Advanced Engineering Informatics 46 (2020) 101153

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

A neurophysiological approach to assess training outcome under stress: A virtual reality experiment of industrial shutdown maintenance using Functional Near-Infrared Spectroscopy (fNIRS)

Yangming Shi[a], Yibo Zhu[b], Ranjana K. Mehta[b], Jing Du[a,*]

[a] Department of Civil and Coastal Engineering, Engineering School of Sustainable Infrastructure and Environment (ESSIE), Herbert Wertheim College of Engineering, University of Florida, Gainesville, FL 32611, United States

[b] Department of Industrial & Systems Engineering, Texas A&M University, College Station, TX 77843, United States

ARTICLE INFO

Keywords:
Shutdown maintenance training
Virtual reality
Eye-tracking
fNIRS

ABSTRACT

Shutdown maintenance, i.e., turning off a facility for a short period for renewal or replacement operations is a highly stressful task. With the limited time and complex operation procedures, human stress is a leading risk. Especially shutdown maintenance workers often need to go through excessive and stressful on-site trainings to digest complex operation information in limited time. The challenge is that workers' stress status and task performance are hard to predict, as most trainings are only assessed after the shutdown maintenance operation is finished. A proactive assessment or intervention is needed to evaluate workers' stress status and task performance during the training to enable early warning and interventions. This study proposes a neurophysiological approach to assess workers' stress status and task performance under different virtual training scenarios. A Virtual Reality (VR) system integrated with the eye-tracking function was developed to simulate the power plant shutdown maintenance operations of replacing a heat exchanger in both normal and stressful scenarios. Meanwhile, a portable neuroimaging device – Functional Near-Infrared Spectroscopy (fNIRS) was also utilized to collect user's brain activities by measuring hemodynamic responses associated with neuron behavior. A human–subject experiment (n = 16) was conducted to evaluate participants' neural activity patterns and physiological metrics (gaze movement) related to their stress status and final task performance. Each participant was required to review the operational instructions for a pipe maintenance task for a short period and then perform the task based on their memory in both normal and stressful scenarios. Our experiment results indicated that stressful training had a strong impact on participants' neural connectivity patterns and final performance, suggesting the use of stressors during training to be an important and useful control factors. We further found significant correlations between gaze movement patterns in review phase and final task performance, and between the neural features and final task performance. In summary, we proposed a variety of supervised machine learning classification models that use the fNIRS data in the review session to estimate individual's task performance. The classification models were validated with the k-fold (k = 10) cross-validation method. The Random Forest classification model achieved the best average classification accuracy (80.38%) in classifying participants' task performance compared to other classification models. The contribution of our study is to help establish the knowledge and methodological basis for an early warning and estimating system of the final task performance based on the neurophysiological measures during the training for industrial operations. These findings are expected to provide more evidence about an early performance warning and prediction system based on a hybrid neurophysiological measure method, inspiring the design of a cognition-driven personalized training system for industrial workers.

1. Introduction

Industrial shutdown maintenance (hereafter, shutdown maintenance) is an event wherein the entire plant is shut down for a short period of time for renewal [1]. It plays a critical role in renovating America's infrastructure systems [2]. The Energy Information Administration data [3] shows that the shutdown maintenance has become more intensive recently: just in the first six months of 2018, there were

* Corresponding author.
E-mail addresses: shiyangming@ufl.edu (Y. Shi), zhuxx355@tamu.edu (Y. Zhu), rmehta@tamu.edu (R.K. Mehta), eric.du@essie.ufl.edu (J. Du).

https://doi.org/10.1016/j.aei.2020.101153
Received 2 March 2020; Received in revised form 9 June 2020; Accepted 3 August 2020
Available online 07 August 2020

7783 planned outages in the US due to industrial shutdown maintenance works. A common challenge of shutdown maintenance is the excessive stress posted to the workers [1]. For example, 4500 contractors were involved in the second major shutdown at the Muskeg River Mine and Scotford Upgrader by Shell and more than 250 valves were repaired and installed within a two-month timeframe [4]. To minimize the impact of the shutdown schedule, the work is usually done in a 24/7 manner [5]. Typical work schedule is 12 h a day, and 7 days a week [6]. To cope with the tight schedule, workers also need to go through fast while excessive training on the site to digest a large amount of complex information (e.g., engineering and operation instructions) in limited time [6]. This presents potentially significant stress to the shutdown maintenance workers during the on-site training, as well as underexplored implications in their final task performance. Possibly driven by the stress issues (during training and during operations), shutdown maintenance is becoming one of the most dangerous career in the US oil and gas industry, there have been 152 documented major industrial disasters since 2000, causing more than 40 deaths and even more injuries [7]. The US Chemical Safety Board (CSB) has indicated that the lack of inherent safety training principles and processes are the root causes of shutdown maintenance accidents [8] and growing evidence has linked the root causes of these incidents to human errors directly and indirectly tied to excessive stress [9–13].

Since stress is of particular interest due to the high-risk nature of shutdown maintenance and the likely psychological impacts on the workers, there is a pressing need for stress and training assessment in a timely or even real-time manner, for fast responses and early interventions for potential human errors. Most training is assessed afterward (i.e., after the training is done), but given the time constraints of the shutdown maintenance, a pre-training effectiveness assessment is needed. This research focuses on the training assessment during the shutdown maintenance training, since training quality serves as a potential predictor of the final performance and allows a more proactive early intervention [14,15].

This study aims to test a neurophysiological approach for assessing and forecasting workers' training quality of the shutdown maintenance operations based on a temporal analysis of fNIRS data and neural connectivity patterns during the training. It also helps build a knowledge base for an early warning system based on neural analysis during training. The experiment results indicated that stressful training scenario had a strong impact on participants' neural connectivity patterns and physiological metrics, and finally affected participants' task performance. Owing to the difficulty of simulating industry shutdown operation tasks in the real world, a Virtual Reality (VR) system integrated with the eye-tracking function was developed to simulate a typical power plant shutdown maintenance operation. The operation is replacing a plate heat exchanger in both normal and emergency stressful scenarios. Meanwhile, a neuroimaging device - Functional Near-Infrared Spectroscopy (fNIRS) was utilized to collect user's brain activities by measuring hemodynamic responses associated with neuron activation levels. The participants' task performance including operation time and operation accuracy were used as the indicators of training quality, while the temporal analysis of neural connectivity patterns and gaze movement patterns were used to evaluate workers' neurocognitive performance in normal and stressful scenarios. The results suggest that simulated stress during the training can serve as an important adaptive factor for desired training outcomes. We also found a significant correlation between the neurophysiological features including gaze movement and fNIRS data, and the final task performance. Based on the findings, we propose a framework for a classification model that may use fNIRS signals in the review session to estimate individual's task performance following training. These findings are expected to provide more evidence about an early performance warning and performance forecasting system based on a hybrid neurophysiological measure method, inspiring the design of a cognition-driven personalized training system for industrial workers. The remainder of this paper introduces the point of departure of this study, the research method and the experiment, and the findings and recommendations.

2. Literature review

2.1. Assessing mental stress in construction operations

Construction industry is known as one of the most stressful industries because of high physical and mental demands [16,17]. According to a professional survey conducted in the United Kingdom (UK) in 2006 [18], nearly 68% of the construction workers have suffered from obvious stress on construction sites. Haynes and Love [19] found three most significant mental stressors experienced by the construction professionals, including high workload, long working hours, and insufficient time with family. The high level of mental stress amplifies the construction workers' errors and leads to increasing unsafe behaviors [16,20]. Many scholars have proposed different approaches to assess individual's mental stress. The most common method to evaluate workers' mental stress on construction sites is the subjective questionnaire or survey [21,22]. It was proven to be an effective method to evaluate a large number of workers' mental stress status at the same time on construction sites. However, Jebelli [16] pointed two limitations of this assessment method, which were interrupting workers' tasks and imprecise subjective evaluation. Thus, other scholars have explored physiological measurements to evaluate individual's mental stress level such as cortisol and glucocorticoids [23], Electrocardiography (ECG) [22,24,25], and Electroencephalography (EEG) [26–28]. Although these approaches can provide on individual's mental stress status, these methods are hard to be implemented in real-world projects due to the technical complexity and cost. Jebelli [16] indicated that high-quality EEG signals can be only collected in a well-controlled lab environment setting since EEG devices are very sensitive to individual's motion. Despite a variety of signal processing filters and algorithms have been developed to handle the intrinsic motion artifacts (e.g., eye blink, facial muscle movement), it is still very challenge of collecting high-quality EEG data during human locomotion or large scale body movement in dynamic work environments [29–31]. In summary, this study proposed an alternative neurophysiological approach - functional near-infrared spectroscopy (fNIRS) to evaluate individual's stress status.

2.2. Electroencephalogram (EEG) and Functional Near-Infrared Spectroscopy (fNIRS)

EEG measures the electric current density on the scalp due to the task-related neural activity [32]. It offers significant higher temporal resolution compares other neural imaging methods such as functional magnetic resonance imaging (fMRI), positron emission tomography (PET), and functional Transcranial Doppler sonography (fTCD), but lacks spatial resolutions [33,34]. The mobility of EEG system is considered moderate compares to fMRI and PET, which allows for cognitive and controlled motor tasks laboratory based studies [34]. Compared to EEG that offers higher temporal resolution, fNIRS has a relatively low sampling frequency of less than 20 Hz but offers higher spatial resolution which is essential in monitoring mental workload related brain regions [35,36]. fNIRS system also requires significantly less calibration and preparation efforts than both dry and wet electrodes EEG systems [37]. To minimize the impact of motion artifacts on the signal quality, greater computational effort was required for filtering and frequency domain analysis of EEG data. In contrast, less filtering and time domain analysis demands less computational effort for fNIRS signal processing due to its robustness to motion artifacts [37]. The comparison between EEG and fNIRS based previous literature was summarized in the Table 1. The fNIRS devices measure the changes in blood flow and oxygenation related to task-specific neural activities. The basic work principle of fNIRS is that the device sends infrared lights signals between 650 and 1000 nm wavelengths via multiple light

Table 1
The comparison between EEG and fNIRS based on previous literature.

	EEG	fNIRS
Experimental Setup	Laboratory setting; Greater preparation and calibration effort [37]	Both laboratory and field settings; Lower preparation and calibration effort [37]
Signal Quality	High temporal resolution. Vulnerable to motion artifacts	High spatial resolution; Resistance to motion artifacts
Data Analysis	Power Spectra; Coherence [52]	Time-Series; Connectivity [37]
Utilization	Mental Workload Monitoring [53]; Brain Computer Interface (BCI) [52]; Predictive Modeling [52]	Mental Workload Monitoring [54]; Brain Stimulation [55]; Predictive Modeling

emitters into the scalp. The photodetector measures the strength of the received light signals and converts into the concentration changes of both oxygenated (ΔHbO) and deoxygenated (ΔHbR) hemoglobin in the channels formed by the closest emitters and detectors [38]. The emitters and detectors are designed to be placed within 3 cm. The fNIRS devices have been proven to be a non-invasive safe neuroimaging technique since the fiber optics are very suitable for any head position and posture [38]. The fNIRS devices have been widely used in the neuroimage research areas to investigate infants' language development [39–41]. Multimodal neuroimaging studies that compensate EEG system with fNIRS system have been explored to deliver complimentary data on task related neural activations and improve motor task classification model performance [42,43]. fNIRS system is much more robust to motor task-related signal artifacts and therefore, adopted by this study to quantify individual's stress status. Recently, it has also shown the potential as a promising neuroimaging technology that can better be integrated with VR devices to explore human cognitive process [44–47]. In the construction literature, Hu and Shealy [48] used fNIRS devices to investigate sustainable engineering decision-making and design cognition [49]. They also utilized the fNIRS device to explore the cognitive response to hazards on the construction site [50]. Du, Zhu, Shi, Wang, Lin and Zhao [51] used fNIRS to investigate the cognitive load in processing different formats of engineering information. In this study, we used a wireless fNIRS device integrated with the VR system to assess the test subjects' stress status and training quality of the shutdown maintenance operations.

