

## **Discriminating cognitive performance using biomarkers extracted from linear and nonlinear analysis of EEG signals by machine learning**

Shivabalan K R<sup>1</sup>, BROTO DEB<sup>2</sup>, Shivam Goel<sup>3</sup>, Sivanesan S<sup>4</sup>

1. Mahatma Gandhi Medical College and Research Institute, Pondicherry
2. Mayo Clinic: Rochester, Minnesota, USA
3. All India Institute of Medical Sciences, New Delhi
4. Post Graduate Institute of Medical Education and Research, Chandigarh

**Nonlinear dynamics and chaos theory are being widely used nowadays in neuroscience to characterize complex systems within which the change of the output isn't proportional to the change of the input. Nonlinear systems compared to linear systems, often appear chaotic, unpredictable, or counterintuitive, and yet their behaviour isn't random. The importance of the time series analysis, which exhibits a typical complex dynamics, within the area of nonlinear analysis can't be undermined. Hidden important dynamical properties of the physiological phenomenon can be detected by many features of these approaches. Nonlinear dynamics and chaos theory are being employed in neurophysiology with the aim to elucidate the complex brain activity from electroencephalographic (EEG) signals. The brain is a chaotic dynamical system and further, their generated EEG signals are generally chaotic in another sense, because, with respect to time, the amplitude changes continuously. A reliable and non-invasive measurement of memory load which will be made continuously while performing a cognitive task would be very helpful for assessing cognitive function, crucial for the prevention of decision-making errors, and also the development of adaptive user interfaces. Such a measurement could help to keep up the efficiency and productivity in task completion, work performance, and to avoid cognitive overload, especially in critical/high mental load workplaces like traffic control, military operations, and rescue commands. We have measured the linear and nonlinear dynamics of the EEG signals in subjects undergoing mental arithmetic task and measured the cognitive load on the brain continuously. We have also differentiated the subjects who can perform a mental task good and bad and developed a system using support vector machine to differentiate rest and task states.**

Complex task performance requires the integration of knowledge related to the task, working memory, attention and decision making. Cognitive load refers to the amount of mental demand imposed by a particular task on these parameters<sup>1</sup>. A reliable, objective and cost effective method of assessing cognitive load is crucial in understanding the mental capacity of people. Experts have more knowledge or experience with regard to a specific task which reduces the cognitive load associated with the task in contrast to novices who have a heavier cognitive load<sup>2</sup>. Measuring and assessing the cognitive load associated with different tasks is crucial for many applications, monitoring the mental well-being of people who require heavy cognitive load in their occupation. Heavy cognitive load increases the chances of error in the task at hand<sup>3-5</sup> or stereotyping<sup>6</sup>. Subjective measures of cognitive load<sup>4</sup> may provide useful for cognitive load measurement in embodied scenarios if an appropriate survey is chosen. However, the use of cognitive load questionnaires has its own theoretical and practical issues. Different phrasings in cognitive load question items might lead to results that may not be comparable<sup>7</sup>. In addition to behavioral measures, cognitive workload has been assessed using physiological measures, such as pupil diameter<sup>8</sup>. The application of objectively determining the cognitive load has its growing application in ergonomics among the pilots or drivers.<sup>9</sup>

Electroencephalography is an easy, cost effective investigation to assess the brain activity. The focus of this paper is on quantifying cognitive workload using measures based on EEG. EEG-based measures for cognitive load have been described in several studies. The relationship among different spectral features to help predict cognitive load from EEG has been described in many studies<sup>10-12</sup>.

A wide range of methods have been applied for measuring and classifying the memory load using EEG signal. They can use either linear or nonlinear dynamic features to study them. Complex signals are information-rich. Another reason behind our preference to utilization of nonlinear measures in analysis of EEG is that nonlinear (nonstationary and irregular) signals nullify comprehensive understanding by a classic reductionist approach. Even the simplest nonlinear signals (originating from complex systems) will nullify the criteria of proportionality and superposition characteristic for linear systems<sup>13</sup>. The application of non-linear methods in classifying mental tasks is more recent, and measures like correlation dimension (CD)<sup>14-16</sup>, fractal dimension<sup>17-19</sup>, Hurst exponent (HE)<sup>9,15,20</sup>, sample entropy<sup>18</sup>, approximate entropy (ApEn)<sup>15,21,22</sup> and largest Lyapunov exponent (LLE)<sup>14,15</sup> have been used to measure the complexity/irregularity of the underlying brain dynamics during the performance of some cognitive tasks compared with the rest condition. Stated differently, in these studies the brain activity states; such as rest and stimulated have been differentiated. But to date, these measures have not been investigated in the analysis of the varying working memory load and the question whether these approaches could provide some information on discriminating subjects who can perform the arithmetic task well and subjects who can't perform well.

In this study we have tried to understand the chaotic nature of the brain waves in the face of serial cognitive workload in normal subjects and discrimination of the 'good' and 'bad' performers when faced with increasing cognitive load.

In future, this study might be extrapolated to subjects with dementia or minimal cognitive impairment and will detect such degenerative processes at an earlier stage. It could also help in ensuring productivity in task completion and ergonomics.

## Results

### Compare rest with task state

First, we calculated the PSD, SE, FE and FD of five frequency bands (delta, theta, alpha, beta and gamma) and whole frequency signal. The alpha band changes from low-amplitude, non-rhythmic electrical activity to high-amplitude, oscillatory activity during the period of eye-closure<sup>23</sup>. We also calculated the ratio of features on alpha band to features on other bands (delta, theta, beta and gamma). The PSD results were shown in Figure 4.

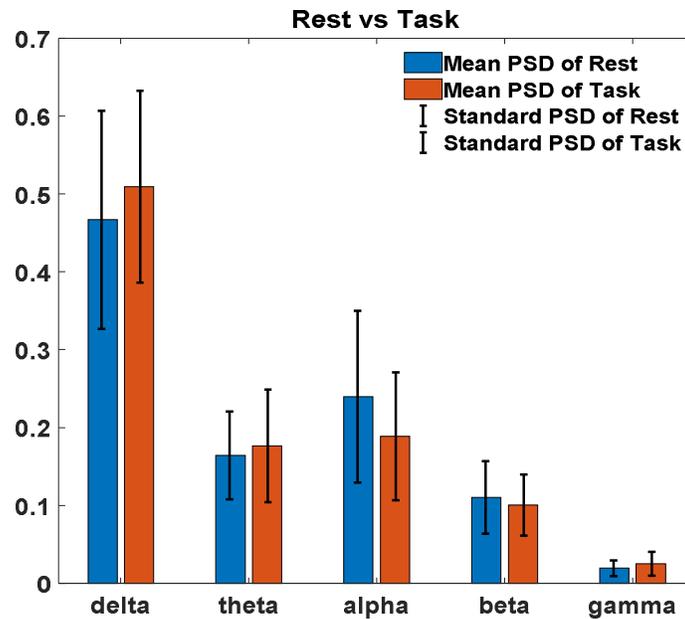


Figure 4. PSD of five frequency bands on FP 1 used to compare rest state with task state.

