wales, Sydney, NSW 2052, Australia. He is also with ATP Research Laboratory, National ICT Australia(email: j.epps@unsw.edu.au).

F. Chen is with ATP Research Laboratory, National ICT Australia(email: fang.chen@nicta.com.au).

the participants were asked to sit still and their hand was placed fixed in a certain position, where they could still make finger movements for clicking the mouse buttons in response to the word stimuli. The participants were also required to refrain from blinking as much as possible during the recording to avoid ocular artifacts (EOG).

B. EEG Recording

The data used for this study consists of multi-channel EEG recordings obtained from the five consenting participants. The data were acquired using an Active Two system [14], at the ATP Laboratory of National ICT Australia in Sydney. The experiment was conducted under controlled conditions in an electrically isolated lab, with a minimum distance of 5 meters from power sources to the experiment desk, and under natural illumination. Each recording contained 32 EEG channels, according to the international 10 - 20 system. The data were recorded in a digital form, at a $f_s = 256$ Hz sampling rate. This dataset was also used in the authors' previous study [9].

III. METHOD

A. EEG Preprocessing and Segmentation

A notch filter of 50 Hz was initially applied to the raw EEG signals to remove the electrical mains contamination. The signals were then passed through a band-pass filter with a pass-band of 0.1-100 Hz. Visual inspection of the recorded signals showed very low ocular artifact occurrences. Therefore, no artifact removal was conducted here. The acquired EEG signals were first segmented using 5s non-overlapping rectangular windows. Each segment is denoted herein as $x[n]$ and contains $N = 1280$ samples.

B. Spectral Feature Extraction

Five spectral features were extracted from each segment of the EEG signals. They are presented as follows:

Spectral Entropy (SpEn): Spectral entropy is a measure of the distribution of normalized spectral energies, in this case within a frequency band. It is given by:

$$\text{SpEn} = -\frac{1}{\log N_f} \sum_f P_f(x) \log_e P_f(x) \tag{1}$$

where $P_f(x)$ is an estimate of the probability density function (PDF) of the EEG segment amplitude spectrum. The PDF is calculated by normalizing the PSD estimate with respect to the total spectral power in each sub-band. N_f is the number of frequency bins in the PSD estimate.

According to the above equation, the spectral entropy attains its peak when all the frequency bins contain the same power. One reason for investigating this feature is that it is essentially independent of the sub-band energy, which has been shown to perform well in previous work [9].

For illustration purposes, the PSDs of the EEG signals acquired from one participant, while the three cognitive
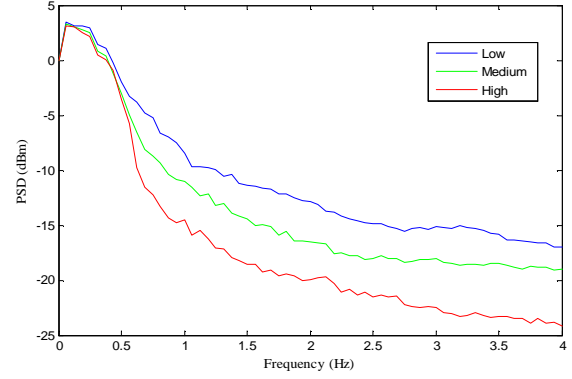

Fig.1. Extracted PSDs from the segmented EEG for participant 1 channel F8, recorded during induced high, medium, and low load levels.

loads induced are shown Fig. 1. As the cognitive load level increases, the signal power is concentrated in a smaller frequency band. Therefore, the spectral entropy decreases. The extracted entropy values in the delta sub-band shown in Fig. 1 for low, medium, and high loads are $0.6171*10^{-7}$, $0.5795*10^{-7}$, and $0.5343*10^{-7}$, respectively. This feature has been used for neonatal seizure detection previously [10].

Sub-band Energy (Enrg): The second spectral feature used here is the energy of the EEG segment in the delta sub-band (0-4 Hz) [9]. The choice of the sub-band was based on the fact that most of the energy of $x[n]$ resides below 4 Hz, which is also seen in Fig.1. In fact, energy is the integral of the signal spectral amplitudes, which is proportional to the signal power. It is an effective feature for EEG signal spatial classification in BCI applications [12].

