Quantifying Cognitive Load in Wayfinding Information Review using EEG

Yanchao WANG[1], Qi WANG, Ph.D.[2], Jing DU, Ph.D.[3], Yingzi LIN, Ph.D.[4]

[1] Urban Informatics and Resilience Lab, Department of Civil and Environmental Engineering, Northeastern University, Boston, .360 Huntington Ave; PH (857) 334-9873; email: wang.yanch@husky.neu.edu

[2] Urban Informatics and Resilience Lab, Department of Civil and Environmental Engineering, Northeastern University, Boston, 360 Huntington Ave; PH (617) 373-7984; email: q.wang@northeastern.edu

[3] Engineering School of Sustainable Infrastructure & Environment, Herbert Wertheim College of Engineering, University of Florida, 460F Weil Hall, Gainesville; PH (352) 294-6619; email: eric.du@essie.ufl.edu

[4] Department of Mechanical and Industrial Engineering, Northeastern University, Boston, 360 Huntington Ave; PH (617) 373-8610; email: yi.lin@northeastern.edu

ABSTRACT

Driven by the increasing complexity of built environments, firefighters are often exposed to extensive wayfinding information which could cause high cognitive load and ineffective or even dangerous decision making. To reduce injuries and fatal incidents in firefighters' line of duty, this study aims at measuring the cognitive load and identifying the source of such cognitive overload in wayfinding information review. We developed a Sternberg Test to induce cognitive load on participants pertaining to working memory development, where participants were required to memorize colors, letters, numbers, directions, icons, words, and letter combinations that are relevant to wayfinding tasks. We used an Electroencephalogram (EEG) device to monitor neural activities especially in frontal, parietal, and occipital areas of brain. The fast Fourier transformation (FFT) was applied to separate the sub-band energy. The speed of response in Sternberg Test and the EEG signals were compared to show the coherence between the results of the two methods in representing the cognitive load in the review test. Results indicate that the cognitive load arises from diverse information can be measured to help customize wayfinding information for controlled cognitive load of firefighters in wayfinding tasks.

INTRODUCTION

Firefighters work in extreme situations and operate highly psychological demanding tasks (Roja et al. 2009; Henderson et al. 2016). According to the report from National Fire Protection Association, 58,835 firefighter injuries occurred in the line of duty and 42 percent of firefighter injuries occurred at fireground in 2017 (Evarts and Molis 2018). Among all the accidents, firefighter disorientation is the root cause of firefighter fatalities defined by the National Institute of Occupational Safety and Health (Brennan 2011). In the investigation for disorientation cases from 1979 to 2001, heavy smoke which caused Prolonged Zero Visibility Conditions was developed in 100%

of the cases (Mora 2003). Therefore, memorization becomes one of the major resources for firefighters to perform wayfinding on low visibility fireground.

However, memorization before and after entering the fireground could both be challenging for the firefighters. Usually, firefighters have to take in a large amount of information all at once and make vital decisions (Davies 2015). This process would introduce tremendous cognitive load. Moreover, the severe pressure arises from limited time, life-threatening environment could further elevate their cognitive load and cause mental overload eventually (Galy 2012). It is critical to monitor firefighters' cognitive load, help them better process demanding information, and thus, improve their performance on wayfinding and rescue tasks.

Despite the criticality, there is a limited understanding on the connection between the types of information needed by firefighters and the cognitive overload on parts of human brain caused by them. We would not be able to identify the causes of cognitive overload without such an in-depth understanding. To narrow the knowledge gap between sources and cognitive overload, cognitive load introduced by information which firefighters frequently exposed to is measured in this research in order to, in the future, understand the relation between cognitive load and various types of information for individuals.

LITERATURE REVIEW

In this research we used Electroencephalography (EEG) as the main neuroimaging tool. EEG is a widely used neuroimaging technique which measure the electrical fields generated by the brain. EEG has proven to be an influential tool in cognitive monitoring (Antonenko et al. 2010). Due to the rapid propagation speed of electric fields, EEG has distinguished time domain resolution in imaging large-scale brain activity (He et al. 2018). However, three disadvantages have limited further application of EEG. Firstly, brain activity signals are often buried by noises from environment and body activities in raw EEG signals, which are called "artifacts" (Roy 2019). Various processing methods must be used to acquire clean brain activity signals. Secondly, high inter-subject variability significantly limits the widely application of EEG devices. The variability originates from physiological differences between individuals. Finally, it is a non-stationary signal (Gramfort et al. 2013), which means data collected on the same subject at different time might hard to be generalized. Despite the existence of three prime drawbacks mentioned above, considerable researches have been conducted on measuring and classifying cognitive using EEG signals.

