EEG-Based Identification of Latent Emotional Disorder Using the Machine Learning Approach

Yaling Deng[1], Fan Wu[1#], Lei Du[2], Renlai Zhou[3*], Lihong Cao[1]

1. Neuroscience and Intelligent Media Institute, Communication University of China, China

2. National Engineering Laboratory, Traffic Management Research Institute of the Ministry of Public Security, China

3. Department of Psychology, School of Social and Behavior Sciences, Nanjing University, China

yalingdeng@cuc.edu.cn, wufan9405@gmail.com, dlbupt@163.com, rlzhou@nju.edu.cn, Lihong.Cao@cuc.edu.cn

Equal as first author.

* Corresponding author: Renlai Zhou, Department of Psychology, Nanjing University, Xianlin Avenue 163, Qixia District, Nanjing 210023, Jiangsu Province, China. E-mail: rlzhou@nju.edu.cn

Abstract—**Emotion influences our daily life to a large extent, especially for those who are undergoing bad mood and have high risk for emotional disorders. It is hard to recognize them, but very important so that we can provide intervention before them getting worse. This study used EEG signals to recognize who has high risk for emotional disorders instead of emotion type only. The proposed machine learning method combined the features of multiple cortex areas and frequency bands to find the high risky group for emotional disorders through a kernel SVM classifier. It achieved the accuracy of 95.20%, with all cortex areas and all frequency bands. Results showed that the frontal cortex, central cortex and temporal cortex have a primary influence on identifying emotional disorder and can be used for the reference information for professional diagnose.**

Keywords—emotional disorder; EEG; machine learning

I. Introduction

Anxiety, fear, anger, sadness, disgust and other kinds of negative emotions influence our life just as what the positive emotions does every day, while in the opposite way. It's believed that without proper intervention in time, negative emotions may lead to emotional disorders which can even make people to hurt themselves [1], [2]. Indeed, it has been found that the common emotional disorders' incidence rate such as depression is almost 10% in lifetime [3]. These emotional disorders have negative effects on individuals, families, even the whole society [2]. Actually, before the formation of major depression or anxiety disorder, there are some symptoms that should not be ignored. It's much more meaningful to identify who has high risk for emotional disorders than just recognizing who has been diagnosed from the mental health perspective.

If the target group can be found at the right time, some serious emotional disorders may be prevented with appropriate intervention. It raised an important research topic that how can we identify the high risky group before they really get it? For this question, the professional psychotherapist can be the best choice. However, the diagnose procedure of emotional disorder made by them always needs a rather long time which makes the high risky group missed the best time for treatment. So as the much more convenient way, neuroimaging technology has attracted many researchers in related field. Because of the noninvasive, fast, and low-cost advantages, functional neuroimaging technique - electroencephalography (EEG) has become the preferred method in studying brain activities. No matter what kinds of characters the person is, the brain activities during emotional experiences would reveal one's inner world which can be utilized on some levels. In fact, many researchers have found some common characteristics associated with the brain activity for the group of those who were in bad mood and had high risk for emotional disorders. For example, compared to healthy controls, the anxiety individuals showed stronger right frontal brain activity [4], depressive individuals showed less left frontal brain activity [5].

Besides these spatial features, other spectrum features of brain activities should not be ignored also.

The goal of finding the target group assisted by selected features mentioned above can be seen as a classification problem. And if accuracy of classifier is high enough, the time of diagnose for professional clinical psychotherapists can be reduced as wish. The present study was inspired by this thought. With the assistance of machine learning theory, the emotional state of person can be acquired through proper signal processing on EEG signals [6]. Lin et al. [7] applied support vector machine (SVM) to recognize four kinds of emotion (joy, anger, sadness, and pleasure), and the accuracy achieved 82.29%. Jatupaiboon et al. [8] argued that frontal cortex area of channels and a high frequency band of gamma wave can reach a better accuracy with SVM on binary emotion classification. Shahabi et al. [9] performed three sets of experiments (joyful and neutral, joyful and depressed, familiar and unfamiliar music), and the accuracy of machine learning classifier was 93.7%, 80.83%, 83.4% respectively. Otherwise, Kroupi et al. [10] used the linear discrimination analysis (LDA) to classify the pleasant, unpleasant, and neutral emotion induced by smell, and the accuracy reached 57% in pleasant, 100% in unpleasant, and 42% in neutral emotion. Mohammadi et al. [11] compared the k-nearest neighbors (KNN) with SVM to recognize the arousal and valence of emotion, which reached to the performance of 84.05% and 86.75% for arousal and valence.

