Psychophysiological Monitoring for Spacewalks

Grace Wusk[1], Andrew Abercromby[2], Hampton Gabler[1]

[1]Department of Biomedical Engineering and Mechanics, Virginia Tech

[2]Human Physiology, Performance, Protection, and Operations Laboratory, NASA Johnson Space Center

Background

A spacewalk, or extravehicular activity (EVA), is one of the most mission critical and physically and cognitively challenging tasks that crewmembers complete. As Scott Kelly stated, reflecting on EVA during his yearlong mission onboard the International Space Station (ISS), "The focus required to do even simple work in space is daunting" (Kelly, 2017). On the ISS, Earth mission control provides significant, real-time decision support to crewmembers during EVA. With next-generation missions to the Moon and Mars, exploration EVA will challenge crewmembers in partial gravity environments with increased frequency, duration, and autonomy of operations. Given the distance from Earth, associated communication delays, and durations of these exploration missions, there is a monumental shift in responsibility and authority taking place in spaceflight; one that is moving from Earth-dependent to crew self-reliant (Caldwell, 2000; Feigh & Pritchett, 2014; Miller & Feigh, 2019).

For the safety, efficacy, and efficiency of future exploration EVAs, there is a need to better understand crew health and performance, particularly for surface operations. With this knowledge, technology and operations can be designed to better support future crew autonomy. As opposed to measuring cognitive workload using surveys or performance metrics, psychophysiological monitoring offers the benefit of objective, passive, and continuous quantification of crew state. The objective of this work was to develop a predictive model to classify cognitive workload using psychophysiological sensing during an operationally relevant EVA task. This work tested the limits of extending cognitive state modeling to a novel virtual reality (VR) task, using commercial wearable devices.

Methods

A sensor suite of commercial wearable devices, including the InteraXon Muse, the Empatica E4, and the Zephyr BioHarness, was selected to stream and record physiological signals in a human research study at NASA Johnson Space Center. Electroencephalography (EEG) data from the Muse was preprocessed using EEGLAB with Artifact Subspace Reconstruction (ASR) and Independent Component Analysis (ICA)(Delorme & Makeig, 2004; Laiti, Wusk, & Gabler, 2021). Custom MATLAB scripts extracted 162 time-synchronized psychophysiological features including EEG bandpowers, heart rate, heart rate variability, breathing rate, breathing variability, tonic and phasic skin conductance, skin temperature, and estimated blood pressure from pulse arrival and pulse transit time. Supervised machine learning with the K-Nearest Neighbor (KNN) algorithm was used to recognize patterns in psychophysiological features to predict crew state. Participant-independent models, trained on all participants, were compared to participant-specific models. The datasets were also subset by device to assess the predictive value of each of the wearables. The training data underwent ten-fold cross-validation repeated five times, and the k hyperparameter was tuned with a grid search from two to ten. All features were scaled and centered in the modeling process, and zero variance and near zero variance features were removed. Feature importance was assessed using distance measures from the KNN models.

Teaching a machine learning model to recognize crew state is not trivial as it is difficult to define "ground-truth" cognitive workload, especially in operational settings. The novel VR Translation Task was developed to control and quantify cognitive demands during an immersive, ambulatory EVA scenario (Wusk, Laiti, Gabler, & Abercromby, 2021). Participants walked on a passive treadmill while wearing a VR headset to move along a virtual lunar surface with constraints on time and simulated resources (Figure 1). During the Translation Task, a heads-up-display projected through the VR headset allowed the participants to monitor systems, such as red and green indicator lights, and manage simulated oxygen capacity. Additionally, they were responsible for identifying and recalling waypoints, flags with specific color-patterns, in the scene. Two configurations of the Translation Task were designed to simulate high and low cognitive workload conditions by varying the frequency and complexity of the system monitoring, resource management, and communication subtasks. Prior to applying any psychophysiological monitoring to the task, the cognitive workload was quantitively assessed using embedded performance metrics, including reaction times, and subjective surveys, including the NASA Task Load Index (TLX). Eight participants completed the high and low cognitive workload configurations of the Translation Task and the benchmark, Multi-Attribute Task Battery (MATB), while wearing the sensor suite. The MATB is a desktop cognitive battery developed by NASA Langley Research Center and has been used in previous studies to train predictive cognitive state models (Comstock & Arnegard, 1992; Harrivel et al., 2016, 2017; Wilson & Russell, 2003).

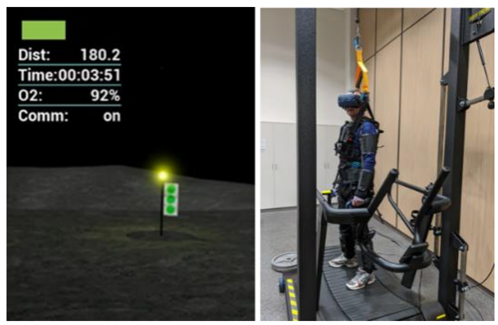

Figure 1. VR Translation Task.