Around 1999, researchers discovered that the electrical activity in the brain have four distinct rhythms separated in frequency, including delta waves, theta waves, alpha waves, and beta waves (Başar 2012). In the research from Berka et al. (Berka et al. 2007), absolute and relative power spectrum were used as the metric for mental workload. A linear discriminant function analysis was applied to correlates the power spectrum to the mental workload assessed by both subjective and objective performance metrics. The research validated the feasibility of using power spectrum as the metric for mental workload by using other performance metrics. But the metric is invented to describe the cognitive load for the whole brain. In 2009, Holm et al. presented a two-channel EEG-based index, theta Fz / alpha Pz ratio. It was later be widely used in generally measuring subject's cognitive load (Dan 2017). But the metric is still measuring the cognitive load for the whole brain, which is not suitable for identification purpose. The Theta to Alpha Ratio (TAR) for the individual channel was also invented to assess the cognitive load for individual channels (Bian et al. 2014). After the flourishing of machine learning techniques, spatial information becomes more and more important for EEG data. Autoencoder was applied to evaluate the binary cognitive task load level (low versus high) for single channel signal (Yin and Zhang 2016). Multi-channel EEG signals were used as the input to the Convolutional Neural Network (CNN) (Zhang and Wang 2017), Recurrent Neural Network (RNN) (Hefron et al. 2017), and the Convolutional Recurrent Neural Network (Hefron et al. 2018). The impact on model performance from individual differences were decreased using Neural Networks and the sufficient diversity of individuals data could largely overcome the non-stationarity of EEG signals. However, the output for Neural Networks are in low resolution and the cognitive load generated by diverse types of information are yet to be identified.

Thus, a method to generate cognitive load from various types of information, an EEG metric, and a classifier are to be found to relate the cognitive load and the type of source information. In this research, the method and the metric will be selected and validated that cognitive load is generated, and the metric is capable to identify the load.

EXPERIMENT

Cognitive load is the load imposed on the cognitive system when human is performing a task (Sweller 1988). It represents the used amount of working memory resources in cognitive psychology. As a result, Sternberg Test (Sternberg 1969), a well-proven working memory test, was applied in this research to simulate task-driven cognitive load scenarios. Because the short-term memory is part of the working memory, subject's cognitive load would be increased throughout the test (Sweller 2011). Then, EEG was used to measure the electrical activity produced by the brain to assess the cognitive load on subjects during the test.

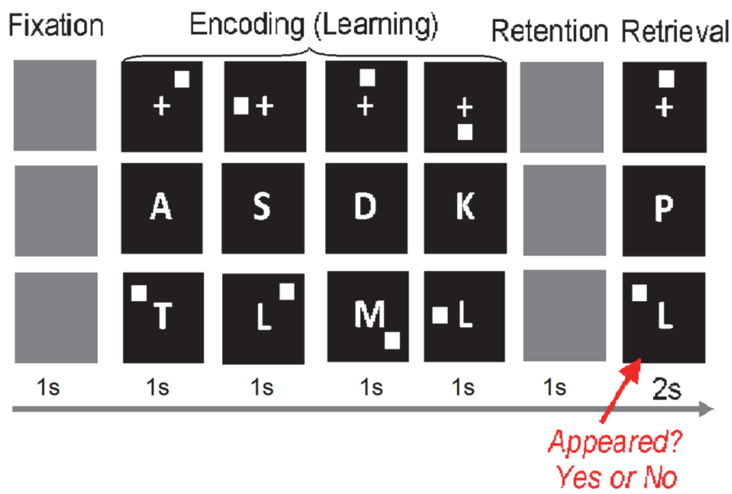


Figure 1 Schematic diagram of Sternberg Test

As illustrated in Figure 1, subjects were asked to take a Sternberg working memory test, which involves presentation of a list of wayfinding information items to memorize (encoding or learning period), followed by a memory retention period during which the subject must maintain the list of items in memory, and a retrieval period in which the subject will answer if a given item appeared. There is a short period of fixation phase between two sessions.

