



USAARL-TECH-TR--2025-13

UNITED STATES ARMY AEROMEDICAL RESEARCH LABORATORY

Relationships Between Individual Differences, Physiological Measures, Cognitive Workload, and Task Performance: Implications for Operator State Monitoring

Amanda M. Kelley, Xiaomin Yue, Chris Aura, Kathryn A. Feltman, Ryan Mackie, & Johnell Brooks

Notice

Qualified Requesters

Qualified requesters may obtain copies from the Defense Technical Information Center (DTIC), Fort Belvoir, Virginia 22060. Orders will be expedited if placed through the librarian or other person designated to request documents from DTIC.

Change of Address

Organizations receiving reports from the U.S. Army Aeromedical Research Laboratory on automatic mailing lists should confirm correct address when corresponding about laboratory reports.

Disposition

Destroy this document when it is no longer needed. Do not return it to the originator.

Disclaimer

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citation of trade names in this report does not constitute an official Department of the Army endorsement or approval of the use of such commercial items.

IRB Determination and Number

This study, USAARL 2023-011, was determined to be research not involving human subjects by the USAARL exemption determination official and approved on 14 March 2024.

REPORT DOCUMENTATION PAGE				<i>Form Approved</i> OMB No. 0704-0188	
<small>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</small> PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 13-12-2024		2. REPORT TYPE Technical Report		3. DATES COVERED (From - To) January 2024 - October 2024	
4. TITLE AND SUBTITLE Relationships Between Individual Differences, Physiological Measures, Cognitive Workload, and Task Performance: Implications for Operator State Monitoring				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER 6.2	
6. AUTHOR(S) Kelley, A. ¹ , Yue, X. ^{1,2} , Aura, C. ¹ , Feltman, K. ¹ , Mackie, R. ^{1,3} , & Brooks, J. ⁴				5d. PROJECT NUMBER MO220103	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Aeromedical Research Laboratory P.O. Box 620577 Fort Novosel, AL 36362				8. PERFORMING ORGANIZATION REPORT NUMBER USAARL-TECH-TR--2025-13	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Development Command Military Operational Medicine Research Program 504 Scott Street Fort Detrick, MD 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S) USAMRDC MOMRP	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A. Approved for public release: distribution is unlimited.					
13. SUPPLEMENTARY NOTES ¹ U.S. Army Aeromedical Research Laboratory; ² Goldbelt Frontier, LLC; ³ Oak Ridge Institute for Science and Education; ⁴ Clemson University					
14. ABSTRACT The potential to monitor operator states and predict future performance deficits in real-time using physiological metrics continues to expand. Technological advancements continue to drive operator state monitoring (OSM) forward in terms of wearable sensors, data synchronization, and feasibility of use in applied settings. Likewise, advancements in understanding the relationships between physiological measurements, performance, and cognitive states are key to the development of a predictive model with acceptable levels of accuracy and classification. Individual differences, both within- and between-operators, introduce error into the model and need to be accounted for during computation in order to boost model performance. The main objective of this study was to identify and control for both stable (e.g., demographics) and dynamic (e.g., baseline physiology) sources of variance. Results indicated support for electroencephalogram derived measures of cognitive workload, demographic information (education level), and differences in perceived workload as predictors of performance deficits. The resultant sample size limited the analyses and thus further research is warranted.					
15. SUBJECT TERMS operator state monitoring, cognitive workload, performance, electroencephalogram					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT SAR	18. NUMBER OF PAGES 25	19a. NAME OF RESPONSIBLE PERSON Loraine St. Onge, PhD
a. REPORT UNCLAS	b. ABSTRACT UNCLAS	c. THIS PAGE UNCLAS			19b. TELEPHONE NUMBER (Include area code) 334-255-6906

This page is intentionally blank.

Acknowledgements

Great appreciation to the data collection team from Clemson University, support from the South Carolina National Guard, and editorial support from U.S. Army Aeromedical Research Laboratory staff. The data collection team from Clemson was supported by Clemson University's Virtual Prototyping of Autonomy Enabled Ground Systems (VIPR-GS), under Cooperative Agreement W56HZV-21-2-0001 with the U.S. Army DEVCOM Ground Vehicle Systems Center (GVSC).

