CHARACTERIZING WORKING MEMORY LOAD USING EEG DELTA ACTIVITY

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

[1]School of EE&T, The University of New South Wales, Sydney, NSW 2052, Australia
[2]ATP Research Laboratory, National ICT Australia, Eveleigh, NSW 2015, Australia
p.zarjam@student.unsw.edu.au, j.epps@unsw.edu.au, fang.chen@nicta.com.au

ABSTRACT

In this paper, we extract a range of features including time-based, spectral-based, and phase-based to characterize working memory load in EEG recordings from a reading task in which different levels of working memory load were induced. It is demonstrated that a subset of time-based and spectral-based features - the mean, cross-correlation, and energy of the EEG signals - recorded from a few frontal channels in the delta frequency band, and also the statistics of selected wavelet coefficients are representative of working memory load and change most consistently in accordance with the induced load. We show classification accuracy of up to 100% for three working memory load levels across all five subjects. This is achieved using a multi-class support vector machine (SVM) trained on the above features from four frontal EEG channels. We present results suggesting that delta frequency sub-band carries most of the information associated with working memory load. Having used the above features, we also demonstrate that shorter window lengths and a smaller number of EEG channels can be successfully applied for similar contexts.

1. INTRODUCTION

Cognitive load is the amount of task load applied on the working memory by a cognitive process. Recently, there has been increasing interest in maintaining an optimal level of cognitive load applied on the brain, especially in critical decision making fields, such as air traffic control, fire command and military operations, where the imposed task load can be very high and may lead to decreased performance or even failed task completion [1].

Electroencephalography (EEG) measures brain activity by recording the neural electrical fluctuations along the scalp. It is reliable, noninvasive, economical, and is highly sensitive to different cognitive task loads and therefore has been successfully applied in on-line monitoring and measurement of different types of mental activities and workloads in cognitive science and psychology [2]. Recently, the usage of EEG has become more feasible for real-world applications with the availability of wireless EEG systems [3].

A variety of features have already been investigated for mental task classification. These include the use of the statistics of the EEG signals in the time domain; e.g. amplitude values (ERP studies) [4], mean and standard deviation [5], root mean squared (RMS) [6], and absolute/relative/maximum power [7]. Features based on spectral characteristics [8], and the entropy of different sub-bands of the EEG signals have also been proposed [9]. Some work has modeled the EEG signals as an auto-regressive model and used the model parameters as features [10]. However, most of the studies which specifically address the problem of determining working memory, cognitive load level or the related problem of assessing task difficulty, mainly consider PSD or ERP based features for classification [4, 7, 8]. Furthermore, they usually analyze the EEG signals only in two frequency bands in classification of cognitive loads, the alpha and beta, while other bands are ignored [5, 7, 8].

In this study, we examine and compare the suitability of different classes of features for discriminating between varying levels of induced working memory load, towards an understanding of which frequency bands, channels and features are optimal for an EEG-based workload classification system. We conclude that a few features extracted from the EEG signals recorded from the frontal channels represent the working memory load well.

2. EXPERIMENTAL METHODS

2.1. Experiment

Five healthy male volunteers, 24-30 years of age, engaged in postgraduate study, participated in the experiment. A silent reading task, displayed and controlled on a laptop PC with a viewing distance of 70 cm to the participant, was adopted. The 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.

The Lexile framework for reading [11] was used to rate the task readability/complexity and ensure that it induced three different difficulty loads (1020, 1090, and 1150 for the low, medium, and high levels) to the participants' working memory.

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 subjects were asked to sit relaxed and keep their eyes open. This signal was recorded and used later as a reference signal (baseline) for calculating some of the features discussed in the next Section. This is also another way of checking that the change in feature values from low to medium to high is not just signal drift of some kind. To minimize any muscle movement artifact (EMG), the participants were asked to silently sit still and their hand was placed fixed in a certain position, where they could still make finger movements in response to the word stimuli. Since the channels of interest include frontal channels susceptible to ocular artifact, subjects were required to refrain from blinking as much as possible during the recording. Each task level lasted for 2 minutes, and between each two levels the participants were given 30 seconds rest (to blink and stretch). The experiment was repeated twice. Visual inspection of the recorded signals showed that no artifact removal was necessary.