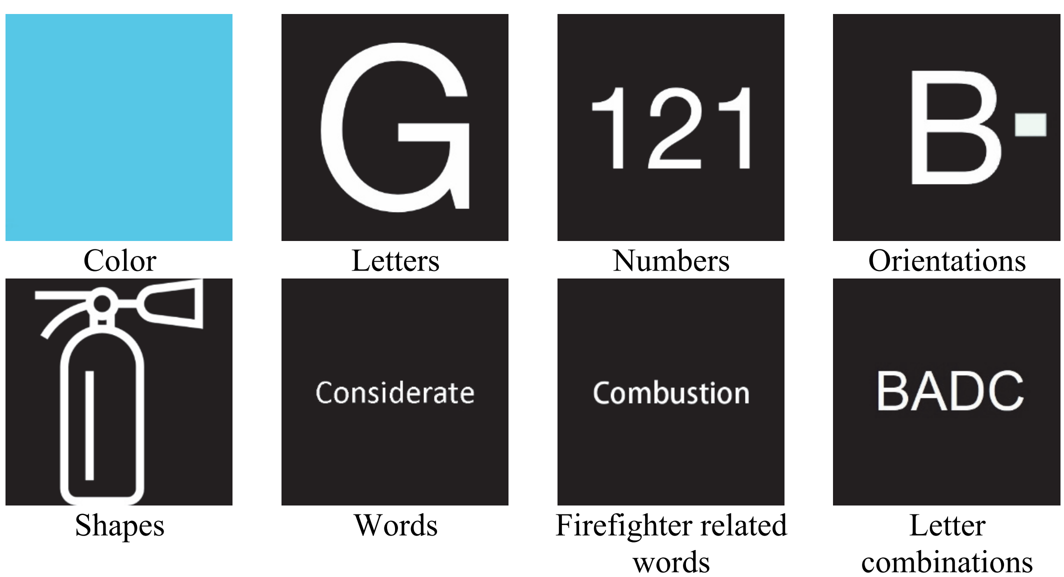


Figure 2 Examples of information items in Sternberg Test

To stimulate cognitive load increases, the Sternberg Test were repeated 153 times with different sets of images for each subject. As the example shown in Figure 2, the 153 tests contain 2 sets of color, 21 sets of letters, 22 sets of numbers, 29 sets of orientations, 25 sets of shapes, 20 sets of daily words, 8 sets of firefighter related words, and 26 sets of letter combinations to induce cognitive load similar to that firefighters potentially could bear. During the test, subjects' response time were recorded and the Enobio 32 from Neuroelectrics® was attached to a subject's head to collect EEG signals. The referenced EEG data were transmitted wirelessly to a computer for further analysis.

DATA ANALYSIS

We collected data of male subjects ($n = 5$) aging from 22 to 28 with the average at 24.8 in this research to find disparities. All of them are graduate students in College of Engineering in Northeastern Univerity. Among them, one subject did two experiments to provide information in investigating personal varieties. The recorded EEG data was processed using EEGLAB, an open source toolbox in Matlab (Delorme and Makeig 2004). A bandpass filter from 1Hz to 50 Hz was applied as the first step to extract clean EEG signals. PREP-pipeline (Bigdely-Shamlo et al. 2015) was applied to the signal to detrend and to remove line-noise. Then, independent component analysis (ICA) (Onton and Makeig 2006) and ADJUST (Mognon et al. 2011) were applied on the data to detect and reject artifacts efficiently and automatically to remove artifacts generated by muscles and eye movements.

After the clean EEG data was acquired after preprocessing, fast Fourier transform (FFT) was applied on the time domain signal to acquire alpha and theta wave. The alpha and theta band energy were calculated on frequency domain using equations below.

$$E_\theta = \sum_{f=4}^{8} X(f)\,, E_\alpha = \sum_{f=8}^{12} X(f)\,, E_{tot} = \sum X(f)$$