This research was supported in part by an appointment to the Research Participation Program at the U.S. Army Aeromedical Research Laboratory administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. Army Medical Research and Development Command.

This page is intentionally blank.

Table of Contents

	Page
Acknowledgements	iii
Introduction	1
Methods.....	3
Participants.....	3
Materials	4
Questionnaires.....	4
Physiological measures.....	5
Driving simulator tasks.....	6
Procedure	10
Statistical Approach	10
Results.....	11
Effects of Workload Level.....	11
EEG.....	11
ECG.....	13
KSS and TLX.....	14
Task performance.....	15
Individual Differences and Dimensionality Reduction.....	16
Correlation analyses.....	16
Principal components analysis.....	17
Task Performance and Cognitive Workload Models.....	18
Discussion	20
Conclusion	21
References.....	22

List of Figures

1. DriveSafety CDS-200 simulator.....	7
2. Example of symbols used in stoplight and steering task	8
3. “E” stimuli positions	8
4. Slider task example screen.....	9
5. The normalized group PSD results from the frontal regions	12
6. Mean workload index values by workload level for each task.....	13
7. Mean HRV by workload level for both tasks	14
8. Mean KSS and TLX scores for both tasks.....	15
9. Mean performance level (percent accuracy) on slider task	16

List of Tables

1. Summary of Physiological Measures and Their Response to High Workload.....	1
2. Examples of Individual Differences That may Confound OSM	2
3. Parameter Specifications for Each Stoplight and Steering Task Workload Level	8
4. Parameter Specifications for Each Slider Task Workload Level.....	10

Table of Contents (continued)

List of Tables (continued)

	Page
5. Individual Differences and Dependent Measures Used in Analyses	11
6. Correlational Analyses for Individual Differences Variables and Slider Task Performance	16
7. Correlational Analyses Between Individual Differences Variables	17
8. Principal Components Analysis: Total Variance Explained.....	17
9. Principal Components Analysis: Varimax Rotated Factor Loadings	18
10. Performance Model Parameter Estimates	19
11. Correlations Between Education Level and TLX Subscale Scores Averaged Across Workload Levels on the Slider task. Pearson's r is Reported in Each Cell	20

Introduction

The potential to monitor operator states and predict future performance deficits in real-time using physiological metrics continues to expand. Technological advancements continue to drive operator state monitoring (OSM) forward in terms of wearable sensors, data synchronization, and feasibility of use in applied settings. Likewise, advancements in understanding the relationships between physiological measurements, performance, and cognitive states are key to the development of a predictive model with acceptable levels of accuracy and classification. Individual differences, both within- and between-operators, introduce error into the model and need to be accounted for during computation in order to boost model performance. The main objective of this study was to identify and control for both stable (e.g., demographics) and dynamic (e.g., baseline physiology) sources of variance.

Research has long established the brain and body connection, supporting the use of objective physiological measures as proxies for latent constructs. Much of this work has been focused on cognitive workload (albeit a variety of definitions) demonstrating generally consistent trends in physiological response concurrent to changes in workload (Borghini et al., 2013; Charles & Nixon, 2019) (Table 1).

Table 1. Summary of Physiological Measures and Their Response to High Workload

Physiological Measure	Response During High Workload
Heart rate	Increases (Unni et al., 2017; Hidalgo-Munoz et al., 2019)
Heart rate variability (HRV)	Decreases (Stuiver et al., 2014)
Pupil diameter	Increases (Feng et al., 2018)
Eye fixation duration	Increases (Feng et al., 2018), Decreases (Schulz et al., 2011)
Theta wave activity – electroencephalography (EEG)	Increases (Wu et al., 2017)
Alpha wave activity – EEG	Decreases (Van Orden et al., 2001; Wu et al., 2017)
Beta wave activity – EEG	Increases (Kurimori & Kakizaki, 1995), Decreases (Xiaoli et al., 2020; Hussain et al., 2021)
Oxygen saturation (rSO ₂)	Increases (Sassoroli et al., 2008)
Electrodermal activity	Increases (Tarabay et al., 2018)
Cerebral blood flow velocity (CBFV)	Increases (Warm & Parasuraman, 2007)

Individual differences in baseline physiology, as well as state and trait characteristics, may confound the utility of models predicting performance using physiological indices of cognitive state. These confounding variables limit accuracy and interpretability of models to identify when performance will decline across a range of individuals. The discrepancy in

performance between generic and individualized (incorporating baseline physiology) models was demonstrated with a stress response model utilizing heart rate variability and electrodermal activity. The addition of baseline or individual physiological variables improved the models' accuracies drastically (Nkurikiyeyezu et al., 2019). In order to improve the predictive validity, as well as reliability, of models aimed at diagnosing operator state and predicting performance deficits, it will be critical to account for individual differences.