Though many studies have used the machine learning approach to recognize specific emotion, few have focused on the identification of whether someone has high risk for emotional disorders. Very recently, Osuch et al. [12] used functional magnetic resonance (fMRI) to predict mood disorder, which showed a very promising way. However, fMRI is much more expensive and less easy to use than EEG, so this work focused on finding the high risky group for emotional disorder instead of classifying the emotion type using EEG signals. It combined both spatial features and frequency features to feed a SVM classifier.

II. Method

A. Participants

Volunteers were recruited at Beijing Normal University through advertisements. A total of 110 applied for taking part in this study. We wanted to discriminate who has high or low risk for emotional disorder. So professional clinical psychotherapists conducted a series of tests on all the participants to make sure their current mood state. Clinical psychotherapists held interviews for each applier. During the interview, demographic information was collected, and psychotherapists asked participants several questions to identify whether they met the criteria of this study. Those who had no personal history of diagnosed psychiatric disorders, no substance abuse during the previous 6 months, no neurological problems, and no use of mood-altering medication (such as amphetamines, coffee, alcohol, and tea are not included here) were selected for this study.

Then psychotherapists tested the selected participators in depression [13] and anxiety (Beck Anxiety Inventory (BAI)), level [14]. Participants were also asked to complete the Eysenck personality questionnaire [15], because many studies have found that the personality, especially neuroticism, is an important predictive factor for emotional disorder [16], [17]. According to the professional clinical psychotherapists' diagnose, those who were in high level of depression or anxiety state were divided into high risk group for emotional disorder. While, those who were in quite low level of depression or anxiety state were sorted into low risk group correspondingly. Finally, 31 participants (9 males and 22 females) were arranged in high risk group and another 31 participants (9 males and 22 females) were sorted in low risk group respectively. The information of all the 62 participants' initial test results was shown in Table 1.

Independent sample T test showed that the high risky group and the low risk group only had significant differences in depression level, anxiety level and neuroticism (see Table 1, t and p value). It is important to note that all the participants in this study did not meet the diagnostic criteria for depression, anxiety or other emotional disorders. The participants in high

TABLE I. DEMOGRAPHIC INFORMATION AND SCORES OF BDI, BAI, EPQ OF THE TWO GROUPS

	High Risk Group (N=31, N_{male}=9)	**Low Risk Group (N=31, N_{male}=9)**	***t***	***p***
Age	21.387±1.856	21.032±1.941	0.736	0.465
BDI score	11.548±6.060	5.032±4.078	4.967	0.000
BAI score	32.387±8.192	26.710±6.972	2.938	0.005
EPQ- Psychoticism (P)	2.355±1.723	2.710±1.811	0.790	0.432
EPQ- Extraversion (E)	7.810±3.449	8.258±2.695	0.574	0.568
EPQ- Neuroticism (N)	9.677±1.558	2.129±0.922	23.223	0.000
EPQ-Social Desirability (L)	3.613±1.944	4.516±2.554	1.567	0.122

risk group had much higher level of depression and anxiety level than the low risk group, and they had much higher risk for emotional disorder according to many previous studies [16], [17], [18]. The experimental procedure was approved by the Institutional Ethics Review Board of the School of Psychology at Beijing Normal University. All the participants provided signed informed consent before participating and got some money as remuneration.

B. Materials and Procedure

After the interview by clinical psychotherapists, those who met the criteria of this study were asked to come to the lab in turns. First the experimenters told them the information of this study and explain the procedure to make them totally understand. There were eight types of emotional film clips: fear, disgust, anger, sadness, neutrality, surprise, amusement, and pleasure. All the participants were arranged to view 16 film clips (2 film clips for each emotion type) on a 14-inch screen. The experiment procedure was controlled by the E-Prime 2.0 software. Participants were told to watch film clips carefully. However, if they felt very uncomfortable to watch they could shut eyes, look away or stop watching. Film clips were showed as scheduled. Before watching each clip, a blank screen for 30 seconds was displayed to tell the participants clear their minds of all thoughts, feelings, and memories. During the film's showing period, participants were told to not make any strong movements to ensure the EEG data's quality. After each film clip, participants were asked to rate the *Valence, Arousal, Pleasure, Disgust, Fear, Sadness, Surprise, Anger, Funny*. The emotional film clips were selected from the Standardized Emotional Film Dataset for Asian Culture [19]. The order of the film clips was organized as [19] too.