$$E_{r\theta} = E_\theta / E_{tot}\,, E_{r\alpha} = E_\alpha / E_{tot}$$

$$TAR = E_\theta / E_\alpha$$

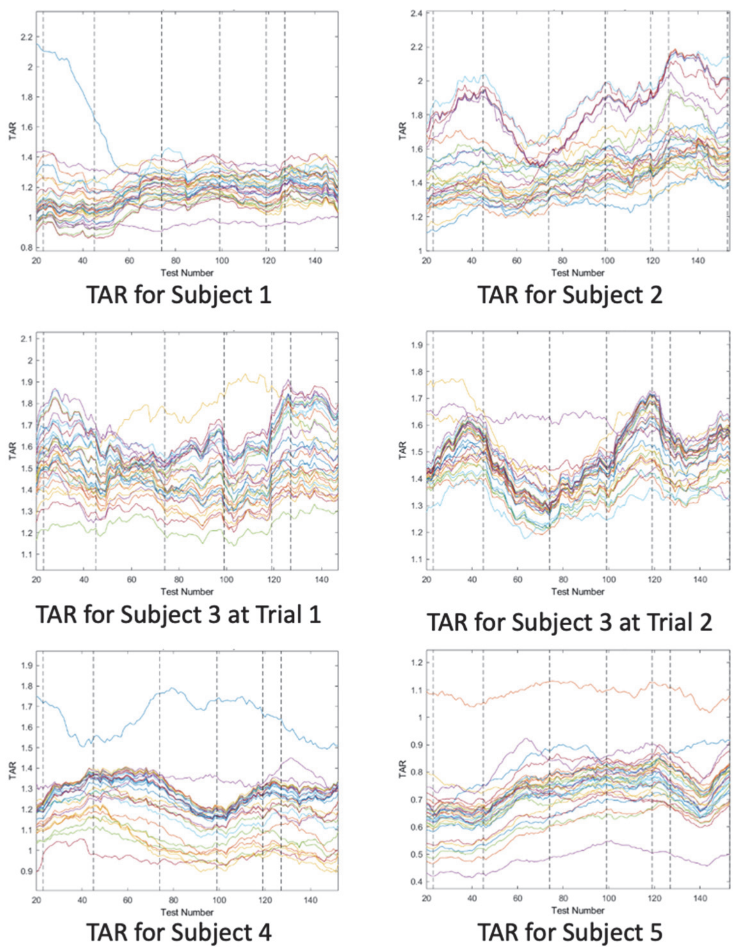
In which, $X(f)$ represents the frequency domain energy from the result of FFT, E_θ, E_α and E_{tot} are, in separate, theta sub band energy, alpha sub band energy, and total energy, $E_{r\theta}$ and $E_{r\alpha}$ are theta and alpha band relatively energy respectively, and TAR is short for the Theta to Alpha Ratio. Pearson correlation is calculated between recorded response time and TAR value to check the reliability of TAR.

RESULTS

The TAR oscillates due to the fluctuation of EEG signal. A Moving-Average (MA) filter is applied on the E_θ , E_α , and TAR through the following difference equation to smoothen the figure and visualize general trends.

$$y(n) = b(1)x(n) + b(2)x(n-1) + \cdots + b(m)x(n-m)$$

in which $y(n)$ and $x(n)$ represent the output from the filter and input to the filter at nth point respectively, $b(m)$ is the mth element of a predefined b array. In this case, the window length is selected to be m and m elements in b are $1/m$. Smoothened TAR results are as illustrated in



TAR for Subject 1 TAR for Subject 2

TAR for Subject 3 at Trial 1 TAR for Subject 3 at Trial 2

TAR for Subject 4 TAR for Subject 5

Figure ***3***.