Then, the significant difference between rest and task states based on these features were also analyzed. First, we tested whether the data satisfied the assumption of normality and the assumption of homogeneity of variance. We chose the parametric test ANOVA1, which is a one-way analysis of variance for the data that satisfied the assumptions. Otherwise, a nonparametric test Kruskal-Wallis was used. It's shown that the feature *log(PSD of alpha/PSD of gamma)* has the best performance and the results are shown in Figure 5. We found that the subjects had a higher alpha and lower gamma during rest state.

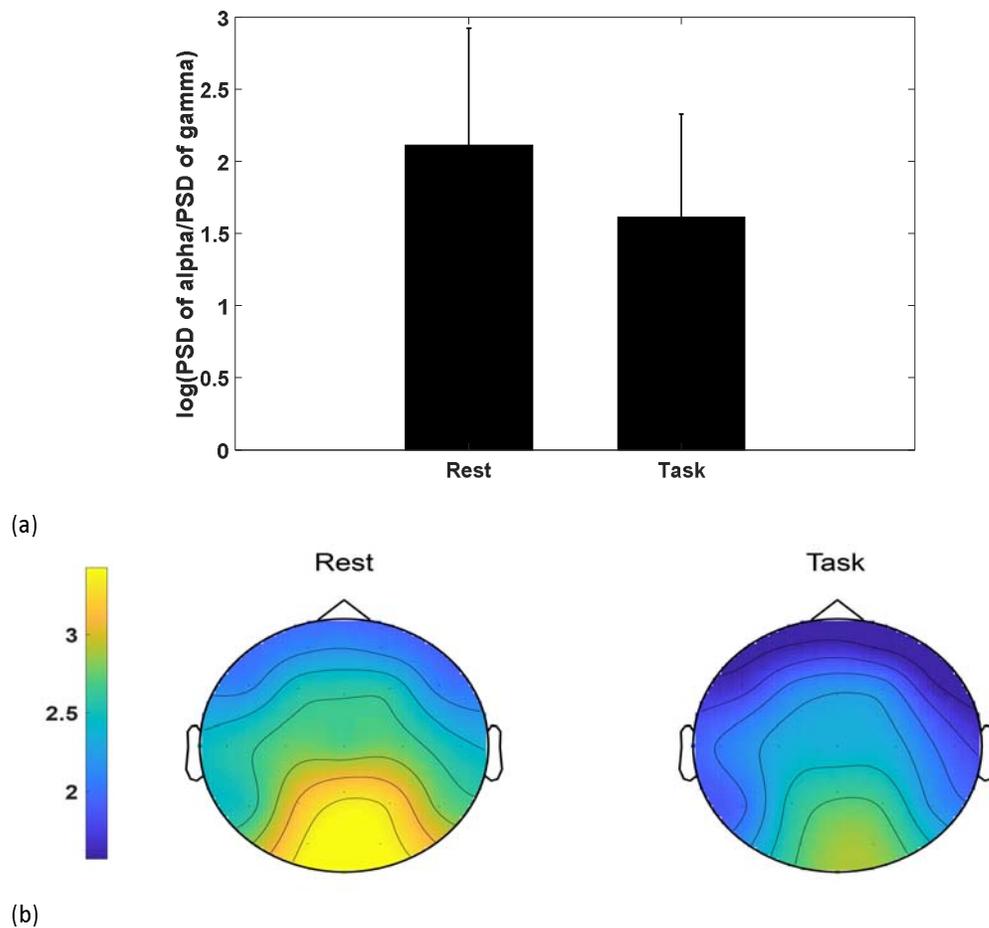


Figure 5. (a) the feature  $\log(\text{PSD of alpha}/\text{PSD of gamma})$  on FP1 for rest and task states (b) Spectral topographic map of power for  $\log(\text{PSD of alpha}/\text{PSD of gamma})$  for rest group and task group. Upper portion for each map shows the nasal and lower portion shows the occipital side. The color bar represents the log-transformed spectral power density ( $10 \cdot \log_{10}(\mu\text{V}^2/\text{Hz})$ ) where yellow represents the maximum and blue represents the minimum values.

#### Compare good with bad quality

First, we calculated the PSD, SE, FE and FD of five frequency bands (delta, theta, alpha, beta and gamma), ratio of alpha to each of the frequency band and ratio of the alpha to the whole frequency band. The PSD results were shown in Figure 6.

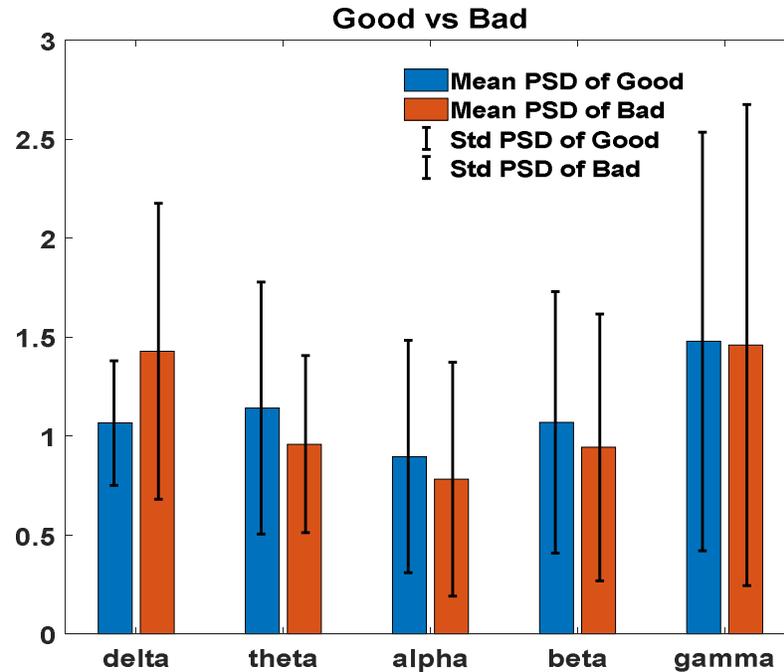


Figure 6. PSD of five frequency bands on FP 1 used to compare good quality with bad quality. Theta and beta band oscillations directly reflect such cognitive processes. So we focused on theta and beta on PSD, SE, FE and FD, the results were shown in Figure 7 and Figure 8. As we can see that, the result is significant in frontal, temporal and occipital lobes. Good group has lower entropy and FD of theta and beta band in frontal lobe than bad group, while good group has higher entropy and FD of theta and beta in occipital lobe and temporal lobe. Good group has lower PSD of theta and beta bands in occipital lobe than bad group. Good group has higher PSD of theta and beta bands in left-frontal lobe than bad group.