2.2. EEG Recording

The participants' EEG signals were recorded using an Active Two acquisition system [12], 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 at a sampling rate of f_s = 256 Hz. The EEG signals were passed through a built-in band-pass filter with cut-off frequencies of 0.1-100 Hz; the attenuation of this filter at DC was not large, hence the utility of the mean feature is discussed in subsequent Sections. Each participant completed 120 seconds of each task level twice. 30 seconds of the beginning of each level was discarded to make sure the participants are engaged in the task and also remove some EMG artifacts observed.

3. EEG SIGNAL ANALYSIS

The EEG signals $x[n]$ were first segmented using rectangular window length of 5 seconds. The i^{th} segment is represented by $x[n], n = 0,1,..., N-1$ where $N = T \times f_s$ with an overlap of 4 seconds between the successive EEG segments. Initially, we investigated the spectral for the working memory load task using a PSD function.

The results given in Fig.1 show that most of the spectral components of the recorded EEG signals lie below 4Hz, corresponding to the delta sub-band (0-4 Hz) of the EEG frequency bands. This matter is discussed further in Subsection 4.2. Therefore, the recorded signals were mainly studied in the delta sub-band, which we denote as $x_\delta[n]$. From the signal $x_\delta[n]$, we computed the following features for each segment.

Time-based features: This includes MEAN, ENERGY, zero-crossing rates (ZCR), and the maximum cross-correlation (MCOR). The MCOR is the maximum cross-correlation between the EEG signal $x_\delta[n]$ and the baseline condition signal $b_\delta[n]$ proposed by the authors. It is given by the following equation:

$$MCOR = \max_{m=1,2,...,2N-1} R_{xb}(m-N);$$

$$R_{xb}(m) = \begin{cases} \sum_{n=0}^{N-m-1} x[n+m]b[n] & ; m \geq 0 \end{cases} \quad (1)$$

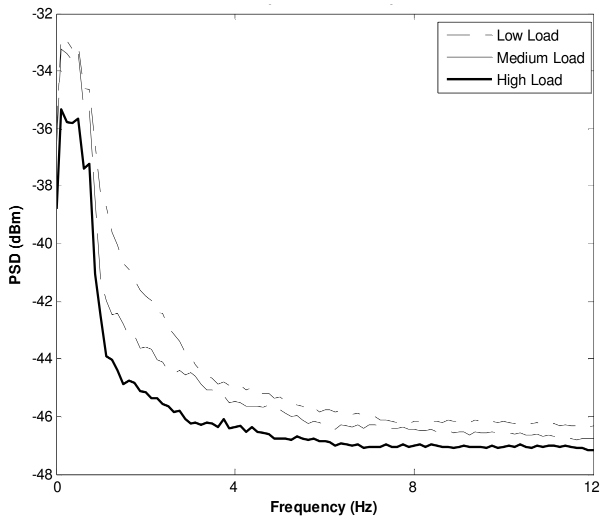

Fig. 1. Example PSDs extracted from the segmented EEG for subject 1, channel Fp1, recorded during imposed high, medium, and low load levels. The EEG frequency bands used are: (0-4Hz) delta, (4-8 Hz) theta, (8-12 Hz) alpha, (12-30 Hz) beta, (30-100 Hz) gamma.

Spectral-based features: This category contains three features; first the frequency at which the PSD attains its peak magnitude value (PSDM), second the signal wavelet decompositions extracted from $x_\delta[n]$. The 5-level wavelet decomposition corresponding to the delta sub-band was also calculated on each EEG segment. The level 5 approximate coefficient was then computed and denoted by $a_\delta[n]$. The wavelet based features proposed herein were extracted from the coefficients of the level 5 approximation $a_\delta[n]$, using the mean (AMEN), minimum (AMIN), and maximum (AMAX).