Results

The results of the Translation Task confirmed the high cognitive workload corresponded to inferior performance and higher subjective survey ratings. During the high workload simulations, participants substantially overused their simulated oxygen resources by walking too fast, identified and recalled more waypoints incorrectly with slower responses, and had more variable reaction times to green indicator lights (1.5 ± 1.1 s in low workload, 3.1 ± 1.9 s in high workload). Participant-specific (boxplots in Figure 2) and participant-independent (triangles in Figure 2) models were able to distinguish between the high and low workload configurations of the Translation Task with 99-100% accuracy when features from all of the wearable devices were included. However, there are limitations on the generalizability of these results. The classification accuracies were much lower when trained on the MATB data and applied to Translation Task data. There was limited improvement from additional normalization and dimension reduction. Trained to recognize high and low workload during the seated MATB tasks, the best model was able to predict the workload configuration of the Translation Task with only 66% accuracy.

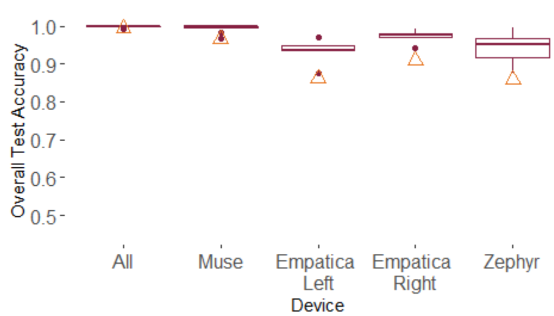

Figure 2. Classification accuracies for models by wearable device when trained within-task.

Discussion

The contributions of this work span the simulation, characterization, and modeling of cognitive state in more realistic settings and scenarios. This work focused on developing the Translation Task, including procedures, performance metrics, and machine learning models. It is part of a larger effort at NASA to assess physiology and performance during an entire VR EVA scenario. The high classification accuracies within the task illustrate the potential predictive power for future laboratory and field studies of cognitive workload using wearable devices. With the machine learning models, which were trained to recognize similarities in the psychophysiological feature sets, there is limited interpretability, and it is difficult to say if the models were detecting differences due to cognitive workload or other factors (i.e. physical workload, sensor signal quality, etc.). Ideally, a cognitive state model could be trained on a seated, benchmark task and tested on an operational task. However, additional work is required to generalize results across tasks by improving signal reliability and model robustness. Ultimately, this work paves the way for future real-time implementation to close the loop between human and automation. The predictions could be used to determine the psychological state of a crewmember performing an EVA. With knowledge of crew cognitive state, systems could appropriately support the crew by automatically providing feedback and assistance.

Acknowledgements
This work was supported by the NASA Space Technology Research Fellowship (Grant 80NSSC17K0171).

References

Caldwell, B. S. (2000). Information and communication technology needs for distributed communication and coordination during expedition-class spaceflight. In *Aviation Space and Environmental Medicine* (Vol. 71, pp. A6-10).

Comstock, J. R. J., & Arnegard, R. J. (1992). The multi-attribute task battery for human operator workload and strategic behavior research.

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*, 9–21.

Feigh, K. M., & Pritchett, A. R. (2014). Requirements for Effective Function Allocation. *Journal of Cognitive Engineering and Decision Making*, *8*(1), 23–32. https://doi.org/10.1177/1555343413490945

Harrivel, A. R., Liles, C., Stephens, C. L., Ellis, K. K., Prinzel, L. J., & Pope, A. T. (2016). Psychophysiological Sensing and State Classification for Attention Management in Commercial Aviation. In *AIAA Infotech @ Aerospace* (pp. 1–8). https://doi.org/10.2514/6.2016-1490

Harrivel, A. R., Stephens, C. L., Milletich, R. J., Heinich, C. M., Last, M. C., Napoli, N. J., … Pope, A. T. (2017). Prediction of Cognitive States during Flight Simulation using Multimodal Psychophysiological Sensing. In *AIAA Information Systems-AIAA Infotech @ Aerospace* (pp. 1–10). https://doi.org/10.2514/6.2017-1135

Kelly, S. (2017). *Endurance: A Year in Space, A Lifetime of Discovery*. New York: Alfred A. Knopf.

Laiti, J., Wusk, G., & Gabler, H. C. (2021). Using a Wearable, Low-Cost Brain Monitoring System during an Ambulatory Virtual Reality Extravehicular Activity (EVA) Simulation. In *Human Research Program Investigators' Workshop*.

Miller, M. J., & Feigh, K. M. (2019). Addressing the envisioned world problem: A case study in human spaceflight operations. *Design Science*, *5*(January). https://doi.org/10.1017/dsj.2019.2

Wilson, G. F., & Russell, C. A. (2003). Real-Time Assessment of Mental Workload Using Psychophysiological Measures and Artificial Neural Networks. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*(4), 635–644. https://doi.org/10.1518/hfes.45.4.635.27088

Wusk, G., Laiti, J., Gabler, H. C., & Abercromby, A. (2021). Psychophysiological Monitoring during an Ambulatory Virtual Reality Extravehicular Activity (EVA) Simulation. In *Human Research Program Investigators' Workshop*.