Intensity Weighted Mean Frequency (IwMf): This feature measures a weighted mean of the frequencies present in the PSD estimate for each EEG segment [10]:

$$\text{IwMf} = \frac{\sum_{i=1}^{N_{f/2}-1} p_i \, idf}{\sum_{i=1}^{N_{f/2}-1} p_i} \tag{2}$$

where p_i is the estimated spectral power in frequency bin of i, f_s the sampling frequency, N_f the total number of frequency bins, and. $df = f_s / N_f$.

Intensity Weighted Bandwidth (IwBw): The associated bandwidth of the IwMf feature can be calculated by [10]:

$$\text{IwBw} = \sqrt{\frac{\sum_{i=1}^{N_{f/2}-1} p_i (\text{IwMf} - idf)^2}{\sum_{i=1}^{N_{f/2}-1} p_i}} \tag{3}$$

Spectral Edge Frequency (EdFr): This feature is defined as the frequency below which 90% of the signal power resides. This measure has been used previously for quantifying a pathological state (i.e. the depth of anesthesia in adults or white matter injury in neonates) [11].

C. Analysis

Since the features of interest are spectral, we initially investigated the spectral components of the recorded EEG signals. It appears that 90% of the energy of the spectral components resides in the 0-3.8 Hz region, computed by the extracted EdFr. This is also confirmed visually by the PSD shown Fig. 1. Clearly, the delta sub-band practically provides the most separation between the three load levels induced in the experiment. Therefore, the performance of all features namely; SpEn, Enrg, IwMf, IwBw, and EdFr were examined in the delta sub-band and compared for all the 32 EEG channels for all participants. We initially calculated the median of all the features for each EEG channel recorded, and then compared the effectiveness of each feature using a Kruskal-Wallis test.

IV. RESULTS AND DISCUSSION

A. Feature Comparison

TABLE I lists the EEG channels in which the features' medians calculated across all participants have shown consistent trends. It displays that the SpEn and Enrg exhibit a consistent decreasing trend as load level increases, in selected channels. However, the IwMf, IwBw, and EdFr exhibit an increasing trend as load level increases.

In order to examine how effective the extracted features are in the different load levels separation from the EEG channels, we used the Kruskal-Wallis test. The benefits of this test are: it examines more than 2 groups, is a non-parametric method, and is not affected by variations in a small portion of the data [15].The largest calculated *p*-values (indicating the worst case scenario) across the selected channels for each feature for all participants are displayed in TABLE II. As the *p*-values suggest, the Enrg feature shows a great statistical significance in differentiating the cognitive load levels in all the EEG channels. It is followed by the SpEn feature, with the second lowest set of *p*-values. As seen, most of the selected channels in TABLE I are confirmed in TABLE II, statistically.

TABLE I
Variations of the extracted features' medians from the EEG channels in the delta sub-band, indicating the channels that follow a consistent trend associated with the 3 load levels induced, across all participants.

Feature	EEG channels	Trend with increasing load
SpEn	Fp1, AF3, F7, CP5,P7,Pz, P8, CP6,CP2, T8, F4, F8, AF4	decreasing
Enrg	Fp1, AF3, F7, T7, CP3, P7, Pz, P8, CP6, CP2, T8, F4, F8, AF4	decreasing
IwMf	F7, FC5, T7, C3, P7, P3, Pz, P8, CP2, FC6, FC2, Fp2, Fz	increasing
IwBw	Fp1, AF3, F7, T7, C3, CP5, P3, Pz, PO3, P4, CP2, FC6, FC2, F4, Fp2, Fz	increasing
EdFr	Fp1, AF3, F7, P3, Pz, CP2, Fp2, Fz	increasing

TABLE II.
The EEG channels with a *p*-values<0.01 for allparticipants in the delta sub-band. The biggest calculated *p*-values display the worst case scenario.

Feature	EEG channels	Maximum *p*-value
SpEn	Fp1, AF3, T7, CP5,O1, P8, CP2,F8, Fp2	0.00357
Enrg	All 32 channels	5.0005372e-05
IwMf	P3, Pz, Oz, CP2, Fp2, Fz	0.0076619
IwBw	P3, Pz, CP2, F4, AF4, Fz	0.0091394
EdFr	P3, Pz, CP2, F4, AF4, Fp2, Fz	0.00847211

For illustration purposes, this significant statistical difference among the three load levels for the SpEn is also displayed in Fig. 2 for channel F4. As observed, the values of the SpEn feature are different for three load levels. Also, there is no overlap between the three box plots indicating complete separation among the three load levels.