3. Methodology

3.1. Experiment apparatus: VR, eye tracker, and fNIRS

Owing to the difficulty of simulating industrial shutdown maintenance operations in the real world, a VR system integrated with eye-tracking function and neuroimaging function was developed based on our previously well-validated VR systems [56–59]. In order to collect high-precise and high-resolution gaze movement data, the Tobii Pro eye tracker integrated with HTC VIVE Head Mounted Display (HMD) [60] was used. The Tobii Pro VR integration is manufactured by Tobii and uses advanced Pupil Centre Corneal Reflection (PCCR) remote eye-tracking technique to capture eyeball movement and pupil size [61]. The Near-infrared illuminators in the eye tracker are used to create the reflection patterns on the cornea and pupil of the eye. The cameras in the eye tracker are used to capture high-resolution images. Finally, the advanced image-processing algorithms and a physiological 3D model of the eye are implemented to estimate the position of the eye in the virtual environment and the user's pupil size [61]. The Tobii Pro VR integration eye-tracker has an accuracy of 0.5° and the maximum gaze data output frequency is 120 Hz [62]. To achieve the eye-tracking and playback functions in the virtual environment, we developed several C# scripts based on the Tobii Pro Software Development Kit (SDK) [62] and the application programming interface (API) in Unity. Fig. 1 shows the eye-tracking in the virtual environment. Fig. 1(a) shows the eye-tracking data collection mode in the virtual environment. Fig. 1(b) shows the playback visualization function of showing the gaze movement in the virtual environment. The white-purple lines indicate the gaze movement trajectories. In the virtual environment, the system collected participants' gaze movement data, body movement data, hand movement data, and pupil diameter data with a frequency of 90 Hz. The gaze and pupil tracking serve as supplementary evidence of the stress assessment, and the body and hand movement data is used to evaluate task performance. After each VR experimental trial, the developed VR system automatically generated a CSV file with all the raw data. Fig. 2(a) shows the eye-tracking function in the immersive virtual environment. Meanwhile, the Cerebral hemodynamic response of each participant was monitored using an 18-channel portable fNIRS system NIRSportTM (NIRx Medical Technologies, NY, USA). The system consists of 8 emitters (in red, Fig. 2(b)) and 8 detectors (in blue, Fig. 2(b)). The emitters and detectors were designed to be located less than 3 cm. The infrared light signals were emitted in two wavelengths (760 and 850 nm) and collected at a sampling frequency of 7.81 Hz through the detectors. The VR and fNIRS systems were synchronized by the Psychopy software during the experiment. In order to avoid the interference of the infrared light generated from the VR lighthouses with the fNIRS device, a black shower head was used to cover the emitters and detectors of the fNIRS device during the experiment.

3.2. Virtual environment and experimental task

Two immersive virtual training scenarios were created in this study, i.e., the normal training scenario and the stressful training scenario, as shown in Fig. 3. To control the undesired influence of the virtual environment on participants' task performance, we designed the same plate heat exchanger in a virtual operation room and set the environment lighting the same in both conditions. In the simulated space, each participant could see the limited space boundary, and they were told not to go beyond the boundary when they performed the task. The participant could interact freely with the virtual plate heat exchanger and each valve using the HTC controllers. The experiment task was designed to let each participant memorize sequences for turning or closing the valves before they replaced the plate heat exchanger. Each participant was asked to memorize two different operation sequences in both normal and stressful scenarios and then perform the operations in both normal and stressful conditions. Each pre-start-up sequence to cut off/open the hot water and cold water consisted of 10 steps, which were developed based on the operation instruction manual of Alfa Laval plate heat exchangers, as listed in Table 2 [63]. The pipe operation sequence as shown in Table 2 were consistent to each participant for each experiment condition. The experiment was conducted in a well-controlled virtual environment. Since we used a within-participant experiment design, to avoid the learning effect from the previous sessions, we designed two different pipe operation sequences with different valve positions. The two sequences were carefully designed to reflect the same level of difficulty. For the normal review session as shown in Fig. 3(a), the instruction of pipe operation sequence was placed on the left side the virtual plate heat exchanger model and each valve was marked with a valve number on the virtual model. The participant could navigate freely in the virtual environment. We added serval stressors in the stressful review session as shown in Fig. 3(b). Since we used a within

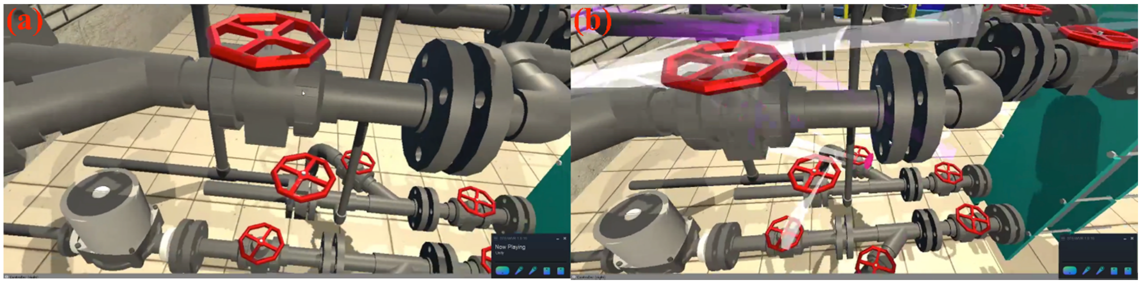

Fig. 1. The eye-tracking in the virtual environment. (a) data collection mode; (b) data visualization mode.

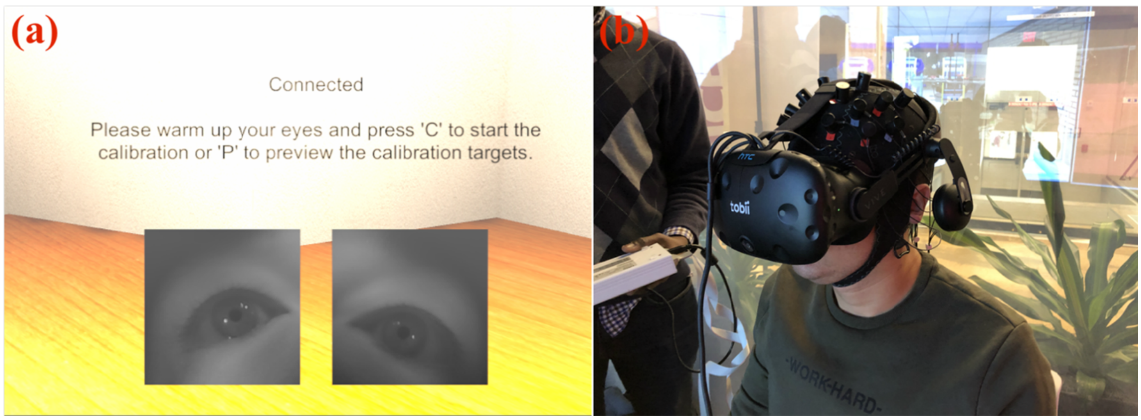

Fig. 2. Experiment setting. (a) eye-tracking function in the virtual environment; (b) equipment setting in the real-world.

participant experiment design, to rule out learning effects, we used a different pipe operation sequence in the stress condition and redesigned all the valve positions on the virtual heat exchanger model. The stressors in the stressful review scenario included simulated smoke gradually occluding the vision, virtual fire propagation, virtual smoke propagation, sudden structural collapse sound, and fire burning sound in the distance. The purpose of adding these stressors in the virtual environment was to simulate the stressful shut-down maintenance scenario as realistic as possible. For the normal operation session as shown in Fig. 3(c), the same virtual plate heat exchanger model was placed in the middle of the operation room, but there were not valve numbers displayed for each valve. The virtual setting was the same as the normal review session for a controlled experiment. Finally, in the stressful operation session we also added the same stressors as demonstrated in the stressful review session as shown in Fig. 3(d).

3.3. Experiment procedure

All the participants were asked to memorize two different 10-step pipe operation sequences with the virtual pipe model in both normal and stressful virtual scenarios, and then to perform the pipe operation in two scenarios respectively. To avoid the influence from previous

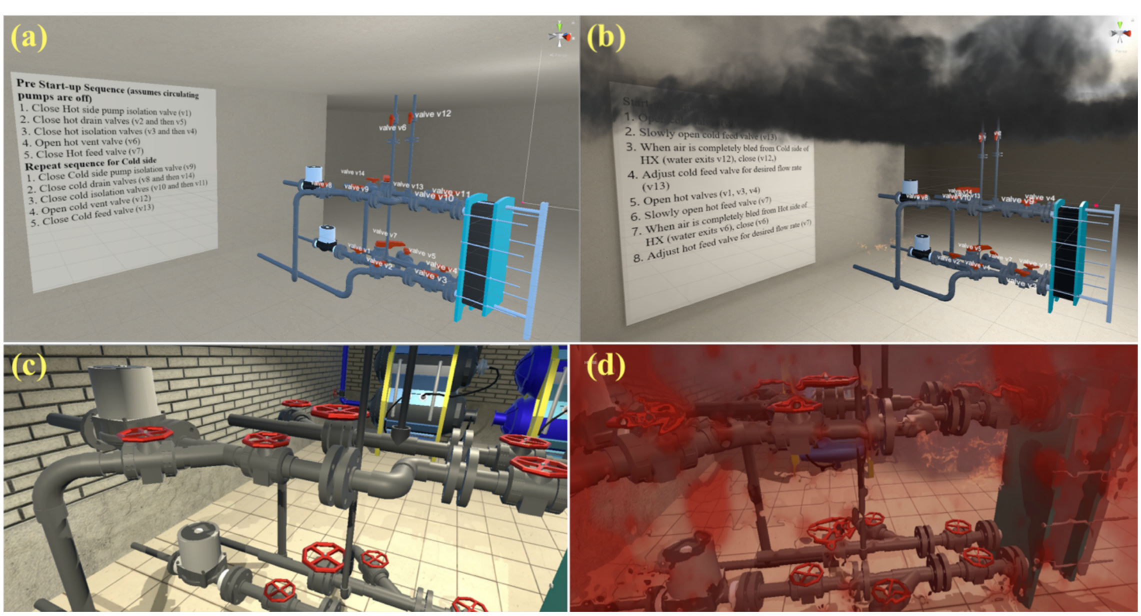

Fig. 3. The training scenarios. (a) normal review scenario; (b) stressful review scenario; (c) normal training scenario; (d) stressful training scenario.

Table 2
10-steps of operation sequence for the pre-start-up plate heat exchanger across two conditions.

Normal Training Scenario	Stressful Training Scenario
Step 1: Close hot side pump isolation valve (v1)	*Step 1: Open Cold Valves (v9 and v10)*
Step 2: Close hot drain valves (v2 and then v5)	*Step 2: Open Cold Valve (v11)*
Step 3: Close hot isolation valves (v3 and then v4)	*Step 3: Slowly open cold feed valve (v13)*
Step 4: Open hot vent valve (v6)	*Step 4: When air is completely bled from cold side of HX, close (v12)*
Step 5: Close hot feed valve (v7)	*Step 5: Adjust cold feed valve for desired flow rate (v13)*
Step 6: Close cold side pump isolation valve (v9)	*Step 6: Open Hot Valves (v1 and v3)*
Step 7: Close cold drain valves (v8 and then v14)	*Step 7: Open Hot Valve (v4)*
Step 8: Close cold isolation valves (v10 and then v11)	*Step 8: Slowly open hot feed valve (v7)*
Step 9: Open cold vent valve (v12)	*Step 9: When air is completely bled from hot side of HX, close (v6)*
Step 10: Close cold feed valve (v13)	*Step 10: Adjust hot feed valve for desired flow rate (v7)*

scenario, we used a counter-balanced approach randomly choosing to start with the normal or stressful scenarios for each participant. It means that some of the participants started with normal sessions while others started with stressful sessions. The participants were also instructed that their compensation would be determined by the task performance. The purpose of this experimental design was to motivate participants to memorize the pipe operation sequence and perform the pipe operation as accurately as possible. The experiment consisted of five sessions: (1) preparation session, (2) device calibration and VR training session, (3) review session, (4) retention session, (5) operation session. The preparation session (5–10 min) was designed to allow participants to familiarize the procedure and potential benefit or risk of the experiment. Participants' demographical information including age, gender, major, degree level, previous game and VR experience, and knowledge level of the HVAC system were also collected in this session. The device calibration and VR training session (10–20 min) were designed for participants to familiarize themselves with the fNIRS system, eye-tracking system, and interactions/navigation in the virtual environment. In this session, all the participants were first instructed to wear on the fNIRS device, and the investigators were able to ensure each probe of the fNIRS device accurately collected the neuroimaging data from the target brain regions. The participants were asked to stay claim in a chair and let the fNIRS device to set up the baseline data for each participant. After the fNIRS calibration, the participants were asked to set up the VR headset and the experiment investigators were also able to ensure participants' eyeball movements were accurately captured by the eye tracker integrated with the VR headset after serval five-point calibration in the virtual environment. Participants were also given instructions about how to use the two controllers to interact with the virtual valves. The review session (15 min) was used for participants to review and memorize the pipe operation sequence. The review session was divided into 10 trails (1 min for each review trail) for both normal and stress scenarios. For each trial in the review session, each participant was given 60 s to review and memorize the pipe operation sequence and pipe model. Between each review trial, there was a retention session (30 s), including 25 s of break time and 5 s of stand-by time. In the retention session, participants were told to sit to calm down. The purpose was to settle down participants' neural activities and minimize the influence on the following sessions. The participants were not asked to perform the operation in the review session. Instead, we designed a separate session for participants to perform the pipe operation after the 10 review trials. After the review session, participants were asked to perform the pipe maintenance task in normal or stress virtual environment (with no time limit). At the end of all experimental stages, participants were asked to provide comments and feedback on the experiment. We used NIRSTIM which is a programmed experiment instruction software to coordinate the collaboration between the fNIRS device and VR system. Fig. 4 shows the experiment procedure. The sequence between normal scenario and stress scenario was counter-balanced assigned to each participant. All the experiments were done at the same location (Francis Hall Room 101 – BIM CAVE at Texas A&M University), with the same devices. The environmental effects can be ruled out as well. Given that the experiment settings were well controlled, the final operation performance was indeed an indicator of the memory quality. We admit that in reality the task performance is also affected by other factors, such as the motor skills of the worker. However, as a study focusing on the impact of stress on knowledge-based learning (in this case, memorizing the correct sequence of valve operations), factors other than the memory quality are out of the scope. In summary, the experiment was designed in a way that memory quality based on review sessions was critical to the final performance. Difference in motor skill, for example, was deliberately removed with the same simple control mechanisms, i.e., touching the valves with two HTC controllers. The only influential factor is the difference in the use of stressors. We also made sure that the experimental stimuli were clear to participants, without any possible vague interpretations.