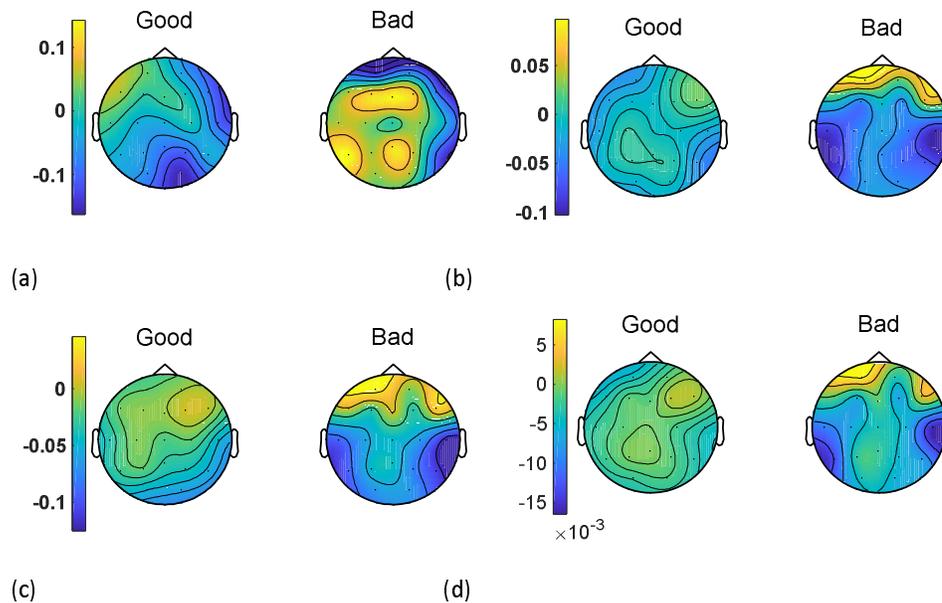


Figure 7. Spectral topographic map for theta band for good group and bad group. Upper portion

for each map shows the nasal and lower portion shows the occipital side. The color bar represents the log-transformed spectral power density ( $10 \cdot \log_{10} (\mu\text{v}^2 / \text{Hz})$ ) where yellow represents the maximum and blue represents the minimum values. (a) PSD (b) SE (c) FE (d) FD.

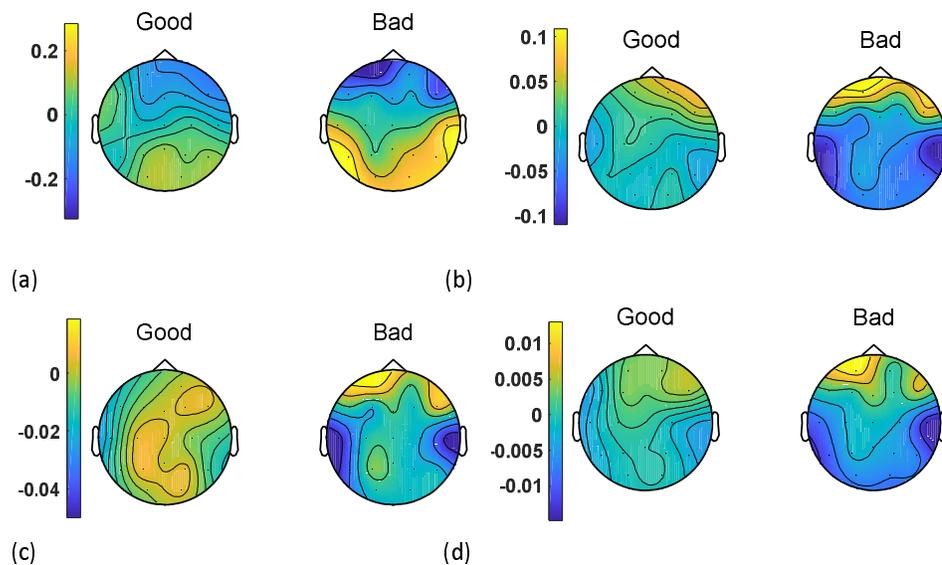


Figure 8. Spectral topographic map for beta band for good group and bad group. Upper portion for each map shows the nasal and lower portion shows the occipital side. The color bar shows the log-transformed spectral power density ( $10 \cdot \log_{10} (\mu\text{v}^2 / \text{Hz})$ ) where yellow represents the maximum and blue represents the minimum values. (a) PSD (b) SE (c) FE (d) FD.

#### Classification

The extracted feature  $\log(\text{psd of alpha} / \text{psd of gamma})$  was fed into the Support vector machine (SVM), k-nearest neighbor (KNN) and Linear Discriminant Analysis (LDA) for classification of rest and task state in the MATLAB (R2018b) (MathWorks Inc., Natick, MA). The 10-fold cross-validation was used before LIBSVM to divide the samples into 10 parts. One of the 10 parts was used as a testing set and the remaining nine parts were used as a training set. To avoid information leakage, the samples from the same subject were divided into either training set or testing set. The k-fold cross validation method is used to divide the actual sample into k subsamples. Here k is chosen as 10. The mean value of 10-fold cross-validation will be used as the accuracy of this model. SVM produced the highest accuracy **85.31%**, while the accuracy of KNN and LDA were **83.44%** and **77.05%**.