A wide range of individual factors can impact cognitive and physiological responses to stressors (such as changes in task demands). For the purpose of this study, focus is given to individual differences known to correlate with or impact physiological responses and may further complicate interpretation of changes in responses concurrent to changes in task demands (cognitive workload) (Table 2). These are factors that ultimately will need to be accounted for in OSM algorithms and systems. Broadly, these factors can be categorized as stable, relatively consistent or unchanging, or dynamic, routinely changing or varying. A comprehensive list of candidate mediating or moderating variables is unnecessary given the target end-user population (military aviators). Strict aeromedical standards and routine exams preclude the possibility of many medications or conditions in rated aviators. Of the applicable individual difference variable, anxiety is a good example of a construct that can be static or dynamic in nature. Static levels of anxiety are less likely in military aviators as anxiety disorders are disqualifying conditions. However, military aviators, like all humans, are susceptible to periods of acute anxiety, such as those related to one's home life or work demands. Symptoms resulting from an acute anxiety period can have implications for not only how one responds to a stressful situation, but also in how they respond physiologically. For example, Wheelock et al. (2016) found support for individual differences in self-reported stress and acute anxiety related to activity within the prefrontal cortex. Identifying which of these factors are important to capture and account for in an OSM system is the primary purpose of this paper.

Table 2. Examples of Individual Differences That may Confound OSM

Construct or Condition	Potential Impact
Treatment for hypertension	Some treatments can lower heart rate
Anxiety and depression symptoms	Acute experiences can elevate heart rate or impact sleep resulting in daytime sleepiness
Fatigue/daytime sleepiness	Fatigue impacts multiple physiological responses
Experience	Less experience may increase psychological stress and physiological responses
Age	Healthy aging may impact physiological responses including increased EEG signal complexity

Individualized approaches to workload monitoring, that account for within- and between-individual variability, are gaining traction in the literature. Multiple methods and approaches can be used to individualize models. For instance, Teo et al. (2020) state that individuals show differences in physiological response sensitivity, with some showing greater differences in one physiological measure over another in response to changes in task demand. Thus, the model can be specified to apply weights by sensitivity of any single index. Another method is for the model to capture measures that show large changes as defined by a criterion threshold (Teo et al., 2020). Alternatively, Ma et al. (2024) included experience level as a moderating factor between task demand level and physiological responses to changes in task demands. While these studies all demonstrate positive findings regarding the relative approaches taken, they were all conducted in different contexts, none of which included military, rotary-wing aviators. One of these approaches may translate to this highly specialized subpopulation; however, it is likely that a model of acceptable reliability and validity will need to be customized to the cockpit, mission demands, and tailored to user-profiles. Addressing individual variability within an Army aviation population will require the use of Army aviators as the sample to develop the model, given that aviators tend to fall within narrow ranges on the spectrums of individual difference measures already (e.g., intelligence, personality) (Causse et al., 2011). Incorporation of stable and dynamic sources of variance could enable the development of robust algorithms that are able to predict and/or identify operator state changes in real-time.

Studies of military aviators tend to have small sample sizes (e.g., 8-10 participants) and oversample pilots early and late in their careers. This poses a significant barrier to studying individual differences and identifying which variables need to be factored into any resultant models. Thus, this study did not limit to aviators, but rather collected a large sample of data from National Guard Soldiers. Given that the participants were not aviators, a simulated flight was not appropriate, thus we evaluated performance on a set of tasks adapted for a driving simulator. The primary objective was to model changes in performance using individual differences and physiological variables as predictors.