3.4. Data collection

We collected three types of data for post-experiment analysis: task performance, gaze movement, and neural activity. Task performance indicators include participants' operation time (s) and pipe operation accuracy (%). Pipe maintenance accuracy was defined as the accuracy in performing correct steps and directly represents how well the participants memorized and performed the pipe operation task. Pipe maintenance accuracy is recorded in a range from 0% to 100%. The operation time was defined as the time participants used to complete the task in the virtual environment. This indicator represents how efficiently the participants finished the task.

For the data analysis of gaze movement in the virtual environment, we extracted gaze transition approximate entropy (ApEn) [64] as a feature to evaluate participants' general attention patterns in the virtual environment. The ApEn was selected to evaluate the regularity and unpredictability of the fluctuations over participants' gaze movement data. ApEn is defined as a technique to quantify the regularity and complexity of the noisy time-series data [65]. This method is widely used in the data analysis of physiological time-series data such as heart rate [66,67], EEG [68,69], and endocrine hormone [70,71]. A higher value of gaze movement entropy indicates more irregularity and unpredictability of gaze movement, suggesting that participants just randomly look around in the environment. On the other hand, a lower value of gaze movement entropy shows a more regular and relatively stable gaze focus transitions [72]. Although we cannot conclude that distinct gaze movement patterns are results of different cognitive processes, but at least, we shall be able to claim that distinct gaze movement patterns, such as entropy of visual scan pathways, indicate the use of different scan patterns. This is supported by Hartley, Maguire, Spiers and Burgess [72] finding that the eye movement is associated with the forward motion and turning during the navigation. Jyotsna and Amudha [73] also found that the gaze movement is associated with the stress level. Therefore, we evaluated participants' visual scan patterns

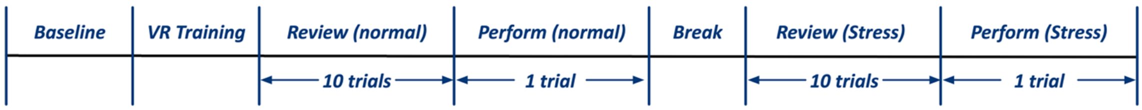

Fig. 4. Experiment procedure. The sequence between normal scenario and stress scenario was counter-balanced assigned to each participant.

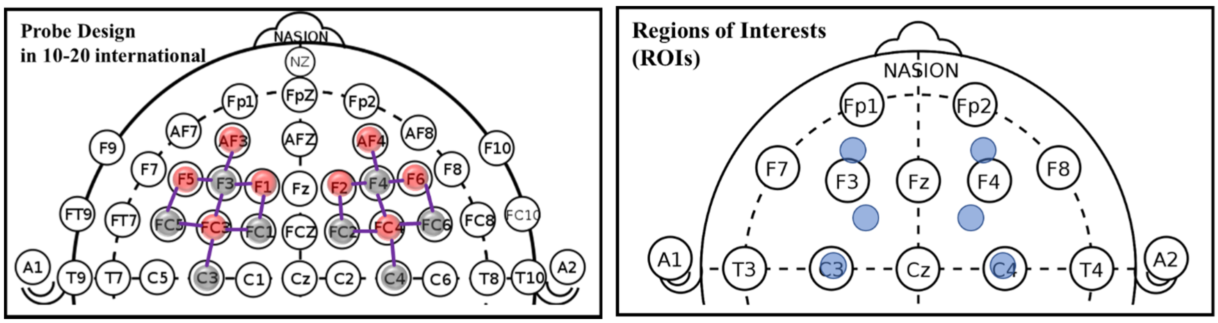

Fig. 5. fNIRS probe design in the international 10–20 system (left); fNIRS probe Regions of Interest (ROIs, right).

based on participants' gaze movement entropy in vertical and horizontal directions in this study.

The neural activities were measured with fNIRS. The fNIRS probe design is presented according to the international 10–20 system format with the probe cap placed on the vertex (Cz) of each participant (Fig. 5 left). The 18-channel system monitored six regions of interest (ROIs) which were specified by averaging the nearby channels (Fig. 5 right). These ROIs include left and right dorsal-lateral prefrontal cortex (L/R-DLPFC), left and right premotor cortex (L/R-PMC), and left and right primary motor cortex (L/R-M1). The dorsal-lateral prefrontal cortex has been shown to be associated with complex mental activities such as cognitive control network (CCN), dorsal attention network (DAN), and verbal episodic retrieval [74–76]. The premotor cortex has been shown to be associated with more complex and cognitive controls, such as the alternation of motor plans, task switching, acquisition of new motor skills, and motor selection [77–79]. The motor cortex has been well proven to be associated with motor movements [80–82]. ROIs corresponded Brodmann Areas and their functions are shown in Table 3.

The cerebral hemodynamic data collected by the fNIRS device was filtered by the band-pass filter. High-frequency noise was rejected at 3 Hz and motion artifacts caused by physiological noise such as heartbeat and slow-wave drift were corrected at 0.5 to 0.016 Hz [83]. Kurtosis wavelet algorithm [84] and spline interpolation [85] were used to reject abrupt motion artifacts and smooth the cerebral hemodynamic signals. At last, oxygenated (ΔHbO) and deoxygenated (ΔHbR) hemoglobin at the 18 channels was calculated by the modified Beer-Lambert law [83]. Fig. 6 shows a sample of participant #1′s post-processed fNIRS data related to the Dorsal-lateral prefrontal cortex across different conditions (red represents normal condition and orange represents stress condition).

In this study, oxygenated (ΔHbO) hemoglobin was used to analyze functional connectivity [83]. Functional connectivity measures task-related interactions among multiple cortical regions using covariance analysis of time series ΔHbO signals [83,86] as shown in Fig. 7. Pearson correlations, R, are calculated across all ROIs to find the correlation coefficients [87,88].

$$R_{ij}=\frac{\mathrm{cov}(x_i,x_j)}{s_i s_j} \tag{1}$$

$$z_{ij}=\frac{1}{2}\ln\left(\frac{1+R_{ij}}{1-R_{ij}}\right) \tag{2}$$

Eq. (1) represents the Pearson correlation coefficient, R values, calculated between the *i*th and *j*th signals where $i,j \in$ (LDLPFC, RDLPFC, LPMC, RPMC, LM1 and RM1), x_i and x_j are both ΔHbO signals, cov (x_i,x_j) represents the covariance between the *i*th and *j*th signals, and s_i and s_j stand for the standard deviations of the *i*th and *j*th signals. The calculated R values are then converted to Fisher's z-scores, Z values, to determine the strength of correlations following Eq. (2) [83]. Functional connectivity with z-score between 0.4 and −0.4 were identified as not connected. Nodes with solid edge indicate intra-hemispheric connectivity [89,90]. Nodes with doted edge indicate inter-hemispheric connectivity.

4. Results

4.1. Overview

In total, 16 participants (15 males, 1 female) participated in the study, including 1 undergraduate student and 15 graduate students. All participants were recruited via the university emailing list. Participants were from a variety of disciplines, including computer science, civil engineering, construction science, and other engineering majors. We performed a power analysis for the paired test [91]. We found that 16 participants can achieve a power of 80% and a level of significance of 5% (two sided), for detecting an effect size of 0.8 between pairs. In addition, we found that many existing neural research studies have been based on a similar sample size, such as [87,92]. As a result, the selection of the sample size was also following the literature standard. Participants were also asked to report their previous knowledge of the HVAC system and none of the participants have previous pipe maintenance knowledge. None of the participants felt VR sickness while performing the pipe operation task in the virtual environment. The

Table 3
Brodmann Area and Functions.

ROIs	Brodmann Areas and Functions
Dorsal-lateral prefrontal cortex (DLPFC)	Area 8, 9; motor planning, complex mental activities
Premotor cortex (PMC)	Area 6; planning of complex and coordinated motor movements
Primary motor cortex (M1)	Area 4; motor movements

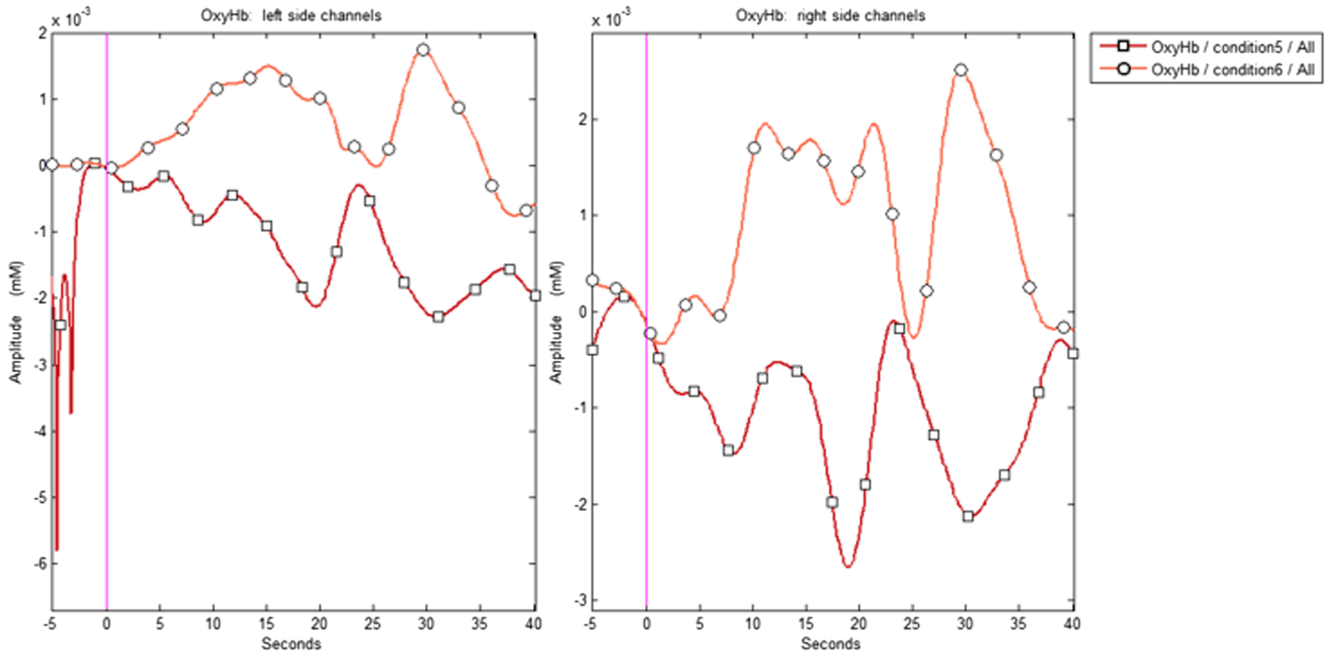

Fig. 6. A sample of participant #1′s post-processed fNIRS data across different conditions. The red line represents the normal condition and the orange line represents the stressful condition. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

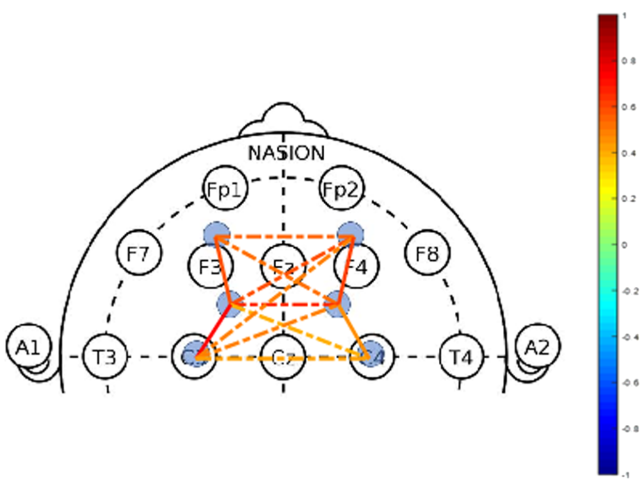

Fig. 7. A Sample of Functional Connectivity based on fNIRS data.

experimental procedure took approximately 40–60 min for each participant. Each participant got a $15 Amazon gift card after they finished the experiment. Fig. 8 shows one participant was using the proposed VR system.