Thereafter, we analyzed the performance of SVM in detail. The performance of the SVM for classification depends on the choice of a kernel. Optimization of parameters of a kernel can train our classifier for a given dataset and improve classification accuracy of a classifier. In this paper, the Polynomial Kernel is chosen. The degree of Polynomial Kernel has a direct influence on flexibility of resulting classifier<sup>24</sup>. The Polynomial Kernel is a global kernel which has a good generalization ability. Parameters of a kernel have a significant effect on the decision boundary. It can classify data with nonlinear boundaries as well as of high dimensions<sup>24</sup>. After the SVM provides the results, the confusion matrix, which contains **precision**, **recall**, **sensitivity**, and **specificity**, is constructed. The results were shown in Figure 9 and Table 3. The receiver operating characteristic curve (ROC) of SVM had an AUROC of 88.2% as shown in Figure 10.

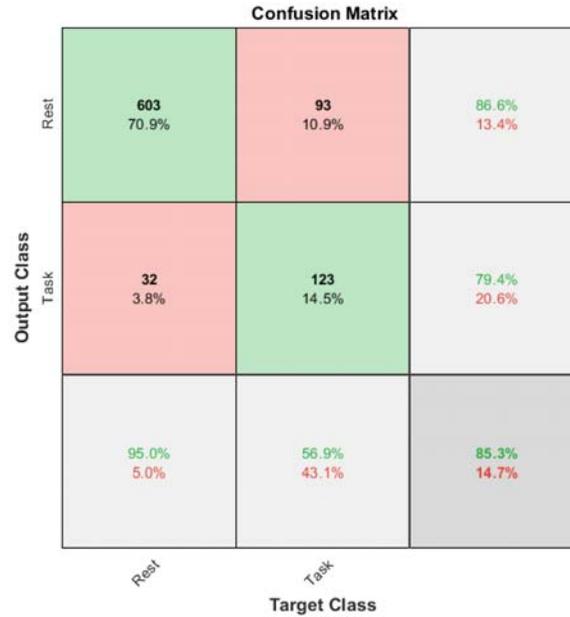


Figure 9. Support vector machine (SVM) polynomial kernel model confusion matrix

Table 3. Precision, recall/sensitivity, specificity and accuracy of SVM model

Methods	Evaluations	Value
Log(PSD of alpha/PSD of gamma)	Precision	86.64%
	Recall/Sensitivity	94.96%
	Specificity	56.94%
	Accuracy	85.31%

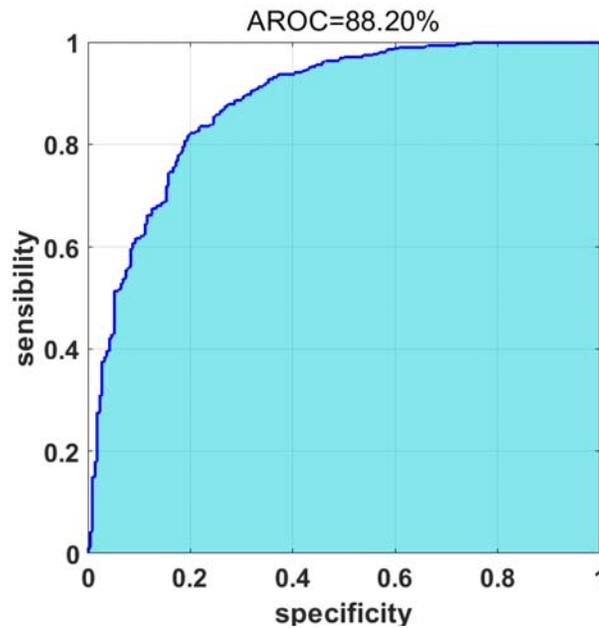


Figure 10. Receiver operating characteristic curve (ROC) of  $\log(\text{psd of alpha/psd of gamma})$  for discriminating rest and task state.

## Discussion

In this paper we have elucidated the chaotic nature of the brain waves in the face of serial cognitive workload in normal subjects and discrimination of the 'good' and 'bad' performers when faced with increasing cognitive load. The broader objective of this work is to develop a biomarker which can discriminate good performers from bad performers in a mental arithmetic task.

Beta and theta band oscillations directly reflected cognitive processes. This finding wasn't very much congruent with the papers which reported that alpha and theta oscillations were reflective of cognitive and memory performance<sup>25-27</sup>. Although it was in congruence with the findings described by Molnar et al<sup>28</sup>. We observed that, there were significant differences among the frontal, temporal and occipital lobes. Good group had lower entropy of theta and beta band in frontal lobe but higher in occipital and temporal lobe when compared to the bad group. This implies that the complexity of the brain dynamics is less in the frontal and more in the occipital and temporal groups in complex cognitive processes. Or in other words, higher the mental work load on the memory, lower the entropy of the EEG signals in the frontal and higher in the occipital and temporal lobes. A decreased value of entropy with increased task load implies higher predictability and less irregularity in the brain activity. This finding is supported by Zarjam et al who showed that entropy decreases with increase in cognitive load mainly in the frontal lobe but it was more prominent in the delta sub band and they had measured approximate entropy and spectral entropy<sup>20</sup>. Both the studies show that brain behaves in a more focused and regular manner when performing more difficult tasks. Entropy correlated with fractal dimension positively. Or in other words Fractal dimension behaved similarly to that of entropy. Good group had lower entropy and FD of theta and beta band in frontal lobe than bad group, while good group had higher entropy and FD of theta and beta in occipital lobe and temporal lobe. While Wang et al showed that Fractal dimension increased with increased cognitive load mainly in the frontal lobe<sup>29</sup>. This contradiction couldn't be explained by us. PSDs of theta and beta bands were higher in frontal but lower in occipital lobe when compared to bad group. This shows that higher PSD correlates with higher mental workload in frontal lobe. Zafar et al showed the involvement of delta oscillations in occipital region during cognitive tasks and they also showed that during the cognitive task the delta band has a different behavior compared to the rest task<sup>30</sup>. And they didn't classify the task group into good and bad performers. This might be explained by the fact that their experiment was done with the subjects' eyes' closed.

But we have developed a marker which can differentiate rest and task states and the same was fed into SVM which had a greater predictive capacity. Due to the processed non-linear nature and a good generalizability of the feature set, the polynomial kernel function was used in our SVM model, as such approaches has been reported in similar situations. The feature set was subjected to various feature selection methods to narrow down more prominent features before using a SVM model. ***Log(psd of alpha/psd of gamma)*** had a better predictive capacity with a good area under ROC to discriminate rest from task states. SVM produced the highest accuracy **85.31%**, while the accuracy of KNN and LDA were **83.44%** and **77.05%**. Attallah used the same dataset to classify mental stress states and non-stress states using Principal Component analysis(PCA)<sup>31</sup>. PCA reduces the features extracted from such electrodes to lower model complexity, where the optimal number of principal components is examined using sequential forward procedure. It also examines the minimum number of electrodes placed on the site which has greater impact on stress detection. They showed that using only 58 and 15 principal components, the accuracy of detecting stress reached 100% and 99.8% using cubic SVM

and KNN classifiers respectively whereas our study showed an accuracy of **85.31%** and **83.44%** with SVM and KNN respectively.