Methods

This study was reviewed and approved prior to execution by the U.S. Army Aeromedical Research Laboratory (USAARL) Regulatory Compliance Officer as research not involving human subjects. The data analyzed were collected by Clemson University and provided to USAARL deidentified under a cooperative research and development agreement. The data analyzed are a subset of a larger study.

Participants

Participants were 53 South Carolina National Guard Soldiers (9 females, 38 males, 6 missing data). Mean age was 33.10 years ($SD = 13.51$, 8 missing data). Education levels included were high school or high school equivalent ($n = 12$), associate degrees ($n = 1$), some college ($n = 14$), bachelor's degrees ($n = 12$), some graduate level training ($n = 4$), and graduate degrees ($n = 5$) (5 missing data). Majority of participants were White ($n = 37$) with the remaining participants (6 missing data) reporting as African American ($n = 7$), Hispanic ($n = 2$), and Asian ($n = 1$).

Materials

Instruments and tasks used in this study are divided in three categories, questionnaires, physiological measures, and driving simulator tasks.

Questionnaires.

All instruments were administered electronically.

Demographics.

Participants reported age, gender, ethnicity, education level, and, if applicable, flight hours.

Adult Attention Deficit/Hyperactive Disorder (ADHD) Self Report Scale Symptom Checklist (ASRS).

The ASRS contains 18 items and requires 2 minutes for completion. It was developed in conjunction with the World Health Organization (WHO) and the Workgroup on Adult ADHD (Kessler et al., 2005) and is used as a screening tool with adult patients. The items are consistent with the Diagnostic and Statistical Manual of Mental Disorders, version IV criteria (American Psychiatric Association, 2000). Two scores are outputted, hyperactivity and inattentiveness.

Beck Depression Inventory (BDI-II).

Depression symptoms were measured using the Beck Depression Inventory-II (BDI-II; Beck et al., 1996). The BDI-II is a commonly used 21-item, multiple-choice self-report tool that captures affect, cognition, and physical symptoms of depression over the most recent two-week period. Higher scores indicate greater endorsement of depression symptoms.

State-Trait Anxiety Inventory (STAI).

Anxiety symptoms were measured using the STAI (Spielberger et al., 1983), which is a 40-item, self-report anxiety inventory rated on a 4-point Likert-type scale that captures two types of anxiety, state, event-dependent anxiety, and trait, persistent demonstrations of anxiety as a personal characteristic. The primary outcome measures are state score and trait score, where higher scores reflect greater endorsement of anxiety symptoms.

NASA Task Load Index (TLX).

Subjective workload was captured using the TLX (Hart & Staveland, 1988). Following each task and difficulty level, participants rate their experience, using a 100-point scale, on the following dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. The outcome measures are ratings on each subscale as well as a total score.

Karolinska Sleepiness Scale (KSS).

The KSS is a well-validated single item questionnaire that asks subjects to rate how sleepy they feel in the moment (Kaida et al., 2006). The KSS measures daytime sleepiness with higher scores indicating greater daytime sleepiness. Specifically, responses range from 1 “extremely awake” to 9 “extremely sleepy – fighting sleep.” It was administered after each level of workload to assess participants’ subjective sleepiness.

Physiological measures.

Both EEG and electrocardiogram (ECG) data were collected using the B-Alert X-24 wireless wet electrode system with 20 channels corresponding to scalp locations according to the International 10-20 system (frontal channels: Fp1, Fp2, F7, F3, Fz, F4, F8; central channels: C3, Cz, C4, T3, T4; parietal and occipital channels: P3, POz, Pz, P4, T5, T6, O1, O2). Power spectral density (PSD) values were computed using the automated algorithms provided through the B-Alert Live software. Prior to computing PSD values, artifacts were identified and removed using the Advanced Brain Monitoring (ABM) algorithms for artifacts associated with electromyography (EMG), eye blinks, excursions, saturations, and spikes.

EEG data processing approach.

A customized data processing pipeline was developed to process the raw EEG data and compute PSD values. Initially, the raw EEG data were bandpassed between 1 to 40 Hertz (Hz) to eliminate the power line noise. During the prolonged recording sessions, some channels may detach or experience increased impedance, which can introduce significant noise and lead to inaccurate results that do not accurately reflect cognitive processes in the brain. To address this, the RANdom Sample Consensus (RANSAC) algorithm, a robust statistical method (Fischler & Bolles, 1981), was employed to detect and remove bad channels from further analysis. After the removal of these channels, the remaining data were re-referenced using an average reference.