4.2. Task performance

First, we evaluated whether task performance was sufficiently different between normal and stressful conditions. Two task performance indicators were used including pipe operation accuracy (%) and operation time (s). We evaluated participants' operation accuracy by checking each operation trail by dividing the pipe model into five sections. For the pipe operation accuracy as shown in Fig. 9(a), we used a matched pairs *t*-test to evaluate the pipe operation accuracy for each participant. We found a significant difference ($p = 0.014 < 0.05$) in pipe maintenance accuracy between the two conditions. The results indicated that, on average, the participants performed 26.43% better in the normal scenario than the stressful scenario. Thus, the stressful virtual environment could make participants perform worse in this experiment. As for the operation time as shown in Fig. 9(b), the matched pairs *t*-test did not find any significant difference ($p = 0.7144 > 0.05$) in the operation time between the two conditions. The results reveal that the stressful scenario did not have a strong impact on the participants' operation time.

4.3. Visual scan pattern

Second, we evaluated whether the visual scan pattern was sufficiently different between normal and stressful conditions. As mentioned in the methodology part, we extracted gaze ApEn as a feature to evaluate participants' general attention patterns in the virtual environment. A higher value of gaze movement entropy indicates more irregularity and unpredictability of gaze movement, suggesting that participants just randomly visual scan in the environment. On the other hand, a lower value of gaze movement entropy shows a more regular and relatively stable gaze focus transitions. In this study, we evaluated participants' visual scan patterns in the horizontal direction (x-axis) and vertical direction (y-axis). As illustrated in Fig. 10(a), we did not find any significant difference in ApEn at horizontal direction between the normal and stressful conditions by using a two-sample *t*-test ($p = 0.2345 > 0.05$). As illustrated in Fig. 10(b), we found a significant difference in ApEn at vertical direction between the normal and stressful conditions by using a two-sample *t*-test ($p = 0.0328 < 0.05$). A higher value of ApEn indicates a more frequent gaze scan pattern in the vertical direction when reviewing the instructions. In other words, participants in stressful condition tended to scan information more quickly and repeatedly across different task steps in the vertical direction. In summary, the results confirmed that participants tended to perform different visual scan patterns in different training scenarios. A higher value ApEn may serve as an indicator of high stress level. Since the review instructions were designed to be listed in the vertical direction, participants in this experiment demonstrated a high value of ApEn in the vertical direction. This is supported by Jyotsna and Amudha [73]'s finding that the gaze movement is associated with the stress level. Although we cannot conclude that distinct gaze movement patterns are results of stress scenario. But at least, we shall be able to claim that distinct gaze movement patterns, such as entropy of visual pathways, indicate the participants experienced the stress scenario.

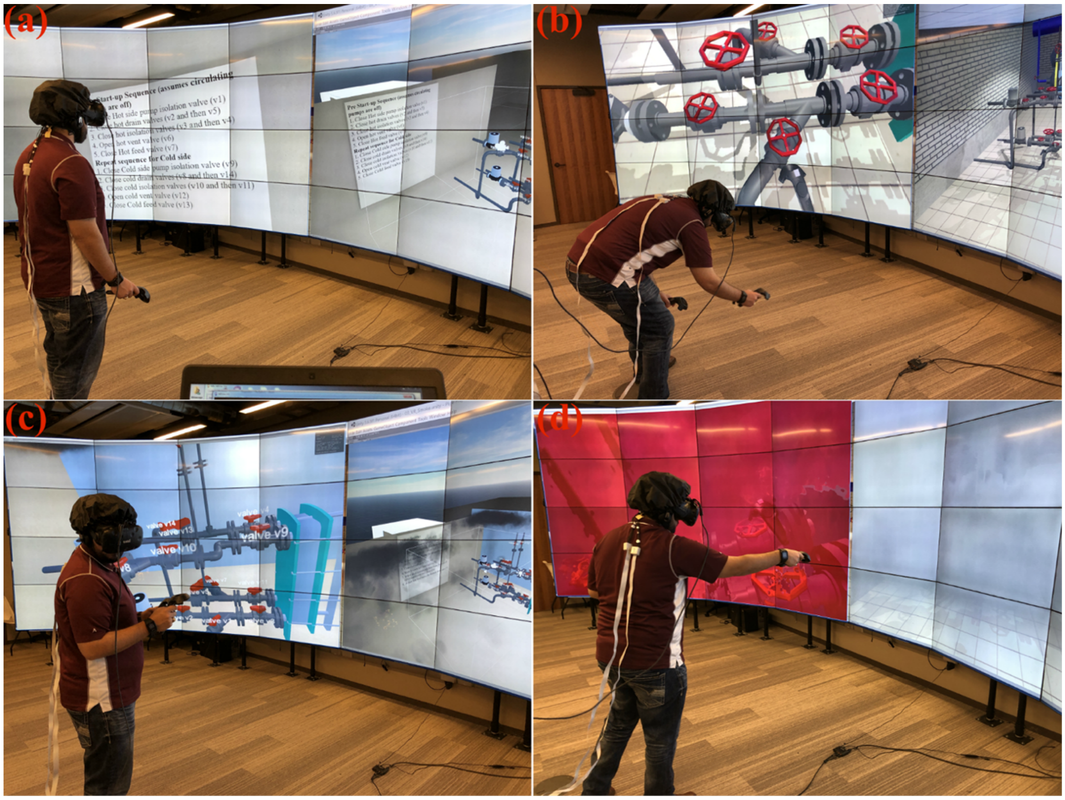

Fig. 8. The participant was using the proposed VR system. (a) review session-normal; (b) operation session-normal; (c) review session-stressful; (d) operation session-stressful.

4.4. Functional connectivity

Third, we analyzed the functional connectivity strength between normal and stressful conditions. Pearson correlations are calculated across all ROIs to find the correlation coefficients. All the calculated R values are then converted to Fisher's z-scores to determine the strength of correlations. Functional connectivity between two regions were determined based on predetermined threshold of 0.4. Fig. 11 illustrates the average Fisher's z-score (functional connectivity strength) across different conditions (normal and stressful) and different phases (early and late).

Significant increases in connectivity strength were observed for the stressful condition compared to the normal condition in Fig. 12. We analyzed the functional connectivity of two phases, including the Early (first 5 training trials) and Late phases (last 5 training trials). Six blue nodes in each graph of Fig. 10 indicate six regions of interest. The color of each line indicates the strength of functional connectivity based on the color scale on the right. Solid lines indicate intra-hemispheric connectivity and dashed lines indicate inter-hemispheric connectivity. The middle column shows significant changes between two groups of sessions. All the lines in the middle column are in dark red (positive correlations) which indicates significantly stronger connectivity for the stressful condition compared to the normal condition. As shown in the result, at the early phase, significant increases of multiple

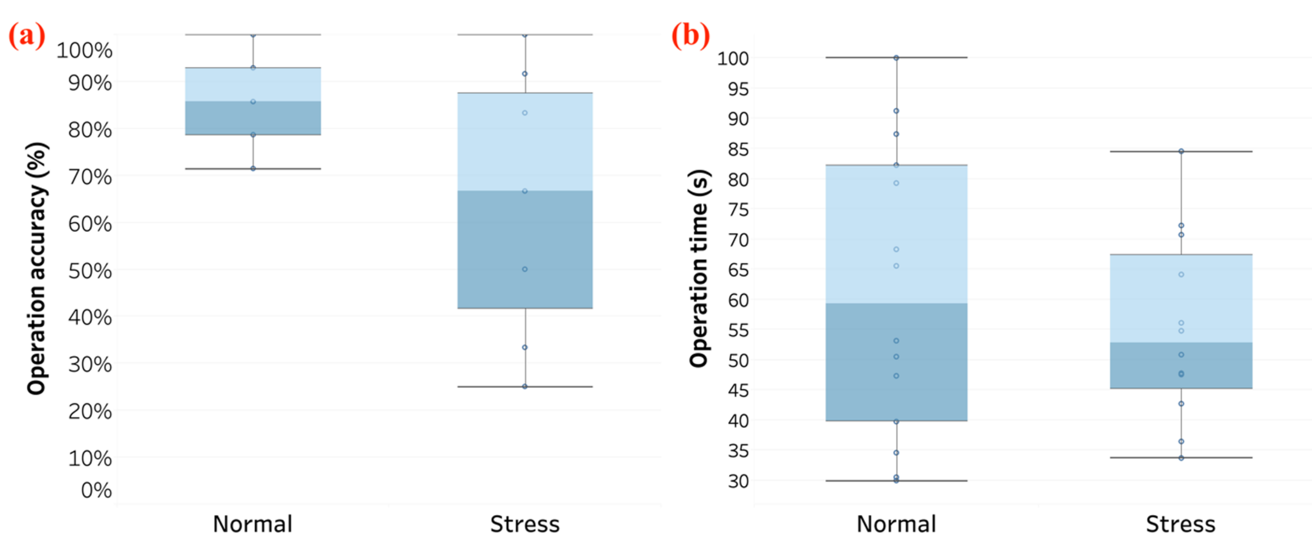

Fig. 9. The results of task performance across different conditions. (a) operation accuracy; (b) operation time.

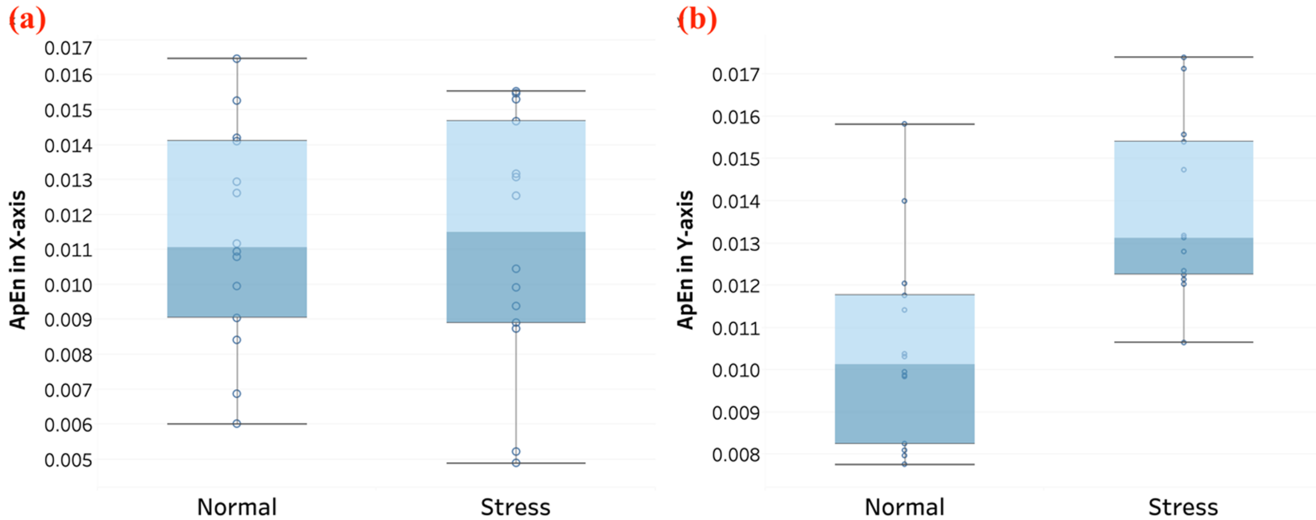

Fig. 10. The results of ApEn across different conditions. (a) horizontal direction (x-axis); (b) vertical direction (y-axis).

interhemispheric connections were identified from normal and to stressful condition by using two-sample *t*-test ($p = 0.005–0.026 < 0.05$). At the late phase, significant increases were identified within the right hemisphere from normal and to stressful condition ($p = 0.015–0.022 < 0.05$). Interhemispheric connection between LM and RM also increased significantly at the late phase ($p = 0.041 < 0.05$). Fig. 13 shows a sample of participant #10′s neuroimage data between normal and stressful conditions.