B. Sub-divisions of the Delta Band

Following our investigation of the delta frequency sub-band; we divided this sub-band into finer frequency sub-bands. Therefore, the delta sub-band (0-4Hz) was split into three sub-bands of δ_0 (0-1 Hz), δ_1 (1-2 Hz), and δ_2 (2-4 Hz) using wavelet decomposition. The feature medians were recomputed for the δ_0, δ_1, and δ_2 sub-bands. It was observed that for the SpEn, and Enrg δ_0, and δ_1 sub-bands provided the same results as TABLE I, showing that the lower sub-bands seem to underlie most of the load level variations. For the IwMf, IwBw, and EdFr the medians in the finer sub-bands did not provide any significant results, indicating the load level information extracted by these features is distributed all over the delta sub-band. This suggests that the performances of the extracted spectral features are highly dependent on the frequencysub-divisions.
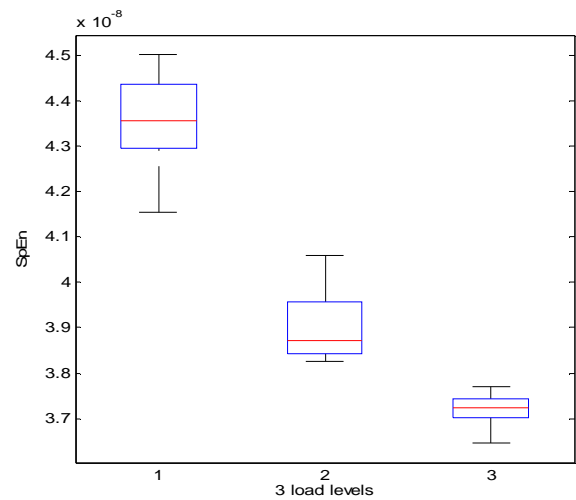

Fig. 2 Boxplot of the SpEn feature extracted from the segmented EEG data across all participants for channel F4. On each box, the red mark is the median; the edges of the box are the 25th and the 75th percentiles. Low level is denoted by 1, medium by 2, and high by 3 here.

C. Classification

To measure the classification accuracy of the extracted features, we used a multi-class support vector machine (SVM) as a classifier. The spectral features were used with three SVMs in a pair wise strategy. One SVM was used to separate low from medium, one used to separate low from high, and the third one was deployed to separate medium from high. The SVM used a linear kernel, and compared the results of the three load level classification for all 32 channels, applied on a per-subject basis. 80% of the data (for each task level for each participant) were used for training and the remaining 20% for testing. We first examined the performance of each feature individually, but it yielded low classification accuracy. Therefore, we combined all the features into the classifier which resulted in a superior performance (higher accuracy rate). This can suggest that each feature captures different information/characteristics of the signal. In other words, they complement one another in defining the cognitive load variations. The classification was performed only on 5channels with the smallest *p*-values from TABLE II. The classification results for all features, on the selected channels which are averaged across all participants are displayed in TABLE III. As shown, 3 channels, namely; Pz, Cp2, Fp2 channels present a high classification accuracy of over 95%.

V. CONCLUSIONS

In this study, we investigated the high ability of few spectral features in discriminating three cognitive load levels. This discrimination was found to be statistically significant using the Enrg, and SpEn as the best performing features, followed by IwMf, EdFr, and IwBw, respectively. The Enrg, and SpEn features trend appeared to decrease as the cognitive load level increased. Although, this trend seemed to increase for the IwMf, EdFr, and IwBw features as the cognitive load level increased. The optimal frequency band was also found to be highly dependent on the spectral features in use. The delta sub-band yielded the best performance for the IwMf, IwBw, and EdFr features. However, this highest performance was achieved for the SpEn and Enrg in the δ_0, and δ_1 sub-bands. Clearly, this needs to be further validated on other databases. Combination of all the features into the classifier resulted in superior performance compared to one feature taken alone, suggesting that each feature captures different information/characteristics of the signal. Therefore, the feature combination defines the cognitive load variations better.

TABLE III
Accuracy of the 3 load level classification for all features by a SVM with linear kernel, averaged over all participants.