Third, we computed the spectral coherence (SPC) given by the following equation [13]:

$$SPC = \frac{\left|\langle C_{XB}(f)\rangle\right|^2}{\langle C_{XX}(f)\rangle\langle C_{BB}(f)\rangle} \quad (2)$$

In eq. (2), $C_{XX}(f)$ and $C_{BB}(f)$ are the PSDs of $x_\delta[n]$ and $b_\delta[n]$, $C_{XB}(f)$ is the cross PSD, and $\langle \cdot \rangle$ is the averaging operator. The signal $b_\delta[n]$ represents a segment of the same size from the baseline EEG (i.e. resting condition, no cognitive load) in the delta frequency band for the same channel as $x_\delta[n]$. The SPC has a value in the interval of $[0,1]$. A value of "1" means that the corresponding frequency components of the two signals are identical, except for a multiplicative amplitude difference and a constant time relation (phase delay). A "0" value indicates that the corresponding frequency components of the both signals are not correlated.

Phase-based features: The first feature of this kind represents the mean of the instantaneous frequency (IFME). The instantaneous frequency (IF) of a given non-stationary signal shows how its frequency content changes with time and the mean is given by the following equation [14]:

$$IFME = \frac{1}{N-1}\sum_{n=1}^{N-2}\frac{f_s}{4\pi}(\arg z[n+1]-\arg z[n-1]) \bmod 2\pi \quad (3)$$

where $z[n]$ is the analytic signal associated with $x_\delta[n]$, i.e. $z[n] = x_\delta[n] + H\{x_\delta[n]\}$, $H\{.\}$ denotes the Hilbert transform.

The second phase-based feature examined here is the phase locking value (PLV), applied previously in BCI. It measures the variability in instantaneous phase difference $\varphi_x[n]-\varphi_b[n]$ between the EEG channel signal $x[n]$ and the baseline signal $b[n]$ [13]:

$$PLV = \left|\left\langle \exp(j(\varphi_x[n]-\varphi_b[n]))\right\rangle\right| \quad (4)$$

The instantaneous phases are determined using the Hilbert transform, e.g. as $\varphi_x[n] = \arctan(H\{x_\delta[n]\}/x_\delta[n])$. PLV has also a value in the interval of $[0,1]$. When the phase difference is constant, there is phase synchronization so the PLV is equal to "1". If the phase differences are randomly distributed over $[0, 2\pi]$, the PLV will be "0". Here we consider the PLV between the task condition signal and the baseline condition signal.

4. RESULTS

4.1. Feature Comparison

The performance of all features introduced in Section 3 was examined and compared for all the 32 EEG channels. Table 1 displays the statistical evaluation for all features extracted from channel Fp1 for subject 1, using a paired t-test to generate *p*-values. Here, the values of the extracted features ZCR, IFME, PSDM, SPC, and (to a lesser extent) PLV show significant overlaps between different load levels and are not able to discriminate between different levels of load, seen also in the high *p*-values. Hence, phase-based features were broadly found to be less promising for this application. On the other hand, the values of the remaining time and spectral-based features; namely: the MEAN, ENERGY, MCOR of $x_\delta[n]$ and AMEN, AMIN, and AMAX of $a_\delta[n]$ in different load levels have no overlap; these features were found to distinguish different load levels completely in some below channels for all subjects, also indicated by *p*-values close to zero. These discriminative features were also subsequently applied to a classifier for classification evaluation, in Subsection 4.5.

As seen in Table 1, the MEAN, AMEN, AMIN, and AMAX of channel Fp1 exhibit a consistent increasing trend as load level increases. This shows that level shifts in these features are proportional to the task load level. However, the ENERGY and MCOR show decreasing trends. Similar trends were observed across all frontal channels (Fp1, Fp2, AF3, AF4, Fz, F3, and F4) for all subjects as working memory load increases.

For illustration purposes, the AMIN feature extracted from the delta band of the EEG signals representing three different task difficulties for one subject is shown in Fig 2. It reveals that as the task difficulty increases, the AMIN of the EEG signal tends to increase, and therefore each load level is clearly distinguishable from other levels.

4.2. Working Memory Load Discrimination by Frequency Band

Although we carried out initial examination of the working memory load discrimination as a function of frequency in Section 3, we investigated this matter further, to quantify the differences between bands. Accordingly, we examined the working memory load discrimination capability of different frequency bands, using an energy feature in each case. Fig. 3 shows the means and 95% confidence intervals of the energies extracted for each of the delta, theta, alpha, beta and gamma bands for the Fp1 channel of subject 1. Clearly the delta frequency band is the only band that provides cognitive/working memory load discrimination for the given experiment.