We also compared the functional connectivity changes between early and late phases as shown in Fig. 14. The color of each line indicates the strength of functional connectivity based on the color scale on the right. Solid lines indicate intra-hemispheric connectivity and dashed lines indicate inter-hemispheric connectivity. The middle column shows significant changes between two groups of sessions. All the lines in the middle column are in dark red (positive changes) which indicates significantly stronger connectivity for the Late phase than for the Early Phase. In the stressful condition, there is no significant changes between early and late phases were observed. However, there

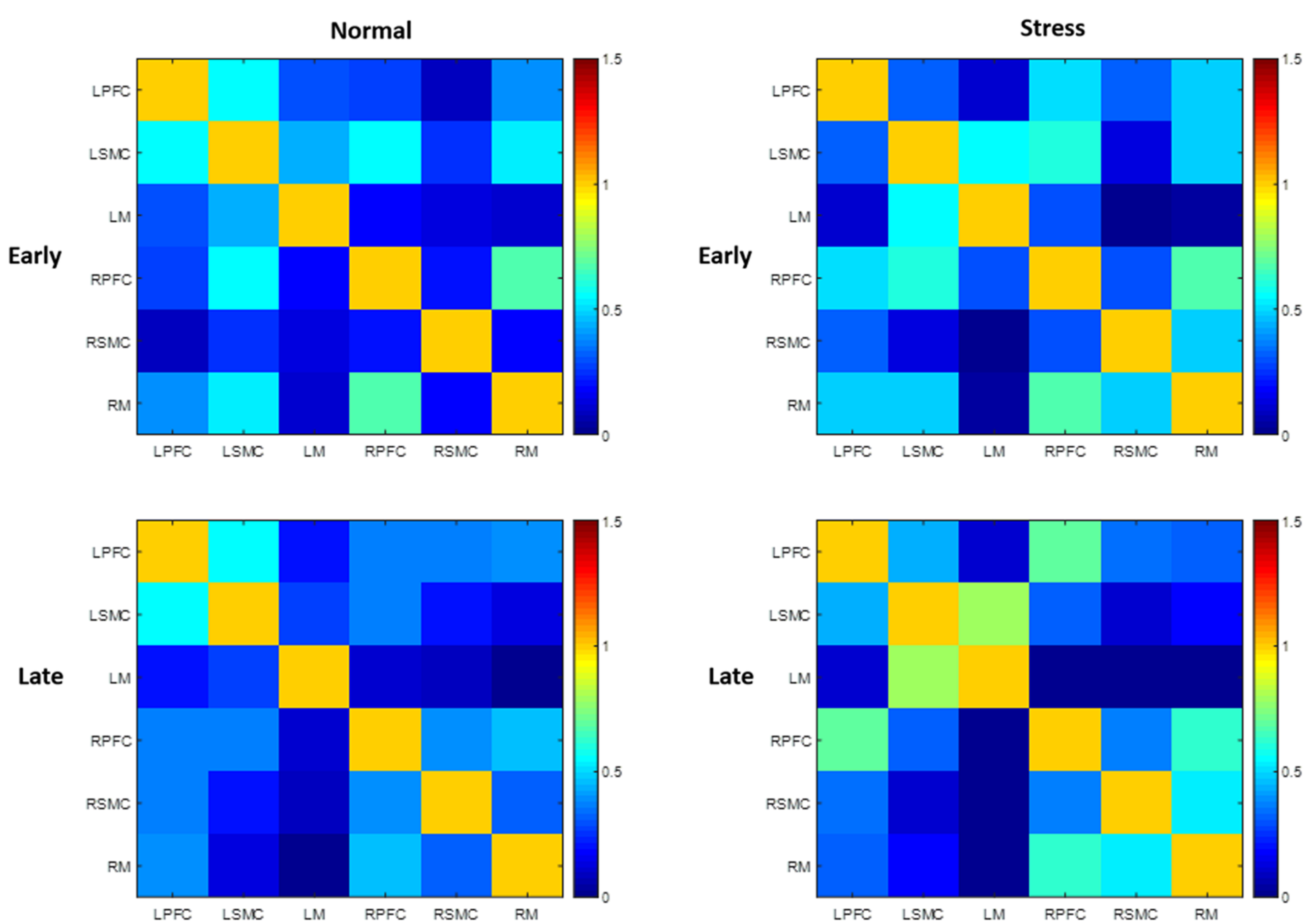

Fig. 11. The average Fisher's z-score (functional connectivity strength) across different conditions (normal and stressful) and different phases (early and late).

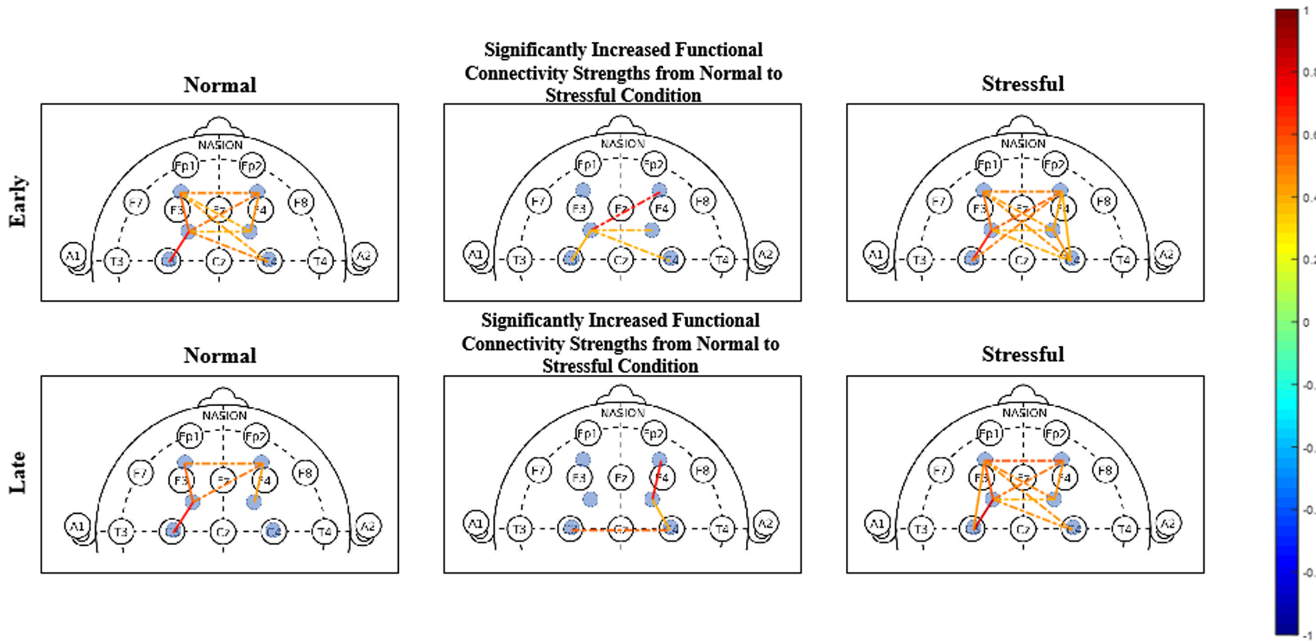

Fig. 12. Functional connectivity maps of the Normal condition (left column), the Stressful condition (right column), and the significantly increased connectivity s from Normal to Stressful Condition (middle column).

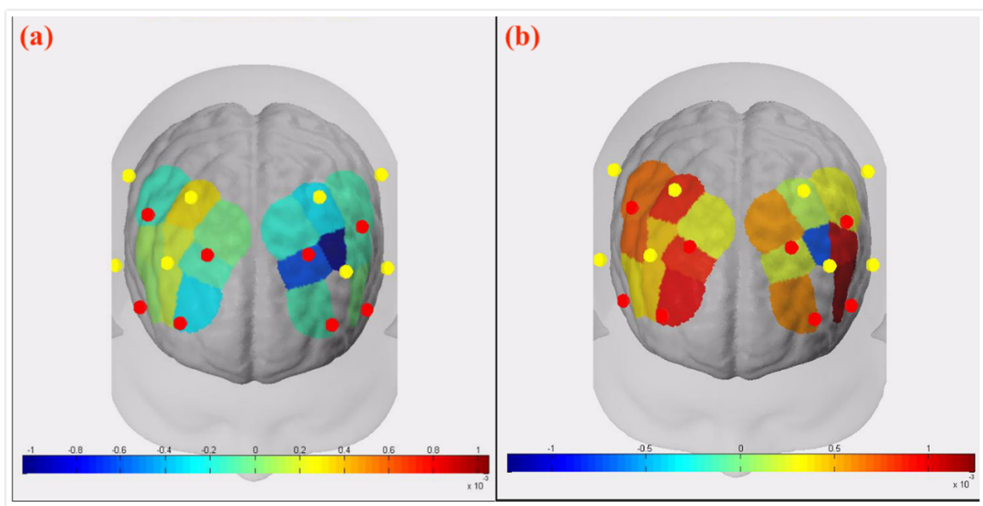

Fig. 13. The participant #10′s neuroimage data between normal and stressful conditions. (a) Normal condition; (b) stressful condition.

are significant increases in the interhemispheric connections in the normal condition. Connectivity among RPMC, RDLPFC, and RM1 were increased significantly from early to late phase by using two-sample *t*-test ($p = 0.005$ and $0.034 < 0.05$).

In summary, the functional connectivity result confirms significant and potentially quantifiable difference between normal and stressful trainings. It constitutes a theoretical foundation for an early warning and performance prediction system based on neural connectivity data. The temporal analysis also shows a significant difference between the early and late phase of the training combined with the neural connectivity pattern difference driven by normal-stress training.

4.5. Correlations between neurophysiological metrics and task performance

The experiment results indicated correlations between a set of neurophysiological measures and the final operation performance (time and accuracy), which set the methodological foundation for a performance early warning and estimating system based on the neurophysiological data during training. First, we found a significant correlation between gaze movement entropy in vertical direction (y-axis) and operation accuracy ($r = -0.388$, $p = 0.019 < 0.05$) as shown in the following Fig. 15. Combining the results we found in Sections 4.2 and 4.3, the participants had lower operation accuracy in the stress condition compared to normal condition and participants had higher of gaze movement entropy in vertical direction in the stress condition compared to normal condition. These results revealed that there was a negative correlation between gaze movement entropy in vertical direction and operation accuracy, which means that the higher value of gaze movement entropy in vertical direction might reduce participants' operation accuracy.

Then, we extracted the peak HbO of six ROI including LDLPFC, RDLPFC, LPMC, RPMC, LM1 and RM1 for each review trail from the fNIRS data as the neural activation features across different conditions. We found significant correlations between RSMC and operation accuracy ($r = 0.205$, $p = 0.019 < 0.05$), between LM and operation accuracy ($r = 0.208$, $p = 0.0176 < 0.05$), and between RM and

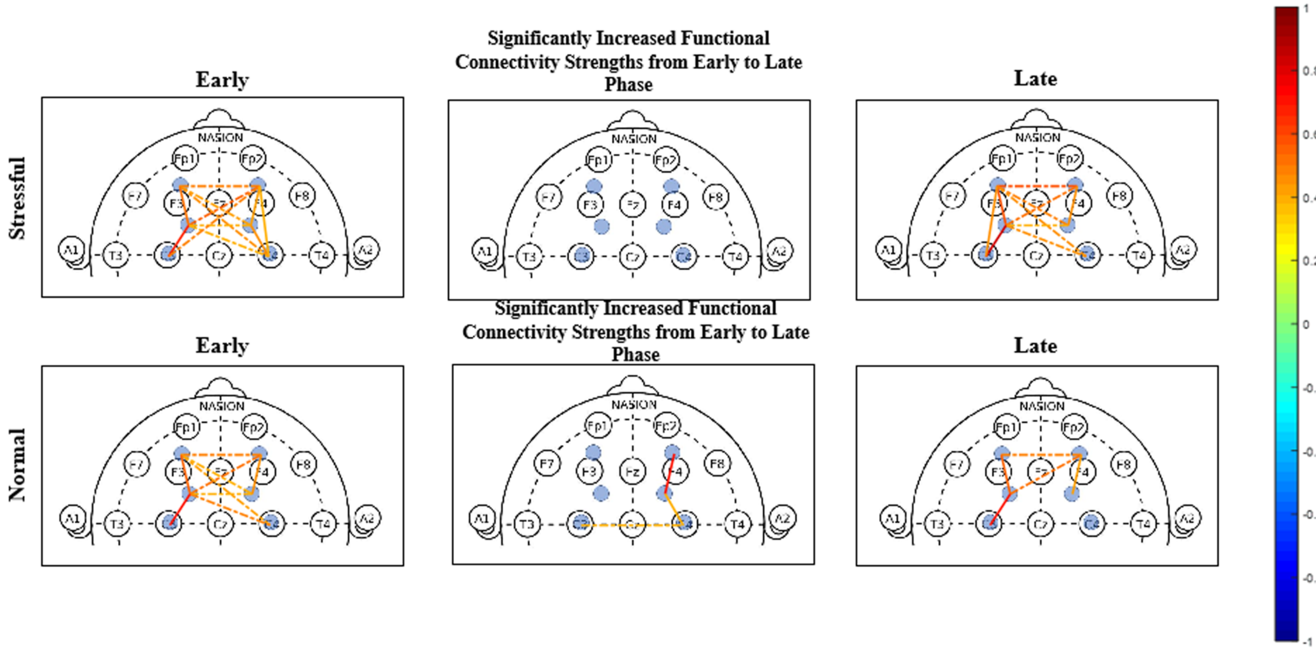

Fig. 14. Functional connectivity maps of the Early phase (first 5 training trials; left column), the Late phase (last 5 training trials; right column), and significantly increased functional connectivity strengths from Early to Late Phase.