EEG channel	P3	Pz	CP2	Fp2	Fz
Classification accuracy %	86.62	95.55	95.55	97.77	82.21

We also determined the few channels which present the highest classification accuracy among the three load level induced. This suggests that a smaller number of EEG channels may be needed for future similar work. However, this needs to be further validated on a larger database. Future work includes collection of EEG data across a larger number of cognitive load levels, which will pose a substantially more difficult classification task.

REFERENCES

[1] F. Paas, J. E. Tuovinen et al., "Cognitive load measurement as means to advance cognitive load theory". In Educational Psychologist, vol. 38, no.1, pp. 63-71, March 2003.

[2] P. Antonenko, F. Paas et al., "Using Electroencephalography to measure cognitive load". In Educational Psychology Review, DOI 10.1007/s10648-010-9130-y, Springer Science and Business MediaLLC 2010.

[3] C. Berka, D. J. Levendowski et al., "Real-Time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset". In Intr. Journal of HumanComputer Interaction, vol. 17, no. 2, pp. 151–170, 2004.

[4] K. M. SpencerandJ. Polich, "Poststimulus EEG spectral analysis and P300: Attention, task, and probability". In Intr. Journal of Psycho physiological Research, vol. 36, no. 2, pp. 220-232, 1999.

[5] J. Lamberts, V.D. Broek, et al., "Correlation dimension of the human EEG corresponds with cognitive load." In Neuropsychology, vol. 41, pp. 149-153, 2000.

[6] P.F. Diez, E. Laciaret et al., "A comparative study of the performance of different spectral estimation methods for classification of mental tasks". In the proceedings of the 30th Int' IEEE EMBS Conference, pp. 1155-1158, 2008.

[7] D. Erdogmus, A. Adami et al.,"Cognitive state estimation based on EEG for augmented cognition." In the 2nd Int. IEEE EMBS Conference on Neural Engineering, pp. 4-7, 2005.

[8] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis". In Brain Research Reviews, vol. 29, pp. 169–195, 1999.

[9] P. Zarjam, J. Epps et al., "Characterizing working memory load using EEG delta activity." To appear in the proceedings of the 19th European Signal Processing Conference, 29 August- 2 September 2011, Barcelona, Spain.

[10] B.R. Greenea, S. Faula et al. "A comparison of quantitative EEG features for neonatal seizure detection." In Clinical Neurophysiology, vol 119, pp. 1248–1261, 2008.

[11] O.M. Doyle, B.R. Greenea et al., "The effect of frequency band on quantitative EEG measures with Hypoxic-ischaemic encephalopathy." In the proceedings the 29th Int' IEEE EMBS Conference, pp.717-721, 2007.

[12] Sh. Sun, "The extreme energy ratio criterion for EEG feature extraction." In the proceedings of ICANN 2008, part II, pp.919-928, 2008.

[13] The Lexile Framework For Reading (www.lexile.com)

[14] http://www.biosemi.com

[15] M. Hollander and D. A. Wolfe, "Data statistical analysis: Nonparametric Statistical Methods." John Wiley & Sons, Inc.

Measuring Cognitive Workload with Low-Cost Electroencephalograph

Avi Knoll, Yang Wang, Fang Chen, Jie Xu,
Natalie Ruiz, Julien Epps, and Pega Zarjam

National ICT Australia
University of New South Wales
{yang.wang,fang.chen,natalie.ruiz,
julien.epps, pega.zarjam}@nicta.com.au
ajk@cse.unsw.edu.au, jie.jackxu@gmail.com

Abstract. Electroencephalography (EEG) is an important physiological index of cognitive workload. While previous research has employed high-end EEG devices, this work investigates the feasibility of measuring cognitive workload with a low-cost EEG system. In our experiment, EEG signals are recorded from subjects performing silent reading tasks under different difficulty levels. Experimental results demonstrate the effectiveness of cognitive workload evaluation even with low-cost EEG equipment.

Keywords: Cognitive workload, electroencephalography (EEG), physiological index.

1   Introduction

In recent years, research efforts have been geared towards measuring human mental states such as cognitive workload and task engagement. Cognitive workload refers to the amount of mental demand imposed by a particular task on a person [3]. Measuring cognitive workload is an important issue in various research and application areas of human-computer interaction, as it can be utilized to evaluate the efficacy of interfaces and build adaptive interaction systems. With the advance of modern sensing technologies, a variety of physiological measures have been developed for the assessment of cognitive workload. Among these techniques, electroencephalography (EEG) has become a popular physiological index that allows continuous monitoring of subjects' cognitive workload in a convenient way.