C. EEG Recording and Data Analysis

The EEG data of all the participants were collected when they watched the film clips. And it was recorded by the NeuroScan recording system with 40 Ag/AgCl electrodes positioned according to the International 10-20 system. Data were referenced to the left mastoids online. The ground electrode was located at the midpoint of the line between FPz and Fz. The vertical electrooculogram (EOG) was recorded with electrodes placed above and below the left eye respectively, whereas the horizontal EOG was recorded with electrodes placed on the left and the right orbital rim. The sampling rate was 1000 Hz, DC sampling with band-pass filter at DC-200 Hz online, and the low-pass filter offline was 30 Hz (24 dB/oct). Electrode impedance was generally below 10 kΩ. The EEG signals were recorded with the Numps amplifier, and all data were analyzed offline.

Offline analysis was performed with Scan 4.5 software. Data were referenced to the bilateral mastoids offline. The ocular artifacts were rejected first. And the outcome was checked visually to remove any epochs that contained other technical artifacts in all channel. The data were filtered by a hamming windowed-sinc FIR filter with zero phase shift between 1-30 Hz. For each film clip, the power of delta (0.5-3Hz), theta (4-7Hz), alpha (8-13Hz), and beta (14-30Hz) bands at each electrode (see Fig.1.) were calculated.

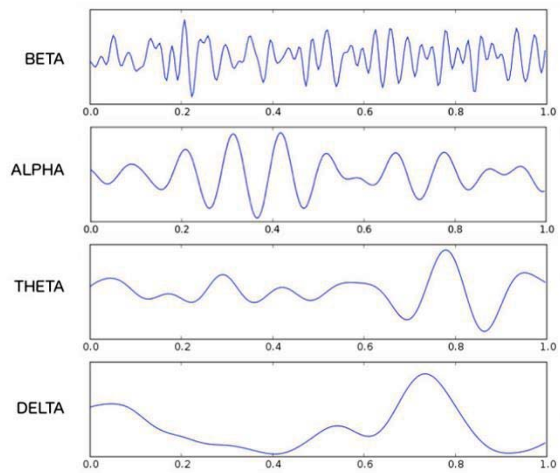

Fig. 1. The four bands of brain activity (cited from Alarcao et al. [6])

D. Machine learning approach

The support vector machine (SVM) was chosen to recognize the low and high risky groups. The core of SVM can be concluded as to construct a hyperplane to maximize the distance of the positive and negative samples. Due to the characteristic of convex, it can reach the global optimum and minimizes the cost of misclassification at the same time. Because of the limitations of the data, the kernel SVM was chosen for classifier instead of linear SVM. It projects the feature of low dimensionality to high dimensionality implicitly, and makes feature to be disentangled in high dimensionality. In order to get more compelling results, we adapted 5-fold cross-validation which separated data to five segments: four for training and one for testing.

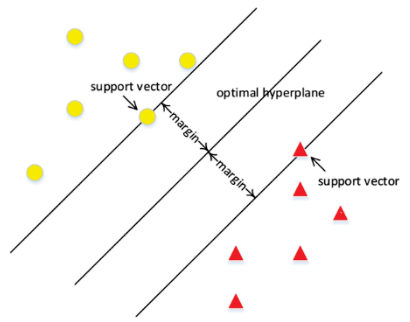

Fig. 2. Support vector, margin, and optimal hyperplane (cited from Jatupaiboon et al. [8])

As Fig.2. shows, the hyperplane in kernel SVM can be described as follows:

$$f(x) = w^T \varphi(x) + b \tag{1}$$

And the radial basis function is:

$$k(x_i, x_j) = \exp\left(-\frac{\left\|x_i - x_j\right\|^2}{2\sigma^2}\right) \tag{2}$$

Where, σ is the width of kernel function, usually, $\frac{1}{2\sigma^2}$ is called as gamma factor. It assumes that all samples are separated, and subjects to the inequation as follow:

$$y_i\left(w^T x_i + b\right) \geq 1 \tag{3}$$

In practice, not all samples can be separated by hyperplane precisely. In order to reduce the influence of this special samples undesired, the approach of soft margin is introduced to SVM. It allows the samples to classify the opposite category in some degree.

$$y_i\left(w^T x_i + b\right) \geq 1 - \xi_i \tag{4}$$

Where, ξ_i is slack variable, representing the degree of every sample deviates from the accurate category. In the phase of optimization, C will be introduced to control to the degree of fitting.