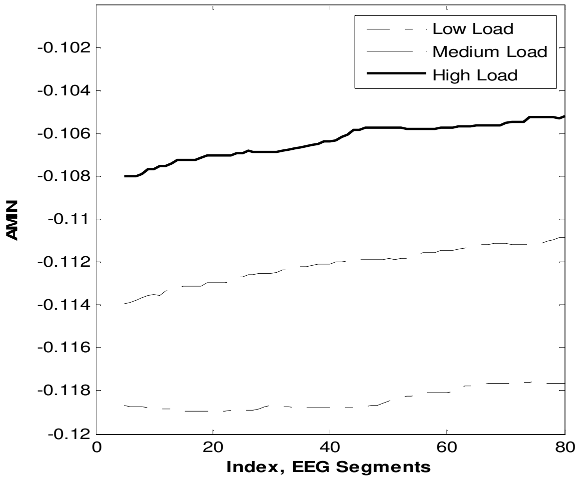

Fig. 2. Values for the AMIN feature extracted from the segmented Fp1 EEG data for the first participant. The top, middle, and bottom curves belong to the EEG signals recorded during high, medium, and low levels of task difficulty, respectively.

Table 1. Results of feature calculation for different load levels of the EEG signals acquired from channel Fp1 of subject 1.

Extracted Feature	Load Level			*p*-value
	Low	Medium	High	
MEAN	-0.0209 ±0.0001	-0.0198 ±0.0002	-0.0188 ±0.0002	1.2346e-089
ENERGY	0.5583 ±0.0069	0.5017 ±0.0118	0.4513 ±0.0098	1.6904e-089
ZCR	10.5000 ±9.5000	10 ±9	10.5000 ±9.5000	0.5225
IFME	0.5227 ±0.4724	0.5033 ±0.4530	0.5250 ±0.4746	0.5065
PSDM	0.4375 ±0.3125	0.6250 ±0.5000	0.5000 ±0.3750	0.9530
SPC	0.9925 ±0.0045	0.9924 ±0.0032	0.9917 ±0.0043	0.0854
MCOR	0.6907 ±0.0051	0.6544 ±0.0068	0.6203 ±0.0064	2.0476e-089
AMEN	-0.1182 ±0.0007	-0.1120 ±0.0013	-0.1062 ±0.0012	2.0119e-089
AMIN	-0.1182 ±0.0007	-0.1121 ±0.0013	-0.1064 ±0.0012	1.8237e-088
AMAX	-0.1180 ±0.0008	-0.1116 ±0.0015	-0.1059 ±0.0012	1.0998e-087
PLV	1.0000 ±1.000	1.0000 ±1.000	1.0000 ±1.000	0.0331

This was supported by similar results for other discriminative features from the frontal channels across all subjects. Given that the task is a silent reading task, this finding confirms previous studies showing that delta activity is an indicator of attention during mental tasks, so that by increasing task demand, participant attention to the task and also the delta band activity increase [15, 16]. So far, just a few related studies analyze the delta band and the remainder ignore it, citing the likelihood of artifacts (especially EOG) appearing in this frequency band. However in this study, we carefully avoided EEG signals with EOG contamination and consider it very unlikely that the very consistent feature changes during the different work load levels across all the subjects are due to EOG.

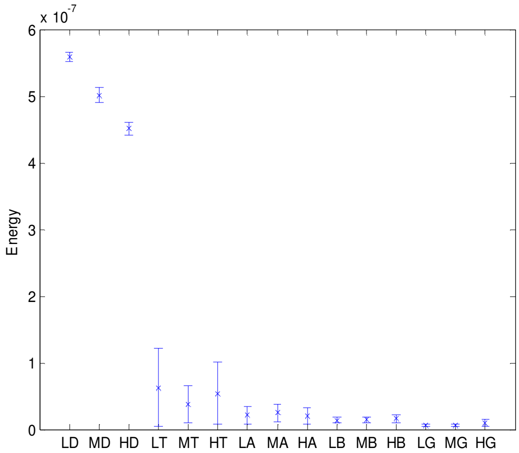

Fig. 3. Means (x) and 95% confidence intervals for energy features extracted for delta (D), theta (T), alpha (A), beta (B) and gamma (G) bands, for low (L), medium (M) and high (H) cognitive loads, for Fp1 channel in subject 1. The delta values have been reduced by a factor of 10^6 for comparison purposes.