Following the RANSAC step, artifact subspace reconstruction (ASR; Mullen et al., 2013) was used to remove artifacts such as blinks and muscle activity. ASR is an adaptive method for online or offline correction of artifacts, particularly effective for cleaning continuous, non-triggered EEG data (e.g., Bulea et al., 2015; Luu et al., 2017) and is recommended for wireless EEG systems (Mullen et al., 2015). In brief, the ASR algorithm identifies a clean EEG segment to compute its statistical properties, such as a covariance matrix. It then applies a sliding window over the EEG data, performing principal component analysis (PCA) on each window. Components with variances that exceed a set threshold (based on the clean segment) are identified as artifacts and removed, with the signal reconstructed using the mixing matrix from the clean segment.

The clean EEG data was subjected to continuous wavelet transformation (CWT) to convert the signals into time-frequency data. The resulting data was then averaged in both the frequency and time domains to generate four PSD values corresponding to each frequency band: delta, theta, alpha, and beta. The B-Alert software was used to collect baseline data to create the participant’s individualized workload index profile. The information collected from the baseline EEG data was then used to make comparisons with the participant’s test data to estimate

workload. The baseline acquisition took place for 10 minutes while participants silently read nonsense text presented on the center screen of the simulator.

ECG data processing approach.

The raw ECG data were preprocessed to remove high-frequency and powerline noise, and baseline drift, as well as to eliminate artifacts (Makowski et al., 2021). The quality of the ECG data was then evaluated using the comprehensive method proposed by Zhao and Zhang (2018), which involves two steps. First, multiple signal quality indices (SQIs) were calculated, each reflecting different aspects of ECG data quality. These SQIs were then mathematically combined into a single index, which were categorized as excellent, barely acceptable, or unacceptable. ECG data categorized as unacceptable were excluded from further analysis.

Subsequently, ECG features in the time domain were extracted (Makowski et al., 2021), including heart rate, mean heart rate variability, and the standard deviation of heart rate variability, for statistical analysis. Those steps were repeated independently for each task and each level of workload.

Since ECG was recorded using the EEG device, its sampling rate is relatively low compared to standalone ECG devices (256 Hz vs. 1000 Hz). Additionally, the EEG device is more susceptible to noise such as motion artifacts, which can significantly impact ECG recordings. As a result, the ECG data were anticipated to be noisy.

Driving simulator tasks.

Performance was measured on two tasks administered using the CDS-200 Clinical Driving Simulator (Figure 1) from DriveSafety Inc. The tasks used were developed by the research team at Clemson University in collaboration with DriveSafety Inc. The interactive tasks (stoplight and steering, and slider) have been previously demonstrated to effectively titrate task demand (e.g., Goodenough et al., 2012).

This space is intentionally blank.



Figure 1. DriveSafety CDS-200 simulator.

Stoplight and Steering® task.

In this task, participants respond to symbols presented on the center screen as well as the outer screens of the simulator. On the center screen, left and right arrows along with green and red traffic lights are presented. Participants respond by tapping on the appropriate pedal (corresponding to the traffic light color) and turning the steering wheel to the left or right (corresponding to the direction of the arrow) (Figure 2). The symbols on the outer screens are an “E” presented in one of four positions: forward, backward, upward, or downward (Figure 3). Prior to the onset of the task, participants are shown a reference “E.” They are told to remember the position of the reference “E.” Then during the task, target “E” stimuli are presented on the outer two screens; both the order of “E” positions and location on the screen are randomized. Participants respond using red buttons on the steering wheel, one on the right and one on the left. Participants press the red button on the side corresponding to the outer screen (right or left) on which the target “E” is presented, in the position that matches the reference “E” position. Reaction times and response accuracies to each symbol presented are recorded; however, accuracies (in percentage) are included in the analyses. If a participant misses a target “E” or responds to an incorrect “E,” that information is also recorded. Workload level is manipulated by the frequency of “E” presentations as well as the duration of time each “E” is presented (Table 3).

This space is intentionally blank.