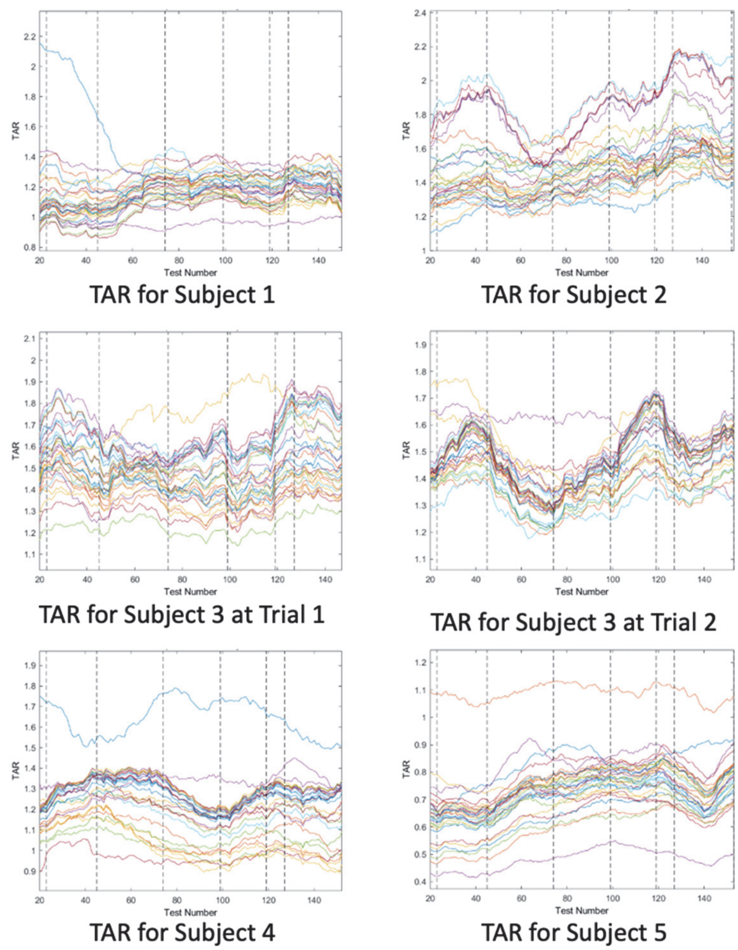


Figure 3 TAR results smoothened by $m = 20$

In the figure above, x axis is the n in the MA filter difference equation and y axis is the TAR. Vertical dash line marks the end of last types of images and the start of a new set of tasks. Lines with diverse colors represent the TAR index for individual channels over the test. Legends were hidden and no additional pattern applied for lines to avoid confusion and keep the figure clean. Changes of cognitive load introduced by information images in Sternberg Test for certain part of brain could be illustrated by the trend of lines. Based on the design of Sternberg Test, the mapping from test numbers to types of information images are listed in **Error! Reference source not found.**.

Table 1 Map between test numbers and types of images

Types	Test Numbers
Color	1 to 2
Letters	3 to 23
Numbers	24 to 45
Orientations	46 to 74
Shapes	75 to 99

Daily words	100 to 119
Firefighter related words	120 to 127
Letter combinations	128 to 153

The result of Pearson Correlation showed relatively low ρ values and comparably high P values because of the oscillation of EEG TAR index. An MA filter with the window size of 20 is applied on both response time and TAR to minimize the fluctuation of TAR data and achieve better Pearson Correlation result. Results are listed below.

Table 2 Pearson Correlation between TAR and Reaction Time

Subject Number	$\boldsymbol{\rho}$ value	$\boldsymbol{P}$ value
1	0.8550	4.8600×10^{-44}
2	0.3589	5.2318×10^{-6}
3 1st trial	0.9093	4.4009×10^{-57}
3 2nd trial	0.8789	2.1384×10^{-50}
4	0.9477	2.9199×10^{-76}
5	0.9525	7.0251×10^{-80}

Table 2 indicates that there is a positive correlation between TAR and response time at a high confidence level except subject 2. Considering the response time has been taken as one of the most significant part of subject performance (Barrouillet et al. 2007), the correlation indicates the Moving Average TAR could be taken as a direct cognitive load indicator.