4.3. Working Memory Load Discrimination by Electrode Position

In order to understand which channels yield the best performance for discrimination purposes, all 32 channels were initially examined using the MEAN feature. Specifically, the difference between feature values between two adjacent load levels was computed, and the results displayed as a spatial heat map. This was repeated across all subjects, and for different discriminative features. It was found that mainly the frontal EEG channels mentioned in Subsection 4.1 were the most discriminative channels for working memory load separation in this experiment, showing highly consistent trends across all 5 subjects, for the successful features identified earlier in this Section.

For illustration purposes, the MEAN features extracted from the EEG signals of one subject representing two different load levels (Low and Medium) is shown in Fig. 4. This reveals an area of maximum difference between feature values (i.e. close to "1") that perfectly covers the frontal lobe, especially Fp2 and F4 on the right side of the frontal lobe and Fp1 and AF3 on the left side; similar results were observed across all 5 subjects.

These results are in line with the related findings indicating that the increasing memory load in the working memory is reflected in the brain frontal lobe [17]. However, this is the current understanding of association of the frontal lobe with the working memory. To our knowledge, the exact corresponding EEG electrodes positions have not been identified previously.

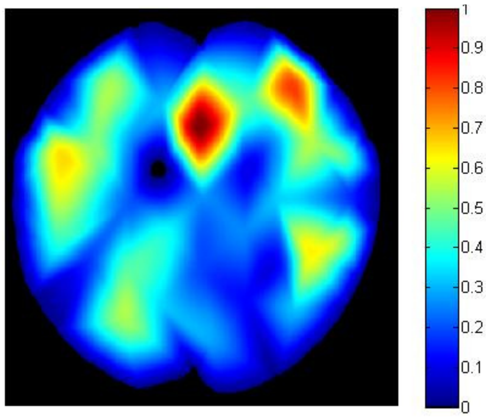
Fig. 4. Spatial heat map for the difference in MEAN feature between low and medium load levels for one subject, with frontal electrodes seen at the top. Note that the scale has been normalized, so that "0" reflects no difference and "1" reflects the maximum difference in feature values.

4.4. The Effect of Window Length

We also investigated the effect of the window length T on the discriminative capability of the features. Specifically, the 95% confidence interval of the MEAN as a function of the window length was calculated for each load level, for $T = 1,2,\ldots,10$ *seconds*. Results showed that shorter windows produced only slightly (<10%) wider confidence intervals than longer windows, suggesting that a window length as small as $T = 1$ second could be used effectively for cognitive/working memory load measurement. This is a preliminary result however, and should be investigated further experimentally using more challenging classification tasks, e.g. either more levels of induced memory load or less separable load levels.

4.5. Working Memory Load Classification

In order to gauge the classification performance for this data set across all subjects, a selection of the features mentioned in Section 3 were then used to train a multi-class SVM. To successfully classify three load levels; Low (L), Medium (M), and High (H), two SVM classifiers were used; one classified L from M and H, and the second SVM classified H from L and M. Here, we deployed SVM with a linear kernel, and compared the results of the three load level classification for all the frontal channels, applied on a per-subject basis. 50% of the data (for each task level for each subject) were used for training and the remaining 50% for testing, and only those features with the smallest *p*-values from Table 1 were considered.

We performed the SVM classification on the frontal EEG channels mentioned in Subsection 4.1. The results showed that mostly the four channels offer the highest classification accuracy. Namely; channels Fp1, Fp2, AF3, and F4 which are displayed in Table 2. As shown, the features MEAN, AMEN, and ENERGY provide the highest accuracy, 100% in this database, followed by AMIN, AMAX, and finally MCOR. It is also shown that the two pre-frontal channels, Fp1 and Fp2, slightly outperform the other two frontal channels, AF3 and F4, in terms of classification accuracy.