But machine learning applied to the good and bad performers to identify a marker couldn't yield a similar statistically significant biomarker. Possibly the sample size was small to achieve a significant discriminating biomarker. Another limitation of the study is that for each arithmetic calculation load the amount of change in the feature couldn't be quantified because of the gross variation for each subject in the calculation they could perform.

**Conclusion** Machine learning by SVM revealed that *Log(psd of alpha/psd of gamma)* can differentiate mental arithmetic task and rest states with eyes closed. Frontal lobe is affected more widely and deeply during mental arithmetic task cognitive load. Entropy and fractal dimension decreased in the frontal lobe proportional to the cognitive load.

## Methods

### Participants

The dataset was obtained from the PhysioNet "EEG During Mental Arithmetic Tasks" database contributed by contributed by Igor Zyma, Sergii Tukaev, and Ivan Seleznev, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Department of Electronic Engineering<sup>32,33</sup>. (Zyma et al, Goldberger et al physiobank)

Totally, 66 healthy right-handed volunteers (47 women and 19 men) were initially involved in the study. All participants are 1st–3rd year students of the Taras Shevchenko National University of Kyiv (Educational and Scientific Centre "Institute of Biology and Medicine" and Faculty of Psychology) aged 18 to 26 years (Mean = 18.6 years, Standard Deviation (SD) = 0.87 years). The details of the participants are described in Table 1.

Subject	Age	Gender	Number of Subtractions	Count Quality
0	21	F	9.7	B
1	18	F	29.35	G
2	19	F	12.88	G
3	17	F	31	G
4	17	F	8.6	B
5	16	F	20.71	G
6	18	M	4.35	B
7	18	F	13.38	G
8	26	M	18.24	G
9	16	F	7	B
10	17	F	1	B
11	18	F	26	G
12	17	F	26.36	G
13	24	M	34	G
14	17	F	9	B
15	17	F	22.18	G
16	17	F	11.59	G

17	17	F	28.7	G
18	17	F	20	G
19	22	M	7.06	B
20	17	F	15.41	G
21	20	F	1	B
22	19	F	4.47	B
23	16	F	27.47	G
24	17	M	14.76	G
25	17	M	30.53	G
26	17	F	13.59	G
27	19	F	34.59	G
28	19	F	27	G
29	19	M	16.59	G
30	17	M	10	B
31	19	F	19.88	G
32	20	F	13	G
33	17	M	21.47	G
34	18	F	31	G
35	17	F	12.18	G

Table 1. Participant details

The participants were eligible to enroll in the study if they had normal or corrected-to-normal visual acuity, normal color vision, and had no clinical manifestations of mental or cognitive impairment or verbal or non-verbal learning disabilities. Exclusion criteria were the use of psychoactive medication, psychiatric or neurological complaints and drug or alcohol addiction.

The study was approved by the Bioethics Commission of Educational and Scientific Centre "Institute of Biology and Medicine", Taras Shevchenko National University of Kyiv (Conclusion from 15 August 2018, project title "Detrended fluctuation analysis of activation re-arrangement in EEG dynamics during cognitive workload"). Each subject signed written informed consent following the World Medical Association (WMA) declaration of Helsinki of 1975 ([http://www.wma.net/Data/2019\\_4\\_14\\_5\\_of\\_6\\_en/30publications/10policies/b3/](http://www.wma.net/Data/2019_4_14_5_of_6_en/30publications/10policies/b3/)), revised in 2008, the Declaration of Principles on Tolerance (28th session of the General Conference of UNESCO, Paris, 16 November 1995), the Convention for the protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine (Oviedo, 4 April 1997).

#### Apparatus/materials



Figure 1. Workflow of the analysis.

The entire analysis is summated in Figure 1. Neurocom monopolar EEG 23-channel system (Ukraine,

XAI-MEDICA) was used to obtain the recording. Silver/silver chloride electrodes were placed on the scalp at symmetrical anterior frontal (Fp1, Fp2), frontal (F3, F4, Fz, F7, F8), central (C3, C4, Cz) parietal (P3, P4, Pz), occipital (O1, O2), and temporal (T3, T4, T5, T6) recording sites as per the 10-20 International System of Electrode Placement. All electrodes were referenced to the interconnected ear reference electrodes. The inter-electrode impedance was below 5 k $\Omega$  and sample rate 500 Hz per channel.

Each recording consists of an artefact-free EEG segments of 180 s for resting state and 60 s for task state. 30 of the 66 initial participants were excluded from the database due to poor EEG quality by a board certified electro neurophysiologist, hence the final number of 36 subjects.

**Procedure.** Tasks in this study involved the serial subtraction of two numbers. Each trial started with the oral communication of the 4-digit (minuend) and 2-digit (subtrahend) numbers (e.g., 7456 and 45, 2343 and 48, etc.). Serial subtraction during 15 min is considered to be a psychosocial stress<sup>34</sup>. In this way, our study design required intensive cognitive activity from the subjects. Intensive mental load is accompanied by a change in the emotional background when the subject makes additional effort to resolve tasks, so one can talk about evoked emotions in this case. During EEG recording, the participants sat in a dark soundproof chamber, comfortably reclined in an armchair. Serial subtraction is a composite measure of auditory attention/ concentration, mental tracking, and computation. Hence to avoid all distractors, before the experiment is begun, participants were instructed to try to relax during the rest state. Then they were asked to count mentally without speaking or using finger movements, accurately and quickly, in the rhythm they had determined. After 3 min of adaptation to experimental conditions, EEG registration of the rest state with closed eyes was made (over the next 3 min). Then the participants performed a mental arithmetic task—serial subtraction—for 4 min.

Participants were interviewed about their strategies and experience after the experiment. These periods were selected since the task performance strategy is being formed simultaneously as the task is executed, and the emotional state of the participants is changing considerably due to intellectual overload.