In our experiments, the penalty factor C is set to 2000 and gamma is set to $\frac{1}{N_f}$ (N_f represent the feature of dimensionality).

E. Experiment I: comparison between different frequency bands

Researchers have found that emotional disorders showed abnormal brain activity especially on alpha band [4], [5], so we explored the influence of different frequency bands to distinguish low risk and high risky group. The power of alpha, beta, theta, delta bands (see Fig. 1) and all four bands combined were selected to classify the two different categories as frequency features.

F. Experiment II: comparison between different brain regions

A large number of researches described that the frontal cortex, central cortex and temporal cortex have a huge impact on emotion [6], [20]. Psychologists also found that individuals with emotional disorders showed abnormal brain activity especially in frontal and central cortexes [4], [5]. We wondered whether the different cortex contributed to the results of classification differently. We first compared different brain regions: frontal cortex, central cortex, parietal cortex, occipital cortex, and temporal cortex (see Fig. 3.).

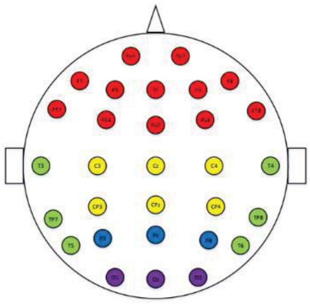

Fig. 3. Location of electrodes (different color represents different cortex, red is frontal cortex, yellow is central cortex, green is temporal cortex, blue is parietal cortex, and purple is occipital cortex.)

III. RESULTS

The accuracy of classification for different frequency bands and different brain regions were shown in Table 2. For different bands, alpha band showed relatively higher accuracy than other bands. As for different brain regions, the accuracy of frontal cortex was higher than central cortex, temporal cortex, occipital cortex and parietal cortex respectively (see Fig. 4.). If we only focused on single band and single brain region, the good performance could achieve in frontal cortex (74.91%) and central cortex (78.88%) at alpha band.

We also calculated the accuracy combining different brain regions and using multiple bands (four bands). The results showed great improvement than single band and single regions. When using all frequency bands, the performance could reach 88.64% in frontal cortex, 82.54% in central cortex and 83.27% in temporal cortex, which also had a competitive performance. When combining several brain regions, the accuracy showed a dramatic improvement, especially for the whole cortexes (95.20%). It was also important to find that the accuracy of frontal and central cortexes (F & C), as well as the frontal, central and temporal cortexes (F & C & T) achieved 93.17% and 93.64% respectively (see Fig. 5.). This was really very close to the whole brain accuracy of 95.20%. However, the combination of temporal, parietal and occipital (T & P & O) did not show this huge improvement.

TABLE II. THE ACCURACY OF CLASSIFICATION (*MEAN* ± *SD*)

	delta	**theta**	**alpha**	**beta**	**All (band)**
Frontal (F)	71.34±3.09	73.9±5.54	74.91±1.33	68.2±1.38	88.64±1.79
Central (C)	65.85±1.65	67.02±2.07	78.88±3.52	69.78±3.60	82.54±2.54
Temporal (T)	63.91±4.11	68.72±2.76	65.94±2.03	70.48±1.69	83.27±2.04
Occipital (O)	58.14±3.95	60.88±2.15	61.64±3.80	58.15±2.36	69.71±1.94
Parietal (P)	58.9±2.64	60.47±3.09	62.95±3.11	61.67±2.34	72.37±2.02
F & C	74.26±2.46	75.13±2.67	83.55±2.29	74.68±3.17	**93.17±2.08**
F & C & T	78.26±1.52	80.20±2.22	85.25±2.22	79.55±4.49	**93.64±3.42**
T & P & O	67.55±4.53	73.78±1.88	74.17±1.92	74.05±2.91	87.68±2.29
P & O	60.58±2.34	69.72±1.18	67.30±2.93	63.70±1.36	78.56±3.35
All (cortex)	77.75±2.17	82.29±2.07	86.50±2.83	80.87±2.41	**95.20±2.14**