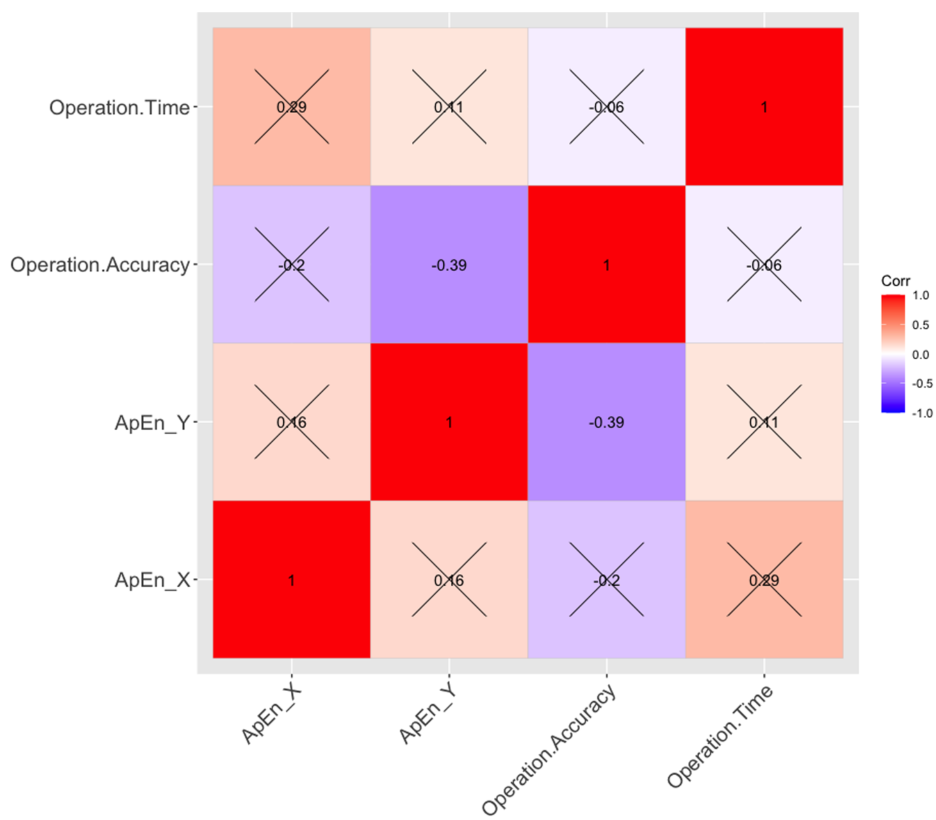

Fig. 15. The correlations between gaze movement entropy and task performance. The boxes crossed out mean that the p-value of the correlations is larger than 0.05.

operation accuracy ($r = -0.196$, $p = 0.0252 < 0.05$) in the normal condition. We did not find any significant correlations between operation accuracy and other ROIs in the normal condition as shown in Fig. 16(a). However, in the stress condition, we detected stronger correlations between the peak Hbo of six ROIs and operation accuracy, we found significant correlations between RSMC and operation accuracy ($r = 0.47$, $p = < 0.001$), between LSMC and operation accuracy ($r = -0.194$, $p = < 0.001$), between RM and operation accuracy ($r = -0.26$, $p < 0.001$), between LPFC and operation accuracy ($r = -0.328$, $p < 0.001$), and between RPFC and operation accuracy ($r = -0.273$, $p < 0.001$) in the stress condition as shown in Fig. 16(b). We detected stronger correlations between the peak HbO of six ROIs and operation accuracy in stress condition compared to normal condition. These results further confirmed that fNIRS has the potential to estimate

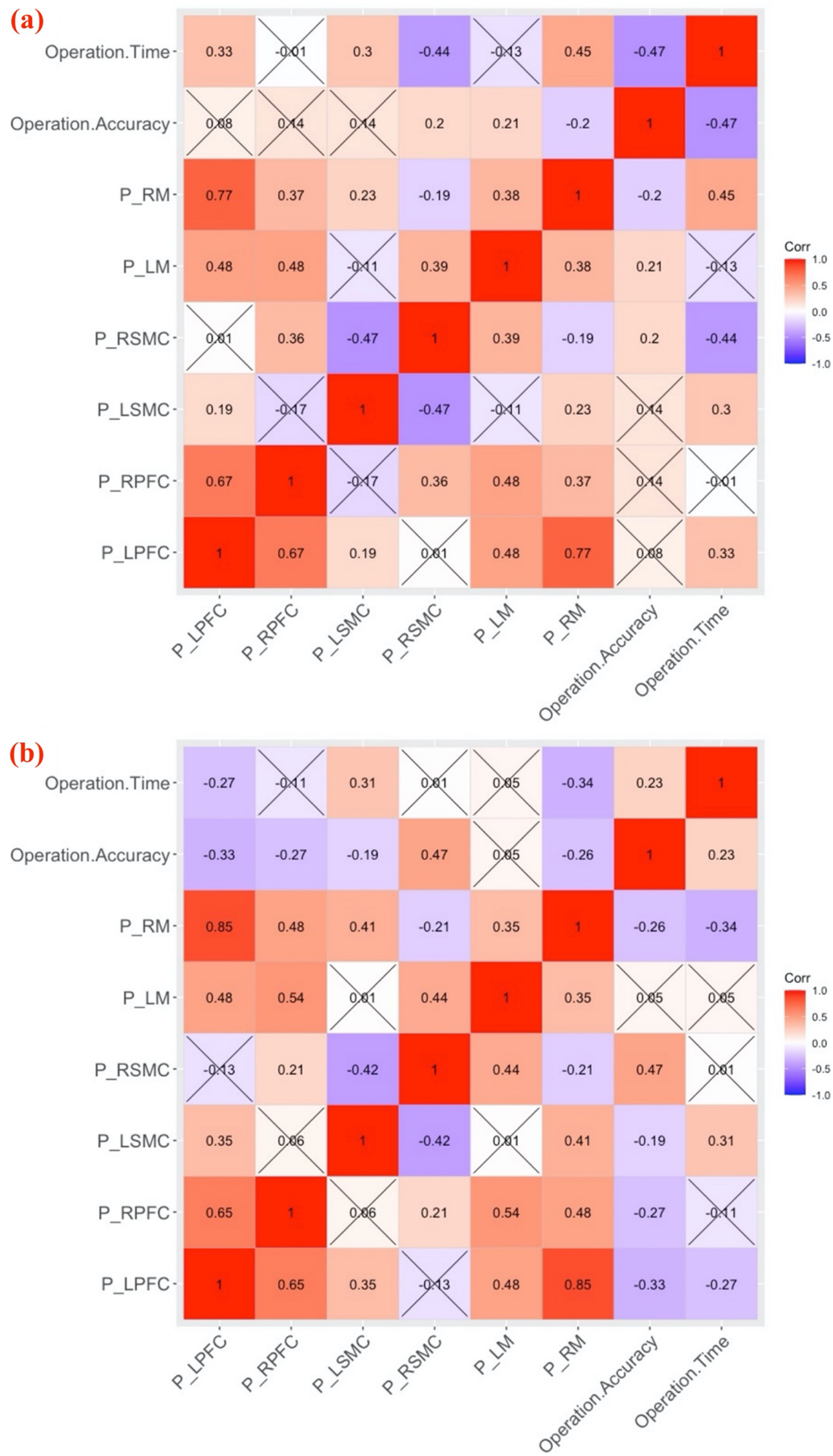

Fig. 16. The correlations between peak Hbo of six ROI and task performance. The boxes crossed out mean that the p-value of the correlations is larger than 0.05.

and project task performance based on training data. It helps us to build a next logic step for using the neurophysiological metrics to estimate final task performance.

4.6. Assessing training outcome

Finally, since the fNIRS data correlated with the task performance in different training scenarios, it is possible to use the fNIRS metrics in the review session to estimate individual's final task performance. We used supervised machine learning method to classify the training results. All the participants have 20 trails of fNIRS data in the review session (10 review trails in normal condition and 10 review trails in stress condition) and all the fNIRS data was labeled by their training outcome of the operation accuracy. The principle of labeling the data is that if the participant achieved 100% operation accuracy, we labeled the data as "satisfactory" otherwise the data was labeled as "unsatisfactory". Our purpose is to create a classification model that can use the fNIRS data in the review session to estimate training outcome. A total of 12 features were selected to classify pipe operation task performance, including 6 fNIRS peak Hbo features (peak Hbo of LDLPFC area, peak Hbo of RDLPFC area, peak Hbo of LPMC area, peak Hbo of RPMC area, peak Hbo of LM1 area, and peak Hbo of RM1 area) and 6 fNIRS average Hbo features (average Hbo of LDLPFC area, average Hbo of RDLPFC area, average Hbo of LPMC area, average Hbo of RPMC area, average Hbo of LM area, and average Hbo of RM area). For the pre-processing of the fNIRS data for the classification model, we followed the fNIRS data preprocessing pipeline proposed by [87]. All the motion artefact of the fNIRS data for each ROIs was removed by the wavelet interpolation and band pass filter (0.01–0.5 Hz) [9]. All the fNIRS features were normalized using the z-score method [92,93]. We compared the classification models by using different machine learning classification algorithms including Decision Tree (DT), Random Forest (RF), K-Nearest Neighbors (KNN), Nominal Logistical Regression (LR), and Naïve Bayes (NB). We also used the K-Folds method ($k = 10$) to cross-validate our classification results. We selected accuracy, precision, recall, F-score, and ROC area to evaluate the classification performance. Table 4 shows the summary of classification models before feature selection. These results indicated that Random Forest can provide more accuracy classification (78.08%) than other classification models. In addition, we further used wrapper method with sequential backward feature selection (SBS) method to find the optimal feature subset to achieve the better classification performance. The K-Folds method ($k = 10$) was also used to cross-validate the feature selection results. Table 5 shows the summary of classification models after feature selection and (n/10) in Table 5 indicates how many times this feature was selected by the wrapper method with sequential backward feature selection during 10-fold. These results also indicated that Random Forest can provide more accuracy classification (80.38%) than other classification models with 6 selected features including LM (Peak Hbo) (10/10), RSMC (Peak Hbo) (9/10), LSMC (Peak Hbo) (8/10), LPFC (Peak Hbo) (5/10), RSMC (Average Hbo) (4/10), and LM (Average Hbo) (4/10). The classification accuracy was improved from 78.08% to 80.38% after feature selection. These results further suggested that the fNIRS features in review session can be used to estimate future task performance during different training scenarios. Combing with the results of functional connectivity, task performance assessment can be used for early monitoring of individual's task performance in pipe maintenance training and provide interactive assessments for performance and learning.

5. Discussion

These experiment results revealed several important findings regarding the stress during shutdown maintenance training. First, our experiment results confirmed that the presence of stressors during learning or training process did affect complex knowledge-based tasks (e.g. pipe maintenance) in a negative way. Performance assessment indicated that although there was no significant difference in task completion time between the two conditions, the performance accuracy under stressful training was much lower. It suggested that participants' cognitive process might have been affected by the stress during the review phase. This result further indicated the importance of assessing task performance under stress training scenarios.

Second, we explored the promising possibility of using neurophysiological measures to assess workers' stress status and task performance under different training scenarios. Stress level assessment and task performance assessment have been recently identified as promising research directions in the construction research area. There is a pressing need for a robust assessment method to predict workers' task performance under different working scenarios. This study echoes recent studies proposed using wireless EEG device (EMOTIV) and other wearable biosensors to assess workers' mental stress on site [16,17,94]. The novelty of this study pertains to using neurophysiological measures during the training to estimate and classify the final performance in industrial operations. We used an alternative portable neuroimage device - fNIRS integrated with the eye-tracking VR system to collect high-quality temporal neuroimaging data and eye-tracking data during the simulated VR training scenarios. Based on the data, we found that several supervised machine learning methods were able to classify the participants' final task performance based on the neurophysiological data in the review phase. It provides innovative methods and knowledge about the role of neural analysis in training evaluation. The novel contribution of this study also lies in the added knowledge about stress in training, which helps scholars and practitioners better understand and leverage stressors to stimulate desired neural activations during training. Specifically, the fNIRS data analysis confirmed significant differences in terms of neurofunctional connectivity between the normal and stressful training conditions. Participants demonstrated a stronger interhemispheric connectivity in the early phase of stress training. We also detected significant increases within the right hemisphere between the normal and stressful conditions in the late training phase. Interhemispheric connections between LM and RM were also increased significantly in the late phase. In summary, the participants who were in the stressful condition tended to have more interhemispheric connections between the left and right hemispheres. The differences in connectivity can be quantified as a leading indicator of detecting the presence of stress. In addition, we also tested the correlations between participants' task performance and activation level of 6 ROIs' fNIRS data. We found significant correlations between RSMC and operation accuracy ($r = 0.205$, $p = 0.019 < 0.05$), between LM and operation accuracy ($r = 0.208$, $p = 0.0176 < 0.05$), and between RM and operation accuracy ($r = -0.196$, $p = 0.0252 < 0.05$) in the normal condition. We detected more correlations between the peak Hbo of six ROIs and operation accuracy in the stressful condition, we found significant correlations between RSMC and operation accuracy ($r = 0.47$, $p = < 0.001$), between LSMC and operation accuracy ($r = -0.194$, $p = < 0.001$), between RM and operation accuracy ($r = -0.26$, $p < 0.001$), between LPFC and operation accuracy ($r = -0.328$, $p < 0.001$), and between RPFC and operation accuracy ($r = -0.273$, $p < 0.001$). These results further confirmed that fNIRS has the potential to assess the task performance in different training scenarios. Based on the correlation results, we tested several supervised machine learning classification models that uses the fNIRS data in the review session to early assess individual's task performance and we utilized k-fold ($k = 10$) cross-validation method to validate our results. The Random Forest classification model achieved an average 80.38% classification accuracy after feature selection to assess participants' training outcome compared to other classification models.