**Data processing and analysis.** EEG Pre-processing. EEG data were first exported to EEGLAB<sup>35</sup> then analyzed mainly using MATLAB (R2018b). The length of rest state EEG signal is 180 seconds and the length of mental arithmetic EEG signal is 60 seconds. They were divided into 10-seconds epochs. However, the rest state EEG signal of subject O4 and subject 31 are shorter than other. Therefore, there are 851 epochs including 635 rest epochs and 216 task epochs.

High pass filter, low pass filter and band stop filter was applied in this dataset. 0.5Hz high pass filter was used to remove baseline. 45Hz low pass filter was used to remove high frequency noise. 50Hz notch pass filter to remove power frequency noise. The Independent Component Analysis(ICA) method was used to eliminate the artifacts (eyes, muscle and cardiac overlapping of the cardiac pulsation). The wave and power spectral density (PSD) of clean pre-processed data was shown in Figure 2. There are 19 channels, including FP1, FP2, F3, F4, F7, F8, T3, T4, C3, C4, T5, T6, P3, P4, O1, O2, Fz, Cz and Pz.

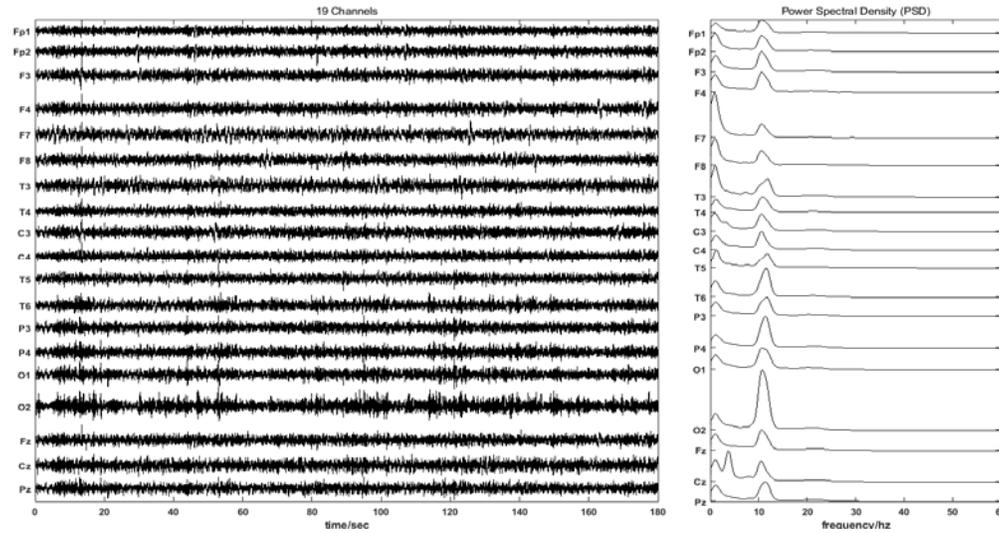


Figure 2. Processed signals of 19 channels with the estimated corresponding power spectral densities.

**Decomposition.** The EEG data was decomposed into several frequency bands (delta, theta, alpha, beta and low gamma) based on wavelet packet transform with wavelet basis 'sym7' and level 7 using the Wavelet Toolbox<sup>TM</sup> software carried out in MATLAB (R2018b) (MathWorks Inc., Natick, MA). According to the following formula<sup>36</sup>(Daubechies), the decomposition was carried out and the results were obtained as shown in Table 2. and Figure 3.

—  
—

Frequency Range	Name
0.5000-3.9063Hz	delta
5.8594-7.8125Hz	theta
9.7656-13.6719Hz	alpha
15.6250-29.2969Hz	beta
31.2500-45.0000Hz	Low gamma

Table 2. Decomposition of the EEG data: Frequency range of corresponding bands.

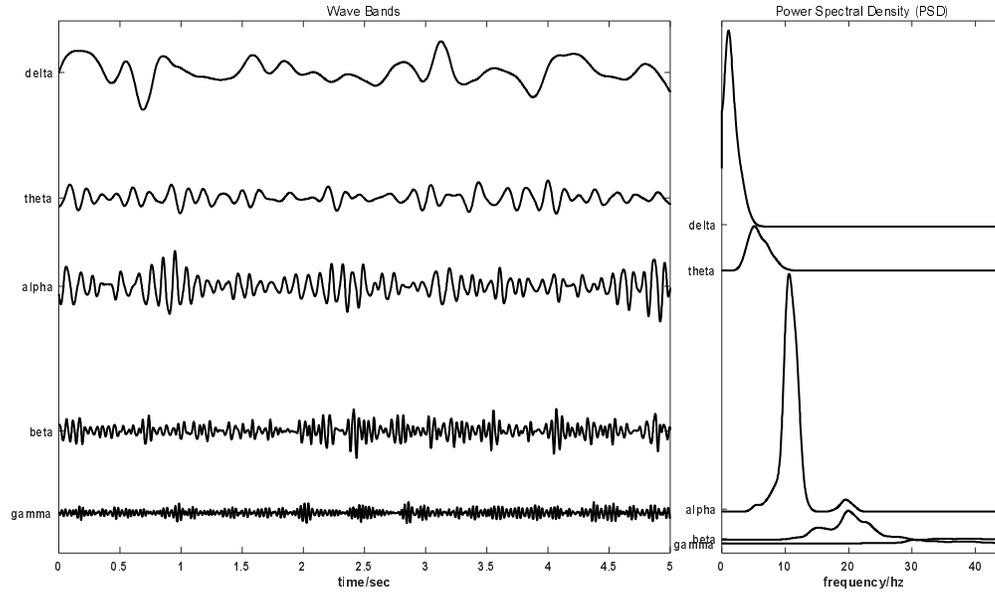


Figure 3. Decomposition of the EEG data into different frequency bands.

**Feature Extraction.** In this study, the power spectral density (PSD), sample entropy (SE), fuzzy entropy (FE) and fractal dimension (FD) were used on five frequency bands (delta, theta, alpha, beta and low gamma) for rest and task states.

**Power spectral density.** *Welch's method*<sup>37</sup> was used for estimating power spectra. It is carried out by dividing the time signal into successive blocks, forming the periodogram for each block, and averaging.

The  $m$ th windowed, zero-padded frame from the signal  $\mathbf{x}$  is denoted by

$$x_m(n) \triangleq w(n)x(n + mR), \quad n = 0, 1, \dots, M - 1, \quad m = 0, 1, \dots, K - 1 \quad (1)$$

where  $R$  is defined as the window *hop size*, and let  $K$  denote the number of available frames.