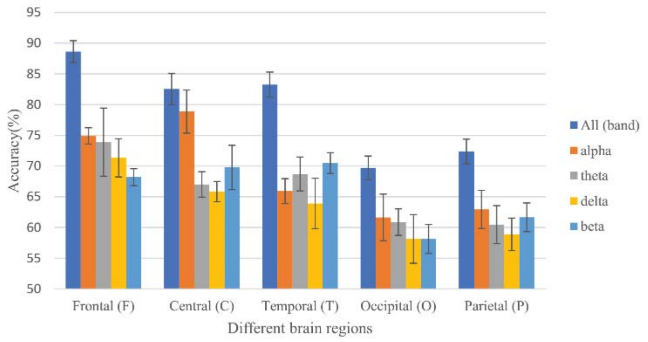

Fig. 4. The accuracy of classification of different frequency bands at different brain regions.

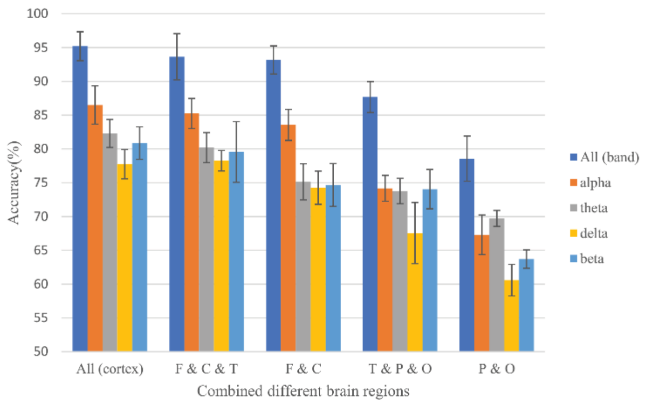

Fig. 5. The accuracy of classification of different frequency bands combining different brain regions.

IV. Discussion

With the rapid development of machine learning approach, many studies have used this approach to recognize specific emotion and got high identification accuracy. However, few have focused on the persons' inner mood state. To identify who has high risk for emotional disorder is meaningful for the mental health perspective. The present research focused on emotional disorder, instead of emotion itself. We hope emotional disorder can get more attention and our study can help to diagnose those who have a high risk for emotional disorder. The earlier to find out, the more we can do to prevent things getting worse. The present study used EEG signals and machine learning approach to identify the person who has high risk for emotional disorder. The accuracy reached 95.20%, which was as good as the fMRI results [12].

According to the results above, we found that using multi-regions data could reach a higher accuracy. The results also showed that alpha band contributed most to the accuracy than other three bands. Alpha band (8-13Hz) is generally thought to be related to brain activation, the lower power of alpha band, the more active processing of brain regions [21], so it was not surprise that alpha band was the most important. Besides, different brain regions showed different importance for recognizing emotional disorder. Frontal cortex, central cortex, and temporal cortex had the primary effect. This result was consistent with the findings of psychologists that frontal cortex and central cortex were important to emotional processing, and individuals who had emotional disorder often showed abnormal activity in these brain regions [4], [5].

Although our study achieved a good performance at 95.20%, it still existed some limitations. Firstly, we just classified two categories and did not divided the level of risk subtly. Secondly, we just used the SVM to classify different risk, instead of more methods. As for source of data, we did not explore different state's influence for classification. Maybe a good performance can be achieved in some specific states, such as joy, sadness. All of those are very important to study emotional disorder, and we will develop further study from this aspect.

V. Conclusion

In this paper, EEG signals were used to find who has high risk for emotional disorders. Both the special features in different brain regions and frequency features in different bands were tested to feed a kernel SVM classifier. And the performance of SVM classifier could achieve 95.20% by combining whole brain regions and all frequency bands. We found that multiple bands and regions had a greater contribution to result than single band and region under the current task. The frontal cortex, central cortex, and temporal cortex were three most important regions to classify emotional disorders and alpha band represented more sensitive than other bands. This approach can be used to help diagnose who has high risk for emotional disorders rapidly and precisely.