Table 2: Accuracy of 3-class classification for different features by SVM with linear kernel, averaged over 5 subjects.

Feature	3 load level classification accuracy %			
	Channel (number)			
	ChFp1 (1)	ChFp2 (30)	ChAF3 (2)	ChF4 (27)
MEAN	100	99.82	97.54	97.01
AMEN	100	100	97.01	97.89
AMIN	100	99.64	98.06	97.19
AMAX	97.19	98.24	95.26	97.71
ENERGY	100	100	98.24	97.54
MCOR	91.57	97.54	83.67	98.06

5. DISCUSSION AND CONCLUSION

In this paper, we investigated the characterization of working memory load using different classes of features extracted from frontal EEG signals. Discrimination between the different workloads induced during a reading task was found to be highly significant using the following time-based features: mean, energy, maximum cross-correlation, and some spectral-based features such as approximate wavelet coefficients (5-level decomposition). The approximate wavelet coefficients and maximum cross-correlation based features were proposed herein. The signal's mean and approximate wavelet coefficient values appeared to increase when the workload increased. This could be due to the fact that DC potentials of the cortical excitation shift in response to changes in attentional demands of the task. These shifts were shown to be proportional with the task level demand. For the cross-correlation feature, as the task load level increased, the cross-correlation with the baseline decreased, implying less similarity between the two signals. It is worth mentioning that in a preliminary experiment, all features except PLV were investigated across all the EEG frequency bands.

The main result of the frequency-band investigation was that the delta frequency band appeared to carry the vast majority of the information relating to working memory load level in this experiment, even though only a few studies relating EEG and mental activity have analysed the delta band previously. So the EEG low frequency activity contains significant electrophysiological correlates of cognitive processing and should receive particular attention. Furthermore, it seems to be related to an increase in the participants' internal concentration during tasks.

We also specifically identified a few frontal channels which contain the most consistently discriminative information related to mental effort, so that a smaller number of EEG channels may be needed for future similar work. The results also suggest that relatively high discrimination among the load levels can be obtained with shorter windows.

In conclusion, obtaining high classification accuracies using the proposed features with shorter windows and a smaller number of channels is a step forward towards using these technologies in real-time and in more realistic environments, at low complexity.

6. ACKNOWLEDGEMENT

The authors would like to acknowledge the volunteers for participating in the experiment, and the assistance of Mr. M. A. Khawaja for testing the designed reading task against the Lexile analyser. This research was supported by the Asian Office of Aerospace Research & Development, Grant No. FA2386-10-1-4029.

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] 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.

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

[4] K. M. Spencer and J. 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, 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.

[6] 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 Proc. of the 30th Int' IEEE EMBS Conference*, pp. 1155-1158, 2008.

[7] 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.

[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] L. Zhiwei and S. Minfen, "Classification of mental task EEG signals using wavelet packet entropy and SVM". *In Proc. of the ICEMI'2007*, vol. 3, pp. 906-909, 2007.

[10] N. Huan and R. Palaniappan, "Classification of mental tasks using fixed and adaptive autoregressive models of EEG signals". *In the IMBS'04 Proc.*, pp. 507 – 510, 2004.

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

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

[13] E. Gysels and P.Celka, "Phase synchronization for the recognition of mental tasks in a brain–computer interface". *In IEEE Trans. on Neural Sys. and Rehabilitation Eng.*, vol. 12, no. 4, 2004.

[14] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal part-II: Algorithms and applications". *In the Proc. of the IEEE*, vol 80, pp.540-568, 1992

[15] T. Harmony, T. FemGndez, and et al. "EEG delta activity: An indicator of attention to internal processing during performance of mental tasks". *In Inter. Journal of Psychophysiology*, vol 24, pp. 161- 17, 1996. 4, pp. 161- 17, 1996.

[16] T. Harmony, T. FemGndez, and et al. "EEG activation patterns during the performance of tasks involving different components of mental calculation". *In Electroencephalography and Clinical Neurophysiology*, vol 94, pp. 175-182, 1995.

[17] A. Stipacek, R. H. Grabner, and et. al, "Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load". *In Neuroscience Letters,* vol 353, pp. 193-196, 2003.