Then the periodogram of the  $m$ th block is given by

$$P_{x_m, M}(\omega_k) = \frac{1}{M} |\text{FFT}_{N, k}(x_m)|^2 \triangleq \frac{1}{M} \left| \sum_{n=0}^{N-1} x_m(n) e^{-j2\pi n k / N} \right|^2$$

as before, and the Welch estimate of the power spectral density is given by

$$S_x^W(\omega_k) \triangleq \frac{1}{K} \sum_{m=0}^{K-1} P_{x_m, M}(\omega_k). \quad (3)$$

In other words, it's just an average of periodograms across time. When  $\mathbf{w}(\mathbf{n})$  is the rectangular window, the periodograms are formed from non-overlapping successive blocks of data.

**Sample entropy.** Sample entropy is a modification of approximate entropy.<sup>38</sup> We have an  $\text{imf}_N = \mathbf{x}_i = (x_1, x_2, x_3, \dots, x_N)$   $i=1, 2, 3, \dots, N$  and use a time interval to reconstruct

series  $X_m(i) = (x_i, x_{i+1}, x_{i+2}, \dots, x_{i+m-1})$   $i = 1, 2, 3, \dots, N-m+1$ . The length of sequence is  $m$ .

The distance function of two sequences is  $d[X_m(i), X_m(j)]$ . For a given embedding dimension  $m$ , tolerance  $r$  and number of data points  $N$ , SE is expressed as:

$$SE(r, m, N) = -\log A_{m+1}(r) B_m(r), \quad (4)$$

where,  $B_m(r)$  is the number of template vector pairs having  $d[X_m(i), X_m(j)] < r$  and represents the similarity between two sequences of length  $m$ ,  $A_{m+1}(r)$  is the number of template vector pairs having  $d[X_{m+1}(i), X_{m+1}(j)] < r$  and represents the similarity between two sequences of length  $m+1$ .

**Fuzzy Entropy.** Fuzzy entropy is the entropy of a fuzzy set, which loosely represents the information of uncertainty<sup>39,40</sup>

For an  $imf_N = x_i = (x_1, x_2, x_3, \dots, x_N)$   $i = 1, 2, 3, \dots, N$ , we reconstruct series with length  $m$ :

$$X_m(i) = (x_i, x_{i+1}, x_{i+2}, \dots, x_{i+m-1}) - 1 \sum_{j=0}^{m-1} x_{(i+j)} \quad i = 1, 2, 3, \dots, N-m+1, \quad (5)$$

The distance function  $d_{mij}$  of two sequences is  $d[X_m(i), X_m(j)]$ . Given  $n$  and  $r$ , calculate the similarity degree  $D_{mij}$  through a fuzzy function  $\mu(d_{mij}, n, r)$ .

$$\mu(d_{mij}, n, r) = \exp(-(d_{mij})^n / r) \quad (6)$$

Define the function  $\emptyset_m$  as

$$\emptyset_m(n, r) = 1/N - m \sum_{i=1}^{N-m} (1/N - m - 1 \sum_{j=1, j \neq i}^{N-m} D_{mij}), \quad (7)$$

$\emptyset_{m+1}(n, r)$  is got similarly.

Lastly, the FuzzyEn( $m, n, r$ ) of the series is shown in equation (8).

$$\text{FuzzyEn}(m, n, r) = \ln \emptyset_m(n, r) - \ln \ln \emptyset_{m+1}(n, r) \quad (8)$$

**Fractal dimension.** Fractal dimension as one of the chaos measurement tools is widely utilized in the determination of chaotic behavior of signals. We have used the box-counting method for the FD calculation. In this method, a network of squares is created on the contour, and the number of squares in this network that includes a part of the curve is calculated. The size of the squares changes, then, the occupied squares are counted again. A point is obtained by calculating the logarithm of the number of counted squares and the logarithm for the reduction coefficient. New points are obtained by reducing the size of the network, which is connected in a curve. Calculating the slope of this curve

indicates the FD. Ahmadlou et al compared two main algorithms for calculating fractal dimension from EEG. HFD was shown to provide better discrimination (91.3%) compared to Katz's Fractal Dimension<sup>41</sup>. In this paper we have used the Higuchi's FD(Higuchi) due to its high level of accuracy<sup>42</sup>. Equation (9) describes how the self-similarity dimension (D), the number of self-similar pieces (a) and the reduction factor (1/S) are related together.

$$a \propto 1^D \quad (9)$$

$$\text{where } D \text{ is defined as: } D = \log(a) / \log(1/S) \quad (10)$$

D can be calculated by estimating the slope of the approximated line for the plot of log (a) vs. log (1/s).

**Statistical software.** All analyses were carried out in MATLAB (R2018b) (MathWorks Inc., Natick, MA). All reported statistical tests in the present study are two-sided tests wherever applicable. The significant difference was defined as the p-value < 0.05. Parametric test ANOVA1, which is a one-way analysis of variance, was used for the data that satisfied the assumptions. Otherwise, a nonparametric Kruskal-Wallis test was used.

### References

1. Sweller, J., van Merriënboer, J. J. G. & Paas, F. G. W. C. Cognitive Architecture and Instructional Design. *Educational Psychology Review* **10**, 251–296 (1998).
2. Murphy, G. L. & Wright, J. C. Changes in conceptual structure with expertise: Differences between real-world experts and novices. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **10**, 144–155 (1984).
3. Moreno, R. & Mayer, R. E. Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology* **91**, 358–368 (1999).
4. Paas, F. G. Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology* **84**, 429–434 (1992).
5. Mousavi, S. Y., Low, R. & Sweller, J. Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology* **87**, 319–334 (1995).
6. Biernat, M., Kobrynowicz, D. & Weber, D. L. Stereotypes and Shifting Standards: Some Paradoxical Effects of Cognitive Load. *J Appl Social Psychol* **33**, 2060–2079 (2003).
7. van Gog, T. & Paas, F. Instructional Efficiency: Revisiting the Original Construct in Educational Research. *Educational Psychologist* **43**, 16–26 (2008).
8. Palinko, O., Kun, A. L., Shyrov, A. & Heeman, P. Estimating cognitive load using remote eye tracking in a driving simulator. in *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10* 141 (ACM Press, 2010). doi:10.1145/1743666.1743701.
9. Barua, S., Ahmed, M. U. & Begum, S. Classifying Drivers' Cognitive Load Using EEG Signals. *Stud Health Technol Inform* **237**, 99–106 (2017).
10. Pope, A. T., Bogart, E. H. & Bartolome, D. S. Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology* **40**, 187–195 (1995).
11. Mills, C. et al. Put your thinking cap on: detecting cognitive load using EEG during learning. in *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* 80–89 (ACM, 2017). doi:10.1145/3027385.3027431.
12. McDonald, N. J. & Soussou, W. QUASAR's QStates cognitive gauge performance in the cognitive state assessment competition 2011. in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 6542–6546 (IEEE, 2011). doi:10.1109/IEMBS.2011.6091614.
13. Goldberger, A. L., Peng, C.-K. & Lipsitz, L. A. What is physiologic complexity and how does it change with aging and disease? *Neurobiology of Aging* **23**, 23–26 (2002).