At last, in addition to fNIRS data, this study also identified a neurophysiological metrics as potential predictor - gaze movement patterns. We found a significant correlation between gaze movement

Table 4
The summary of classification models before feature selection.

Classification Algorithm	Accuracy	Precision	Recall	F-score	ROC Area
Random Forest	78.08%	77.10%	78.10%	0.746	0.759
Logistical Regression	75.38%	72.90%	75.40%	0.727	0.709
Decision Tree	73.08%	68.60%	73.10%	0.679	0.517
K-Nearest Neighbors	68.85%	60.40%	68.80%	0.688	0.498
Naïve Bayes	66.92%	68.40%	66.90%	0.676	0.686

Table 5
The summary of classification models after feature selection.

Classification Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F-score	ROC Area	Selected Features
Random Forest	80.38%	80.1%	80.4%	0.777	0.687	• LM (Peak Hbo) (10/10) • RSMC (Peak Hbo) (9/10) • LSMC (Peak Hbo) (8/10) • LPFC (Peak Hbo) (5/10) • RSMC (Average Hbo) (4/10) • LM (Average Hbo) (4/10)
Logistical Regression	79.61%	82.8%	79.6%	0.75	0.547	• RSMC (Peak Hbo) (10/10) • LM (Peak Hbo) (9/10) • LPFC (Peak Hbo) (8/10) • LSMC (Peak Hbo) (7/10) • RSMC (Average Hbo) (7/10) • RPFC (Average Hbo) (2/10) • LSMC (Average Hbo) (2/10)
Decision Tree	79.23%	79.5%	79.2%	0.755	0.667	• RM (Peak Hbo) (6/10) • RPFC (Peak Hbo) (5/10) • LM (Peak Hbo) (5/10)
K-Nearest Neighbors	79.2%	77.9%	79.2%	0.777	0.682	• LM (Peak Hbo) (9/10) • RM (Peak Hbo) (9/10) • LM (Average Hbo) (8/10) • RSMC (Peak Hbo) (7/10) • RPFC (Peak Hbo) (4/10) • LPFC (Peak Hbo) (3/10)
Naïve Bayes	78.5%	77.8%	78.5%	0.749	0.701	• RPFC (Peak Hbo) (10/10) • RSMC (Peak Hbo) (10/10) • LM (Peak Hbo) (10/10) • LSMC (Peak Hbo) (6/10) • RM (Average Hbo) (6/10) • LM (Average Hbo) (1/10)

entropy in vertical direction (y-axis) and operation accuracy ($r = -0.388$, $p = 0.019 < 0.05$). Specifically, increased entropy of gaze movement in vertical direction may be an indicator of stress. This result suggests that participants tended to vertically scan information more quickly and repeatedly across different task steps in stressful condition compared to normal condition. These findings provide empirical evidence that the neurophysiological features can be used to develop a task performance assessment model under different training scenarios.

6. Conclusions

This study proposed a neurophysiological approach to assess workers' stress status and training outcomes under normal and stressful training scenarios. A VR system integrated with the eye-tracking function was developed to simulate different training scenarios. A neuroimaging device – fNIRS was used to collect user's brain activities by measuring hemodynamic responses associated with neuron behavior. A pipe maintenance task of replacing a plate heat exchanger was selected as the shutdown maintenance training scenario. A human-subject experiment was conducted to test the feasibility and usability of the neurocentric approach. Our experiment results indicated that stressful training scenario had a strong impact on participants' neural connectivity patterns and gaze movement patterns in vertical direction during training scenarios, and finally negatively affected participants' task performance. We also found that the task performance was correlated with neurophysiological features including gaze movement entropy and fNIRS data. At last, we tested several supervised machine learning classification models that uses the fNIRS data in the review session to early assess individual's task performance and we utilized k-fold ($k = 10$) cross-validation method to validate our results. The Random Forest classification model achieved an average 80.38% classification accuracy after feature selection to assess participants' training outcome compared to other classification models. The focus of this study does not only rely on developing a neurocentric training approach via the VR system, but it also aims to make a theoretical contribution of how different training scenarios affect training quality related to neural connectivity. These findings are expected to provide more evidence about an early performance warning and prediction system based on a hybrid neurophysiological measure method, inspiring the design of a cognition-driven personalized training system for industrial workers in the future.

Several research limitations still need to be addressed in the future research agenda. First, there was gender bias in this experiment (15 males, 1 female). Gender difference might have influence on participants' task performance in this experiment. According to previous literature, gender is known to influence both physical and cognitive task performance. Male and female differ in the way when handling stress [95,96]. Previous literatures also have found gender differences in both neural signature and motor task performance [97,98]. Male were found to exhibit higher activation in prefrontal cortex during stress compares to women [99]. However, exploring gender differences in both task performance and neural activities is not the focus of this study. We admitted this gender bias is one of the limitations of this study. We will investigate the effects of diversity of users including disciplines, backgrounds, ages, and gender difference in our future studies. Second, this study was conducted in a well-controlled laboratory environment. In real world, construction sites and construction operations are more complex and unpredictable. Thus, more complex and dynamic scenarios should be tested in future research. At last, we only proposed the implementation of using fNIRS to assess individual's task performance under different training scenarios. Multimodal neuroimaging studies found that compensate EEG system with fNIRS system have been explored to deliver complimentary data on task related neural activations and improve motor task classification model performance [42,43]. A comparison study of using EEG and fNIRS was suggested in our future studies.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This material is supported by the National Science Foundation (NSF) under Grants 1937878, as well as the National Institute of Standards and Technology (NIST) under Grant 60NANB18D152. Any opinions, findings, conclusions, or recommendations expressed in this article are those of the authors and do not reflect the views of the NSF and NIST.

References

[1] S.O. Duffuaa, M. Ben Daya, Turnaround maintenance in petrochemical industry: practices and suggested improvements, J. Qual. Mainten. Eng. 10 (2004) 184–190.
[2] ASCE, 2017 Infrastructure Report Card: Investment, 2018.
[3] E.I.A. EIA, Planned Refinery Outages in the United States: December 2017 – June 2018, 2018.
[4] O.G. IQ, Shutdowns and Turnarounds in the Oil and Gas Industry, 2014.
[5] D.T. Daley, The Little Black Book of Maintenance Excellence, Industrial Press Inc., 2008.
[6] M. Ben-Daya, D. Ait-Kadi, S.O. Duffuaa, J. Knezevic, A. Raouf, Handbook of Maintenance Management and Engineering, Springer, 2009.
[7] Wikipedia, List of natural gas and oil production accidents in the United States, 2018.
[8] Y. Pinchuk, The Dangers of Refinery Turnarounds, 2017.
[9] T.S. Abdelhamid, J.G. Everett, Identifying root causes of construction accidents, J. Construct. Eng. Manage. 126 (2000) 52–60.
[10] J. Garrett, J. Teizer, Human factors analysis classification system relating to human error awareness taxonomy in construction safety, J. Construct. Eng. Manage. 135 (2009) 754–763.
[11] X. Huang, J. Hinze, Analysis of construction worker fall accidents, J. Construct. Eng. Manage. 129 (2003) 262–271.
[12] R.P. Gordon, The contribution of human factors to accidents in the offshore oil industry, Reliab. Eng. Syst. Saf. 61 (1998) 95–108.
[13] N. Meshkati, Human factors in large-scale technological systems' accidents: Three Mile Island, Bhopal, Chernobyl, Indust. Crisis Quart. 5 (1991) 133–154.
[14] C. Viswesvaran, D.S. Ones, Perspectives on models of job performance, Int. J. Sel. Assess. 8 (2000) 216–226.
[15] P.C. Kyllonen, D.F. Lohman, R.E. Snow, Effects of aptitudes, strategy training, and task facets on spatial task performance, J. Educ. Psychol. 76 (1984) 130.
[16] H. Jebelli, Wearable Biosensors to Understand Construction Workers' Mental and Physical Stress, 2019.
[17] H. Jebelli, B. Choi, S. Lee, Application of wearable biosensors to construction sites. I: Assessing workers' stress, J. Construct. Eng. Manage. 145 (2019) 04019079.
[18] F. Campbell, Occupational Stress in the Construction Industry, Chartered Institute of Building, UK, Berkshire, 2006.
[19] N.S. Haynes, P.E. Love, Psychological adjustment and coping among construction project managers, Construct. Manage. Econ. 22 (2004) 129–140.
[20] J.K. Sluiter, High-demand jobs: age-related diversity in work ability? Appl. Ergon. 37 (2006) 429–440.
[21] P. Bowen, P. Edwards, H. Lingard, K. Cattell, Occupational stress and job demand, control and support factors among construction project consultants, Int. J. Project Manage. 32 (2014) 1273–1284.
[22] O.O. Abbe, C.M. Harvey, L.H. Ikuma, F. Aghazadeh, Modeling the relationship between occupational stressors, psychosocial/physical symptoms and injuries in the construction industry, Int. J. Ind. Ergon. 41 (2011) 106–117.
[23] E. Russell, G. Koren, M. Rieder, S. Van Uum, Hair cortisol as a biological marker of chronic stress: current status, future directions and unanswered questions, Psychoneuroendocrinology 37 (2012) 589–601.
[24] R. Castaldo, P. Melillo, U. Bracale, M. Caserta, M. Triassi, L. Pecchia, Acute mental stress assessment via short term HRV analysis in healthy adults: a systematic review with meta-analysis, Biomed. Signal Process. Control 18 (2015) 370–377.
[25] G.G. Berntson, J.T. Cacioppo, Heart rate variability: stress and psychiatric conditions, Dyn. Electrocardiogr. (2004) 57–64.
[26] Y. Choi, M. Kim, C. Chun, Measurement of occupants' stress based on electroencephalograms (EEG) in twelve combined environments, Build. Environ. 88 (2015) 65–72.
[27] J. Alonso, S. Romero, M. Ballester, R. Antonijoan, M. Mañanas, Stress assessment based on EEG univariate features and functional connectivity measures, Physiol. Meas. 36 (2015) 1351.
[28] S.-H. Seo, J.-T. Lee, Stress and EEG, Converg. Hybrid Inf. Technol. 1 (2010) 413–424.
[29] J.T. Gwin, K. Gramann, S. Makeig, D.P. Ferris, Removal of movement artifact from high-density EEG recorded during walking and running, J. Neurophysiol. 103 (2010) 3526–3534.
[30] A.S. Oliveira, B.R. Schlink, W.D. Hairston, P. König, D.P. Ferris, Induction and separation of motion artifacts in EEG data using a mobile phantom head device, J. Neural Eng. 13 (2016) 036014.
[31] R.J. Croft, R.J. Barry, Removal of ocular artifact from the EEG: a review, Neurophysiologie Clinique/Clin. Neurophysiol. 30 (2000) 5–19.
[32] D.A. Pizzagalli, Electroencephalography and high-density electrophysiological source localization, Handbook Psychophysiol. 3 (2007) 56–84.
[33] S. Duschek, R. Schandry, Functional transcranial Doppler sonography as a tool in psychophysiological research, Psychophysiology 40 (2003) 436–454.
[34] R.K. Mehta, R. Parasuraman, Neuroergonomics: a review of applications to physical and cognitive work, Front. Hum. Neurosci. 7 (2013) 889.
[35] T. Nozawa, T. Kondo, A comparison of artifact reduction methods for real-time analysis of fNIRS data, Symposium on Human Interface, Springer, 2009, pp. 413–422.
[36] T. Zama, S. Shimada, Simultaneous measurement of electroencephalography and near-infrared spectroscopy during voluntary motor preparation, Sci. Rep. 5 (2015) 1–9.
[37] Y. Zhu, C. Rodriguez-Paras, J. Rhee, R.K. Mehta, Methodological approaches and recommendations for functional near-infrared spectroscopy applications in HF/E research, Hum. Factors (2019) 0018720819845275.
[38] M. Ferrari, V. Quaresima, A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application, Neuroimage 63 (2012) 921–935.
[39] H. Bortfeld, E. Wruck, D.A. Boas, Assessing infants' cortical response to speech using near-infrared spectroscopy, Neuroimage 34 (2007) 407–415.
[40] Y. Minagawa-Kawai, H. Van Der Lely, F. Ramus, Y. Sato, R. Mazuka, E. Dupoux, Optical brain imaging reveals general auditory and language-specific processing in early infant development, Cereb. Cortex 21 (2011) 254–261.
[41] R.N. Aslin, J. Mehler, Near-infrared spectroscopy for functional studies of brain activity in human infants: promise, prospects, and challenges, J. Biomed. Opt. 10 (2005) 011009.
[42] M. Muthalib, A.R. Anwar, S. Perrey, M. Dat, A. Galka, S. Wolff, U. Heute, G. Deuschl, J. Raethjen, M. Muthuraman, Multimodal integration of fNIRS, fMRI and EEG neuroimaging, Clin. Neurophysiol. 124 (2013) 2060–2062.
[43] S. Fazli, J. Mehnert, J. Steinbrink, G. Curio, A. Villringer, K.-R. Müller, B. Blankertz, Enhanced performance by a hybrid NIRS–EEG brain computer interface, Neuroimage 59 (2012) 519–529.
[44] B. Seraglia, L. Gamberini, K. Priftis, P. Scatturin, M. Martinelli, S. Cutini, An exploratory fNIRS study with immersive virtual reality: a new method for technical implementation, Front. Hum. Neurosci. 5 (2011) 176.
[45] L. Xu, B. Wang, G. Xu, W. Wang, Z. Liu, Z. Li, Functional connectivity analysis using fNIRS in healthy subjects during prolonged simulated driving, Neurosci. Lett. 640 (2017) 21–28.
[46] D. Dong, L.K. Wong, Z. Luo, Assessment of prospective memory using fNIRS in immersive virtual reality environment, J. Behav. Brain Sci. 7 (2017) 247.
[47] S.B. Moro, S. Bisconti, M. Muthalib, M. Spezialetti, S. Cutini, M. Ferrari, G. Placidi, V. Quaresima, A semi-immersive virtual reality incremental swing balance task activates prefrontal cortex: a functional near-infrared spectroscopy study, Neuroimage 85 (2014) 451–460.
[48] M. Hu, T. Shealy, Application of functional near-infrared spectroscopy to measure engineering decision-making and design cognition: literature review and synthesis of methods, J. Comput. Civil Eng. 33 (2019) 04019034.
[49] M. Hu, T. Shealy, J. Grohs, R. Panneton, Empirical evidence that concept mapping reduces neurocognitive effort during concept generation for sustainability, J. Cleaner Prod. 238 (2019) 117815.
[50] M. Hu, T. Shealy, M. Hallowell, D. Hardison, Advancing construction hazard recognition through neuroscience: measuring cognitive response to hazards using functional near infrared spectroscopy, Construct. Res. Congress 2018 (2018) 134–143.
[51] J. Du, Q. Zhu, Y. Shi, Q. Wang, Y. Lin, D. Zhao, Cognition digital twins for personalized information systems of smart cities: proof of concept, J. Manage. Eng. 36 (2020) 04019052.
[52] A. Berger, F. Horst, S. Müller, F. Steinberg, M. Doppelmayr, Current state and future prospects of EEG and fNIRS in robot-assisted gait rehabilitation: a brief review, Front. Hum. Neurosci. 13 (2019) 172.
[53] T.C. Hankins, G.F. Wilson, A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight, Aviat. Space Environ. Med. 69 (1998) 360–367.
[54] R. McKendrick, R. Mehta, H. Ayaz, M. Scheldrup, R. Parasuraman, Prefrontal hemodynamics of physical activity and environmental complexity during cognitive work, Hum. Factors 59 (2017) 147–162.
[55] R. McKendrick, R. Parasuraman, H. Ayaz, Wearable functional near infrared spectroscopy (fNIRS) and transcranial direct current stimulation (tDCS): expanding vistas for neurocognitive augmentation, Front. Syst. Neurosci. 9 (2015) 27.
[56] J. Du, Y. Shi, Z. Zou, D. Zhao, CoVR: cloud-based multiuser virtual reality headset system for project communication of remote users, J. Construct. Eng. Manage. 144 (2017) 04017109.
[57] J. Du, Z. Zou, Y. Shi, D. Zhao, Zero latency: real-time synchronization of BIM data in virtual reality for collaborative decision-making, Autom. Constr. 85 (2018) 51–64.
[58] Y. Shi, J. Du, S. Lavy, D. Zhao, A multiuser shared virtual environment for facility management, Procedia Eng. 145 (2016) 120–127.
[59] Y. Shi, J. Du, C.R. Ahn, E. Ragan, Impact assessment of reinforced learning methods on construction workers' fall risk behavior using virtual reality, Autom. Constr. 104 (2019) 197–214.
[60] Tobii, Tobii Pro VR Integration, 2019.
[61] Tobii, How do Tobii Eye Trackers work?, 2019.
[62] Tobii, Tobii Pro SDK, 2019.
[63] AlfaLaval, Instruction Manual – Plate Heat Exchangers, 2016.
[64] S.M. Pincus, Approximate entropy as a measure of system complexity, Proc. Natl. Acad. Sci. 88 (1991) 2297–2301.
[65] S.M. Pincus, I.M. Gladstone, R.A. Ehrenkranz, A regularity statistic for medical data analysis, J. Clin. Monit. 7 (1991) 335–345.
[66] D.E. Lake, J.S. Richman, M.P. Griffin, J.R. Moorman, Sample entropy analysis of neonatal heart rate variability, Am. J. Physiol.-Regul. Integr. Comp. Physiol. 283