DISCUSSION

Firstly, individual differences are observed. According to the result in

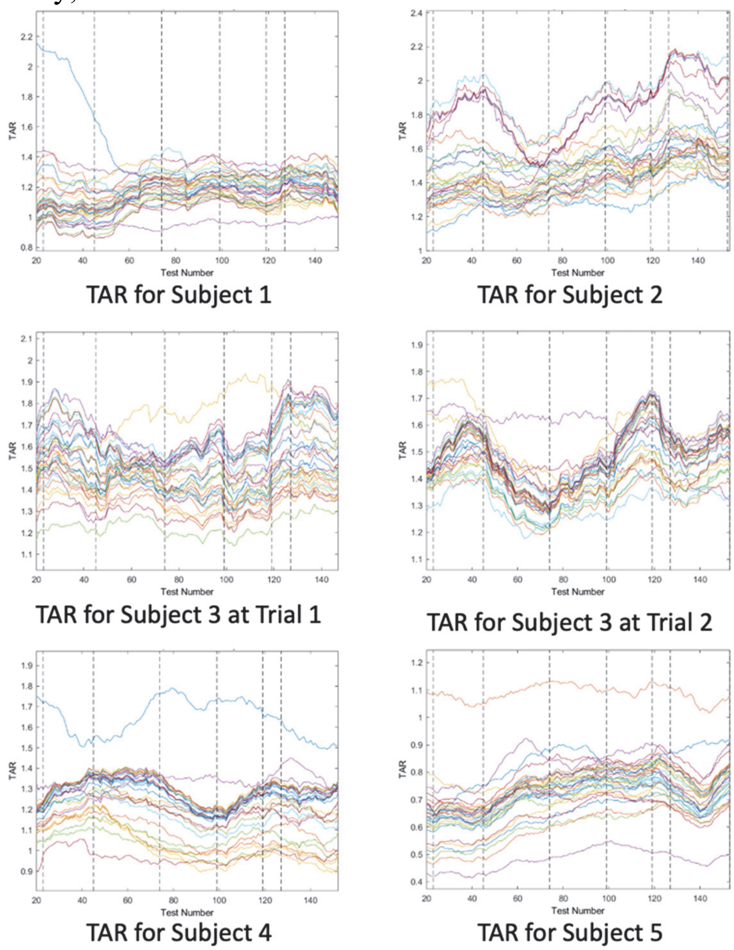


Figure ***3***, it could be observed that five subjects have shows distinct TAR performance during the test. For example, for subject 2, cognitive load increases significantly when the subject was exposed to numbers information. But the cognitive load for subject 4 merely shows a slight increase during the same procedure. Subject 1 and subject 3 at first trial have had a sharp decrease in TAR, indicating the fact that numbers brings relatively low mental burden in working memory development. Similar variations arising from the individual differences are also found in the rest types of images except for the color images due to the lack of tests.

Secondly, the non-stationarity for individual subject could be found in the result. The comparison between two sets of data from subject 3 reveals a potential that subjects may learn from the previous trial. Besides the remarkable contrast on cognitive behaviors in numbers information and letter combination information, subject 3 achieves some progress in the rest part of the test. In orientation tests, cognitive load stayed stable at a level during the first trial. But the cognitive load decreased significantly at the second trial. More significantly, the range of fluctuation of TAR in the second trial is smaller than the first one. This might be the result of both the training effect from the previous trial and the non-stationarity inherited by EEG. However, the sharp drop from test number 99 to 119 in the first trial is not because of the change of mental workload. It is mainly as the consequence of a huge peak existence in the alpha band energy. This might arise from a several-seconds eyes closure.

Lastly, the spatial correlation could be observed from the same performance in most of the channels. In the figure of TAR of Subject 2, some parts of brain have a relatively high cognitive load and those channels follows similar trends which are different from others. Those channels are from parietal and occipital parts of brain and part of them are in the visual cortex. The separation might reveal higher visual demand or the lack of baseline data. In the rest five figures, channels are highly spatially correlated, which are mainly because of the conductance of tissues and skull.

CONCLUSIONS

The massive cognitive load introduced by extensive information was increasing the level of danger in making vital decisions for firefighters. The understandings in the resources of personal cognitive load are necessary to decrease the potentiality of cognitive overload and control the load within a safe level. In this research, we performed an experiment to trigger specific cognitive load arising from a certain type of information and an automatic EEG processing method based on EEGLAB and plugins. Theta to Alpha Ratio was selected as the assessment of the cognitive load separated in 32 channels.

The result of this research reveals that the design of experiment is capable to generate various cognitive, the automatic EEG processing method is sufficient in extracting true brain activity signals, and the TAR is capable to quantify the cognitive load of individual channels.

The study, as a first step of measuring cognitive load in the context of Sternberg tests, can be expanded in multiple ways. First, future work could employ a new design of Sternberg Test which provides the subjects enough time to relax between different types of information to exclude the influence from the previous type. The new test could provide better resource for classification and identification of cognitive load which could help us identify the source of mental workload during firefighter's training period and, as a result, optimize the information delivery for individuals. Second, the automatic EEG proessing method would enable the real-time whole-brain cognition measurement, which will be of help to monitor the cognitive status of firefighters and make decisions to prevent overload. The method could further be applied on workers in other high hazardous conditions to improve their performance in decision making. Also, individual differences and the spatial correlation are to be decreased with the application of Neural Networks with the TAR as input and the effect of non-stationarity could also be attenuated with enough diversity of individuals data.

ACKNOWLEDGEMENTS

This study is supported by the National Institute of Standards and Technology (Grant No. 1806730) and National Science Foundation (Award No. 1761459, 1761950). Any opinions, findings, and conclusions or recommendations expressed in

this material are those of the authors and do not necessarily reflect the views of the National Institute of Standards and Technology and the National Science Foundation.