14. Lamberts, J. *et al.* Correlation Dimension of the Human Electroencephalogram Corresponds with Cognitive Load. *Neuropsychobiology* **41**, 149–153 (2000).
15. Natarajan, K., Acharya U, R., Alias, F., Tiboleng, T. & Puthusserypady, S. K. Nonlinear analysis of EEG signals at different mental states. *BioMed Eng OnLine* **3**, 7 (2004).
16. Boon, M. Y., Suttle, C. M., Henry, B. I. & Dain, S. J. Dynamics of chromatic visual system processing differ in complexity between children and adults. *Journal of Vision* **9**, 22–22 (2009).
17. Lutzenberger, W., Elbert, T., Birbaumer, N., Ray, W. J. & Schupp, H. The scalp distribution of the fractal dimension of the EEG and its variation with mental tasks. *Brain Topogr* **5**, 27–34 (1992).
18. Cabañero *et al.* Analysis of Cognitive Load Using EEG when Interacting with Mobile Devices. *Proceedings* **31**, 70 (2019).
19. P Harne, B. Higuchi Fractal Dimension Analysis of EEG Signal Before and After OM Chanting to Observe Overall Effect on Brain. *IJECE* **4**, 585–592 (2014).
20. Zarjam, P., Epps, J., Lovell, N. H. & Chen, F. Characterization of memory load in an arithmetic task using non-linear analysis of EEG signals. *Conf Proc IEEE Eng Med Biol Soc* **2012**, 3519–3522 (2012).
21. Mazher, M., Abd Aziz, A., Malik, A. S. & Ullah Amin, H. An EEG-Based Cognitive Load Assessment in Multimedia Learning Using Feature Extraction and Partial Directed Coherence. *IEEE Access* **5**, 14819–14829 (2017).
22. Li, X., Jiang, Y., Hong, J., Dong, Y. & Yao, L. ESTIMATION OF COGNITIVE WORKLOAD BY APPROXIMATE ENTROPY OF EEG. *J. Mech. Med. Biol.* **16**, 1650077 (2016).
23. Geller, A. S. *et al.* Eye closure causes widespread low-frequency power increase and focal gamma attenuation in the human electrocorticogram. *Clin Neurophysiol* **125**, 1764–1773 (2014).
24. Achirul Nanda, M., Boro Seminar, K., Nandika, D. & Maddu, A. A Comparison Study of Kernel Functions in the Support Vector Machine and Its Application for Termite Detection. *Information* **9**, 5 (2018).
25. Klimesch, W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews* **29**, 169–195 (1999).
26. Lagopoulos, J. *et al.* Increased Theta and Alpha EEG Activity During Nondirective Meditation. *The Journal of Alternative and Complementary Medicine* **15**, 1187–1192 (2009).
27. Gladwin, T. E. & de Jong, R. Bursts of occipital theta and alpha amplitude preceding alternation and repetition trials in a task-switching experiment. *Biological Psychology* **68**, 309–329 (2005).
28. Molnár, M. *et al.* [Psychophysiologic and clinical aspects of EEG synchronization related to cognitive processes]. *Idegyogy Sz* **58**, 393–401 (2005).
29. Wang, Q. & Sourina, O. Real-Time Mental Arithmetic Task Recognition From EEG Signals. *IEEE Trans. Neural Syst. Rehabil. Eng.* **21**, 225–232 (2013).
30. Zafar, R. *et al.* EEG spectral analysis during complex cognitive task at occipital. in *2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES)* 907–910 (IEEE, 2014). doi:10.1109/IECBES.2014.7047643.
31. Attallah, O. An Effective Mental Stress State Detection and Evaluation System Using Minimum Number of Frontal Brain Electrodes. *Diagnostics (Basel)* **10**, (2020).
32. Zyma, I. *et al.* Electroencephalograms during Mental Arithmetic Task Performance. *Data* **4**, 14 (2019).
33. Goldberger, A. L. *et al.* PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* **101**, (2000).
34. Noto, Y., Sato, T., Kudo, M., Kurata, K. & Hirota, K. The Relationship Between Salivary Biomarkers

- and State-Trait Anxiety Inventory Score Under Mental Arithmetic Stress: A Pilot Study.: *Anesthesia & Analgesia* 1873–1876 (2005) doi:10.1213/01.ANE.0000184196.60838.8D.
35. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* **134**, 9–21 (2004).
  36. Daubechies, I. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.* **41**, 909–996 (1988).
  37. Welch, P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* **15**, 70–73 (1967).
  38. Richman, J. S. & Moorman, J. R. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology* **278**, H2039–H2049 (2000).
  39. Weiting Chen, Zhizhong Wang, Hongbo Xie & Wangxin Yu. Characterization of Surface EMG Signal Based on Fuzzy Entropy. *IEEE Trans. Neural Syst. Rehabil. Eng.* **15**, 266–272 (2007).
  40. Azami, H., Fernández, A. & Escudero, J. Refined multiscale fuzzy entropy based on standard deviation for biomedical signal analysis. *Med Biol Eng Comput* **55**, 2037–2052 (2017).
  41. Ahmadlou, M., Adeli, H. & Adeli, A. Fractality and a Wavelet-Chaos-Neural Network Methodology for EEG-Based Diagnosis of Autistic Spectrum Disorder: *Journal of Clinical Neurophysiology* **27**, 328–333 (2010).
  42. Affinito, M., Carrozzi, M., Accardo, A. & Bouquet, F. Use of the fractal dimension for the analysis of electroencephalographic time series. *Biological Cybernetics* **77**, 339–350 (1997).