Previous research has demonstrated that EEG signals are sensitive to cognitive load changes in various tasks [1]. Gevins and Smith [5] demonstrated that spectral features of the theta and alpha frequency bands correlate with task difficulty levels in simulated flight tasks and n-back tests. Fitzgibbon et al. [4] also found that the power of gamma band could be augmented by various cognitive tasks. Berka et al. [2] employed discriminant function analysis on spectral features for monitoring cognitive workload and task engagement in different tasks including digit span, mental arithmetic, image learning and memory tests. Grimes et al. [6] and Zarjam et al. [9] investigated EEG based classification and evaluation of subjects' working memory

load. A feature selection scheme based on mutual information was proposed in [6] to deal with physiological drift. EEG has also been used to monitor cognitive workload in various military tasks under complex environments [8].

Although EEG is a promising tool for continuous measurement of cognitive workload, most previous research has employed high-end EEG systems costing more than $15,000 (e.g. see www.biosemi.com), which limits their widespread usage in human-computer interfaces. On the other hand, low-cost (under $1000) EEG headsets have become accessible for HCI research in recent years [7]. This work takes an initial step in exploring the feasibility of cognitive workload evaluation using a low-cost multi-channel EEG system.

2   Experiment

Sixteen students and employees (16-46 years old, 4 females) were invited to perform silent reading tasks. Brain waves from each subject were recorded with a low-cost EEG device originally designed for gaming interfaces (Emotiv EPOC, a 14 channel 128Hz neuro-signal acquisition and processing wireless neuroheadset [10], see Figure 1). Channel names based on the International 10-20 locations are: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4. During the experiment, each subject was asked to silently read the text displayed on-screen, with a viewing distance of 70cm (see Figure 2). Similar to [9], different task difficulty levels were employed to manipulate cognitive workload during the experiment.
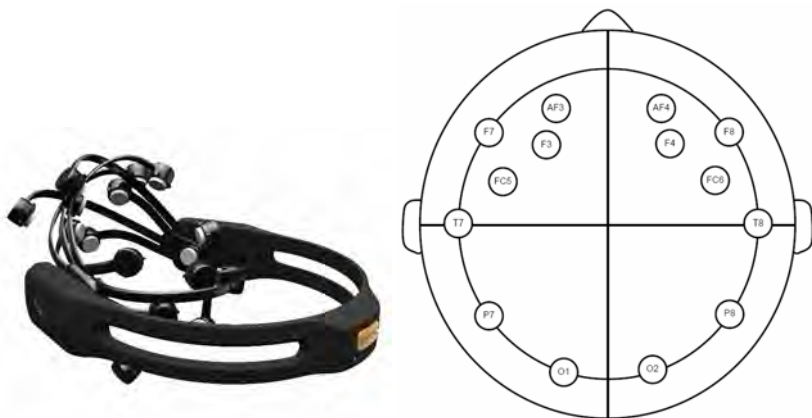

Fig. 1. A low-cost EEG device (Eomtiv EPOC neuroheadset [10]) and layout of EEG channels

There were three levels of task difficulty in total: low (level 1), medium (level 2) and high (level 3). For each difficulty level, 4 text pages were sequentially displayed on the screen, with each page appearing for 30 seconds. In the low level task, the subject was required to press the left mouse button when he encountered any 3 letter word during silent reading. In the medium level task, the subject was required to press

the left or middle button respectively, each time he encountered either a 3 or a 4 letter word. Likewise, in the high level task, the subject was required to press the either the left, middle, or right button when he saw a 3, 4, or 5 letter word respectively. The task difficulty levels were administered in a randomized fashion. There was a 30 second resting period after the task for each difficulty level. One minute baseline data (with the subject looking at a blank screen) was recorded at both the beginning and the end of the whole experiment for each subject. The subject was asked to refrain from eye blinking and to stay as still as possible during the baseline period and task period. However, the subject was free to blink and move their head naturally during each rest period.
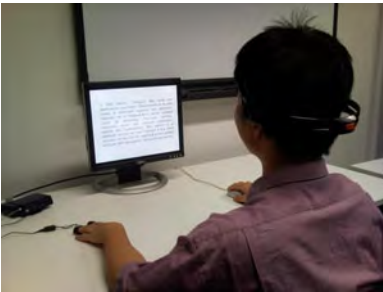