(2002) R789–R797.
[67] S.M. Pincus, R.R. Viscarello, Approximate entropy: a regularity measure for fetal heart rate analysis, Obstet. Gynecol. 79 (1992) 249–255.
[68] V. Srinivasan, C. Eswaran, N. Sriraam, Approximate entropy-based epileptic EEG detection using artificial neural networks, IEEE Trans. Inf Technol. Biomed. 11 (2007) 288–295.
[69] H. Ocak, Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy, Expert Syst. Appl. 36 (2009) 2027–2036.
[70] S.M. Pincus, D.L. Keefe, Quantification of hormone pulsatility via an approximate entropy algorithm, Am. J. Physiol.-Endocrinol. Metab. 262 (1992) E741–E754.
[71] S.M. Pincus, Approximate entropy as a measure of irregularity for psychiatric serial metrics, Bipolar Disord. 8 (2006) 430–440.
[72] T. Hartley, E.A. Maguire, H.J. Spiers, N. Burgess, The well-worn route and the path less traveled: distinct neural bases of route following and wayfinding in humans, Neuron 37 (2003) 877–888.
[73] C. Jyotsna, J. Amudha, Eye gaze as an indicator for stress level analysis in students, in: 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 2018, pp. 1588–1593.
[74] D. Rosenbaum, M. Thomas, P. Hilsendegen, F.G. Metzger, F.B. Haeussinger, H.-C. Nuerk, A.J. Fallgatter, V. Nieratschker, A.-C. Ehlis, Stress-related dysfunction of the right inferior frontal cortex in high ruminators: an fNIRS study, NeuroImage: Clin. 18 (2018) 510–517.
[75] R. Holtzer, C. Schoen, E. Demetriou, J.R. Mahoney, M. Izzetoglu, C. Wang, J. Verghese, Stress and gender effects on prefrontal cortex oxygenation levels assessed during single and dual-task walking conditions, Eur. J. Neurosci. 45 (2017) 660–670.
[76] S. Jahani, A.L. Fantana, D. Harper, J.M. Ellison, D.A. Boas, B.P. Forester, M.A. Yücel, fNIRS can robustly measure brain activity during memory encoding and retrieval in healthy subjects, Sci. Rep. 7 (2017) 1–14.
[77] M. Hatakenaka, I. Miyai, M. Mihara, S. Sakoda, K. Kubota, Frontal regions involved in learning of motor skill—a functional NIRS study, Neuroimage 34 (2007) 109–116.
[78] M. Suzuki, I. Miyai, T. Ono, K. Kubota, Activities in the frontal cortex and gait performance are modulated by preparation. An fNIRS study, Neuroimage 39 (2008) 600–607.
[79] I. Miyai, H. Yagura, I. Oda, I. Konishi, H. Eda, T. Suzuki, K. Kubota, Premotor cortex is involved in restoration of gait in stroke, Ann. Neurol. 52 (2002) 188–194.
[80] B. Biswal, F. Zerrin Yetkin, V.M. Haughton, J.S. Hyde, Functional connectivity in the motor cortex of resting human brain using echo-planar MRI, Magn. Reson. Med. 34 (1995) 537–541.
[81] A. Kami, G. Meyer, P. Jezzard, M.M. Adams, R. Turner, L.G. Ungerleider, Functional MRI evidence for adult motor cortex plasticity during motor skill learning, Nature 377 (1995) 155–158.
[82] Y.-Z. Huang, M.J. Edwards, E. Rounis, K.P. Bhatia, J.C. Rothwell, Theta burst stimulation of the human motor cortex, Neuron 45 (2005) 201–206.
[83] J. Rhee, R.K. Mehta, Functional connectivity during handgrip motor fatigue in older adults is obesity and sex-specific, Front. Hum. Neurosci. 12 (2018).
[84] A.M. Chiarelli, E.L. Maclin, M. Fabiani, G. Gratton, A kurtosis-based wavelet algorithm for motion artifact correction of fNIRS data, NeuroImage 112 (2015) 128–137.
[85] X. Cui, S. Bray, D.M. Bryant, G.H. Glover, A.L. Reiss, A quantitative comparison of NIRS and fMRI across multiple cognitive tasks, Neuroimage 54 (2011) 2808–2821.
[86] J.M. Baker, J.L. Bruno, A. Gundran, S.H. Hosseini, A.L. Reiss, fNIRS measurement of cortical activation and functional connectivity during a visuospatial working memory task, PLoS ONE 13 (2018) e0201486.
[87] K.J. Verdière, R.N. Roy, F. Dehais, Detecting pilot's engagement using fNIRS connectivity features in an automated vs. manual landing scenario, Front. Hum. Neurosci. 12 (2018) 6.
[88] G. Vergotte, S. Perrey, M. Muthuraman, S. Janaqi, K. Torre, Concurrent changes of brain functional connectivity and motor variability when adapting to task constraints, Front. Physiol. 9 (2018) 909.
[89] M. Rubinov, O. Sporns, Complex network measures of brain connectivity: uses and interpretations, Neuroimage 52 (2010) 1059–1069.
[90] T. Nguyen, O. Babawale, T. Kim, H.J. Jo, H. Liu, J.G. Kim, Exploring brain functional connectivity in rest and sleep states: a fNIRS study, Sci. Rep. 8 (2018) 1–10.
[91] N.K.K. Dhand, M.S. Statulator: An online statistical calculator. Sample Size Calculator for Comparing Two Paired Means, 2014.
[92] A. Nemani, M.A. Yücel, U. Kruger, D.W. Gee, C. Cooper, S.D. Schwaitzberg, S. De, X. Intes, Assessing bimanual motor skills with optical neuroimaging, Sci. Adv. 4 (2018) eaat3807.
[93] J. Hennrich, C. Herff, D. Heger, T. Schultz, Investigating deep learning for fNIRS based BCI, 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2015, pp. 2844–2847.
[94] H. Jebelli, S. Hwang, S. Lee, EEG-based workers' stress recognition at construction sites, Autom. Constr. 93 (2018) 315–324.
[95] M.P. Matud, Gender differences in stress and coping styles, Person. Individ. Differ. 37 (2004) 1401–1415.
[96] A. Bonneville-Roussy, P. Evans, J. Verner-Filion, R.J. Vallerand, T. Bouffard, Motivation and coping with the stress of assessment: gender differences in outcomes for university students, Contemp. Educ. Psychol. 48 (2017) 28–42.
[97] S.K. Hunter, The relevance of sex differences in performance fatigability, Med. Sci. Sports Exerc. 48 (2016) 2247.
[98] J. Wang, M. Korczykowski, H. Rao, Y. Fan, J. Pluta, R.C. Gur, B.S. McEwen, J.A. Detre, Gender difference in neural response to psychological stress, Soc. Cogn. Affect. Neurosci. 2 (2007) 227–239.
[99] E.V. Goldfarb, D. Seo, R. Sinha, Sex differences in neural stress responses and correlation with subjective stress and stress regulation, Neurobiol. Stress 11 (2019) 100177.