REFERENCES

Antonenko, P., Paas, F., Grabner, R., & Van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, *22*(4), 425-438.

Başar, E. (2012). *Brain function and oscillations: volume II: integrative brain function. Neurophysiology and cognitive processes*. Springer Science & Business Media.

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., ... & Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, space, and environmental medicine*, *78*(5), B231-B244.

Bian, Z., Li, Q., Wang, L., Lu, C., Yin, S., & Li, X. (2014). Relative power and coherence of EEG series are related to amnestic mild cognitive impairment in diabetes. *Frontiers in aging neuroscience*, *6*, 11.

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K. M., & Robbins, K. A. (2015). The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Frontiers in neuroinformatics*, *9*, 16.

Brennan, C. (2011). *The combat position: Achieving firefighter readiness*. Fire Engineering Books.

Dan, A., & Reiner, M. (2017). EEG-based cognitive load of processing events in 3D virtual worlds is lower than processing events in 2D displays. *International Journal of Psychophysiology*, *122*, 75-84.

Davies, N. (2015). *Decision Making on the Fireground*. Retrieved from https://www.firerescuemagazine.com/articles/print/volume-10/issue-9/firefighting-operations/decision-making-on-the-fireground.html

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9-21.

Evarts, B., & Molis, J. L. (2018). *United states firefighter injuries-2017*. National Fire Protection Association. Research, Data and Analytics Division.

Galy, E., Cariou, M., & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types?. *International Journal of Psychophysiology*, *83*(3), 269-275.

Gramfort, A., Strohmeier, D., Haueisen, J., Hämäläinen, M. S., & Kowalski, M. (2013). Time-frequency mixed-norm estimates: Sparse M/EEG imaging with non-stationary source activations. *NeuroImage*, *70*, 410-422.

He, B., Sohrabpour, A., Brown, E., & Liu, Z. (2018). Electrophysiological source imaging: A noninvasive window to brain dynamics. *Annual review of biomedical engineering*, *20*, 171-196.

Hefron, R. G., Borghetti, B. J., Christensen, J. C., & Kabban, C. M. S. (2017). Deep long short-term memory structures model temporal dependencies improving cognitive workload estimation. *Pattern Recognition Letters*, *94*, 96-104.

Hefron, R., Borghetti, B., Schubert Kabban, C., Christensen, J., & Estepp, J. (2018). Cross-participant EEG-based assessment of cognitive workload using multi-path convolutional recurrent neural networks. *Sensors*, *18*(5), 1339.

Henderson, S. N., Van Hasselt, V. B., LeDuc, T. J., & Couwels, J. (2016). Firefighter suicide: Understanding cultural challenges for mental health professionals. *Professional Psychology: Research and Practice*, *47*(3), 224.

Holm, A., Lukander, K., Korpela, J., Sallinen, M., & Müller, K. M. (2009). Estimating brain load from the EEG. *The Scientific World Journal*, *9*, 639-651.

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, *48*(2), 229-240.

Mora, W. R. (2003). *US Firefighter Disorientation Study: 1979-2001*. WR Mora.

Onton, J., & Makeig, S. (2006). Information-based modeling of event-related brain dynamics. *Progress in brain research*, *159*, 99-120.

Roja, Ž., Kaļķis, V., Kaļķis, H., & Pencis, I. (2009, January). Assessment of firefighters-rescuers' work severity in relation with interaction between physical and mental load. In *Proceedings of the Latvian Academy of Sciences. Section B. Natural, Exact, and Applied Sciences.* (Vol. 63, No. 6, pp. 264-270). Versita.

Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., & Faubert, J. (2019). Deep learning-based electroencephalography analysis: a systematic review. *Journal of neural engineering*.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American scientist*, *57*(4), 421-457.

Sweller, J. (2011). Cognitive load theory. In *Psychology of learning and motivation* (Vol. 55, pp. 37-76). Academic Press.

Yin, Z., & Zhang, J. (2016, July). Recognition of cognitive task load levels using single channel EEG and stacked denoising autoencoder. In *2016 35th Chinese Control Conference (CCC)*(pp. 3907-3912). IEEE.

Zhang, J., Li, S., & Wang, R. (2017). Pattern recognition of momentary mental workload based on multi-channel electrophysiological data and ensemble convolutional neural networks. *Frontiers in neuroscience*, *11*, 310.