Acknowledgments

This work was funded by Communication University of China (Grant Nos. CUC18A001, CUC18A003-3, 3132017XNG1724, 2018CUCTJ031), the Key Project of

Philosophy and Social Science Research in Colleges and Universities in Jiangsu Province (Grant No.2015JDXM001), and NJU National Demonstration Base for Innovation & Entrepreneurship (Grant No. SCJD0406). We would like to express our gratitude for the support of these projects.

REFERENCES

[1] Lecrubier Y . The burden of depression and anxiety in general medicine[J]. Journal of Clinical Psychiatry, 2001, 62 Suppl 8(6):4.

[2] Sheikman M B . The burden of depression[J]. Nature, 2014, 515(7526):163.

[3] Riolo S A , Nguyen T A , Greden J F , et al. Prevalence of Depression by Race/Ethnicity: Findings From the National Health and Nutrition Examination Survey III[J]. American Journal of Public Health, 2005, 95(6):998-1000.

[4] Davidson R J, Marshall J R, Tomarken A J, et al. While a phobic waits: regional brain electrical and autonomic activity in social phobics during anticipation of public speaking.[J]. Biological Psychiatry, 2000, 47(2):85-95.

[5] Henriques J B, Davidson R J. Left frontal hypoactivation in depression.[J]. Journal of Abnormal Psychology, 1991, 100(4):535.

[6] Alarcao S M , Fonseca M J . Emotions Recognition Using EEG Signals: A Survey[J]. IEEE Transactions on Affective Computing, 2017:1-1.

[7] Lin Y P , Wang C H , Jung T P , et al. EEG-Based Emotion Recognition in Music Listening[J]. IEEE Transactions on Biomedical Engineering, 2010, 57(7):1798-1806.

[8] Jatupaiboon N, Pan-Ngum S, Israsena P. Emotion classification using minimal EEG channels and frequency bands[C]// International Joint Conference on Computer Science & Software Engineering. 2013.

[9] Shahabi H, Moghimi S. Toward automatic detection of brain responses to emotional music through analysis of EEG effective connectivity[M]. 2016.

[10] Kroupi E, Vesin J M, Ebrahimi T. Subject-Independent Odor Pleasantness Classification Using Brain and Peripheral Signals[J]. IEEE Transactions on Affective Computing, 2016, PP(99):422-434.

[11] Mohammadi Z, Frounchi J, Amiri M. Wavelet-based emotion recognition system using EEG signal[J]. Neural Computing & Applications, 2017, 28:1-6.

[12] Osuch E, Gao S, Wammes M, et al. Complexity in mood disorder diagnosis: fMRI connectivity networks predicted medication‐class of response in complex patients[J]. Acta Psychiatrica Scandinavica, 2018, 138(5): 472-482.

[13] Beck A T, Steer R A, Carbin M G. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation[J]. Clinical Psychology Review, 1988, 8(1):77-100.

[14] Beck A T , Epstein N , Brown G , et al. An inventory for measuring clinical anxiety: psychometric properties.[J]. Journal of Consulting & Clinical Psychology, 1988, 56(6):893-7.

[15] Qian M, Wu G, Zhu R, et al. Development of the revised Eysenck personality questionnaire short scale for Chinese (EPQ-RSC)[J]. Acta Psychologica Sinica, 2000, 32(03): 317-323.

[16] Lahey B B. Public health significance of neuroticism.[J]. American Psychologist, 2009, 64(4):241.

[17] Neuroticism and the brain: A quantitative meta-analysis of neuroimaging studies investigating emotion processing[J]. Neurosci Biobehav Rev, 2013, 37(8):1518-1529.

[18] Whalley H C , Sussmann J E , Romaniuk L , et al. Dysfunction of emotional brain systems in individuals at high risk of mood disorder with depression and predictive features prior to illness[J]. Psychological Medicine, 2015, 45(06):1207-1218.

[19] Deng Y, Yang M, Zhou R. A New Standardized Emotional Film Database for Asian Culture.[J]. Frontiers in Psychology, 2017, 8:1941.

[20] Huang, Guan, Ang K K , et al. Asymmetric Spatial Pattern for EEG-based emotion detection[C]// International Joint Conference on Neural Networks. IEEE, 2012.

[21] Davidson R J, Chapman J P, Chapman L J, et al. Asymmetrical brain electrical activity discriminates between psychometrically‐matched verbal and spatial cognitive tasks[J]. Psychophysiology, 1990, 27(5): 528-543.