Fig. 2. Experiment setup

3   Analysis

The EEG signals were first divided into segments of 1.5 seconds in length. Statistical features including mean, variance, root mean square (RMS), spectral powers of theta (3-7 Hz), alpha (8-12 Hz), beta (13-29 Hz), and gamma (30+ Hz) frequency bands were then calculated for each data segment.
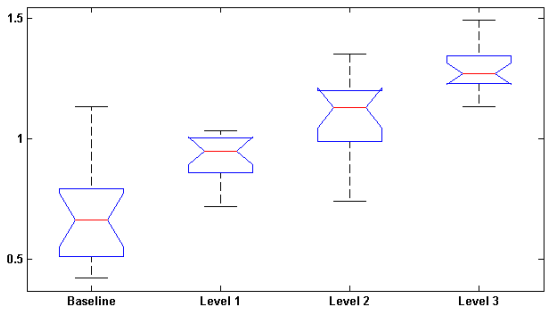

Fig. 3. Box plot of normalized RMS values (sample minimum, lower quartile, median, upper quartile, and maximum) from nodes F3 and F4 at different task difficulty levels

Among the features obtained from different EEG channels, RMS from nodes F3 and F4 exhibit significant correlation with task difficulty ($F > 38$, $p < 0.01$ in ANOVA test). This finding is consistent with previous research indicating that the brain frontal lobes play an important role in cognitive tasks associated with attention and mental effort [5]. Figure 3 plots the distribution of normalized RMS acquired from the two frontal channels at different workload levels for all the subjects. It can be seen that the feature value consistently increases when the task difficulty level is increased.

Moreover, the spectral power of gamma frequency band at nodes AF3 and AF4 shows a statistically significant difference between the baseline condition and task condition ($F > 28$, $p < 0.01$ in ANOVA test), which is consistent with previous study on gamma activation of EEG during cognitive tasks [4]. There is an increase in average gamma power with each rise in task difficulty. However, the difference between task levels is not statistically significant ($p > 0.05$).

4   Conclusion

This work investigates the feasibility of cognitive workload evaluation using a low-cost EEG system. It is demonstrated that cognitive workload could be effectively measured even with low-cost electroencephalograph. The experimental results are consistent with previous research on cognitive workload. We hope that this work will promote the application of EEG-based physiological measures in various HCI areas involving cognitive workload evaluation.

References

1. Antonenko, P., Paas, F., Grabner, R., van Gog, T.: Using electroencepahlography to measure cognitive load. Educational Psychology Review 22, 424–438 (2010)
2. Berka, C., Levendowski, D., Lumicao, M., Yau, A., Davis, G., Zivkovic, V., Olmstead, R., Tremoulet, P., Craven, P.: EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. Aviation, Space, and Environmental Medicine 78, B231–B244 (2007)
3. Cain, B.: A review of the mental workload literature. Technical Report, Defence Research and Development Canada Toronto (2007)
4. Fitzgibbon, S., Pope, K., Mackenzie, L., Clark, C., Willoughby, J.: Cognitive tasks augment gamma EEG power. Clinical Neurophysiology 115, 1802–1809 (2004)
5. Gevins, A., Smith, M.: Neurophysiological measures of cognitive workload during human-computer interaction. Theoretical Issues in Ergonomic Science 4, 113–131 (2003)
6. Grimes, D., Tan, D., Hudson, S., Shenoy, P., Rao, R.: Feasibility and pragmatics of classifying working memory load with an electroencephalograph. In: Proc. ACM CHI Conf. Human Factors in Computing Systems, pp. 835–844 (2008)
7. Lee, J., Tan, D.: Using a low-cost electroencephalograph for task classification in HCI research. In: Proc. ACM Symposium on User Interface Software and Technology, pp. 81–90 (2006)
8. Wilson, G.: An analysis of mental workload in pilots during flight using multiple psychophysiological measures. International Journal of Aviation Psychology 12, 3–18 (2002)
9. Zarjam, P., Epps, J., Chen, F.: Evaluation of working memory load using EEG signals. In: Proc. APSIPA Annual Summit and Conference, pp. 715–719 (2010)
10. http://www.emotiv.com/