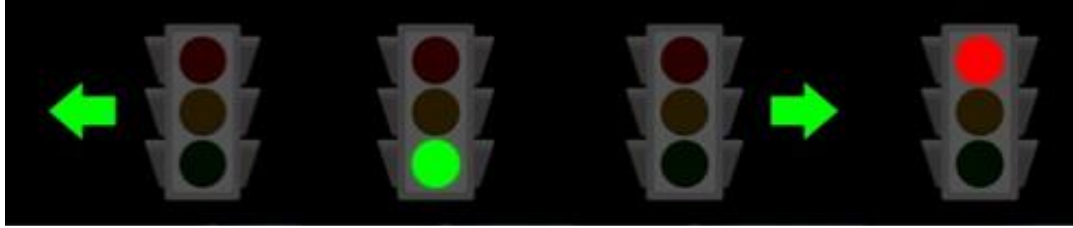


Figure 2. Example of symbols used in stoplight and steering task.

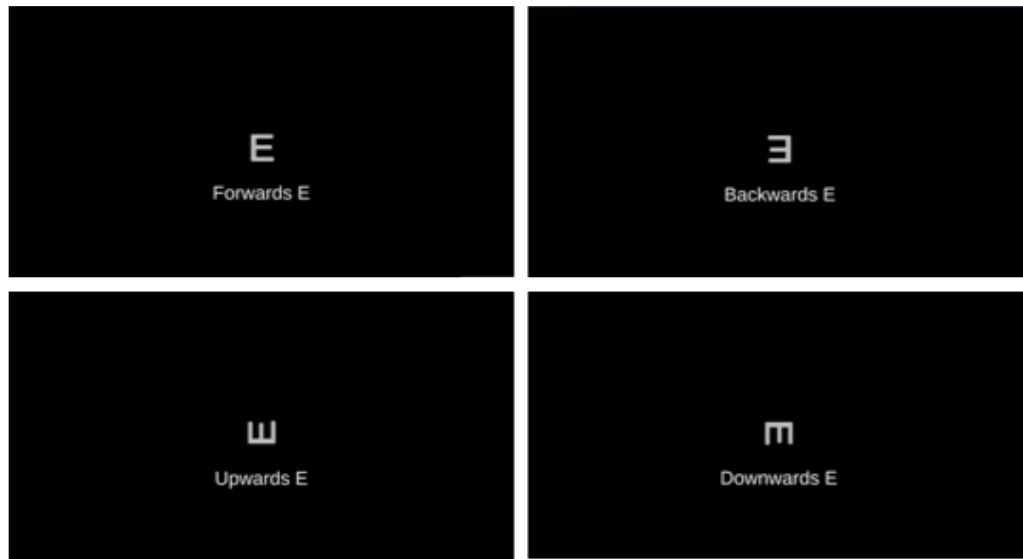


Figure 3. “E” stimuli positions.

Table 3. Parameter Specifications for Each Stoplight and Steering Task Workload Level

Workload Level	Target “E” Presentation Duration (seconds)	Target “E” Presentation Frequency (seconds)
Level 1	2.0	2.0 – 3.5
Level 2	1.5	1.0 – 2.0
Level 3	1.0	0.75 – 1.5

Participants control how long each symbol is presented on the screen; as soon as the input is released, the next symbol is presented. The input is released when the participant’s foot is lifted from the gas or brake pedal or when the participant returns the steering wheel to the home (12:00) position. Once that neutral point occurs, the next symbol is presented. If the participant holds down a pedal or does not return the steering wheel back to the home (12:00) position, then the next symbol will not be displayed. Given that participants control the length of time the symbols are displayed, the task is not defined by a set number of trials, but rather lasts for 10 minutes.

Slider[®] task.

The slider task uses the same symbols as the stoplight and steering task (Figures 2 and 3) which are presented at a set speed (unlike the stoplight and steering task, which allows participants to control the speed at which symbols are presented). Symbols are presented by “entering” the center monitor on the right and then “slide” across the screen to the left (Figure 4). When a symbol enters the white box, participants respond to it using the same inputs described in the stoplight and steering task. Once at least half of the symbol is in the response box, the simulator will “count” the participant’s input. Concurrently, target “E” stimuli are presented on the outer screens and response inputs are the same as in the stoplight and steering task. This task is completed after the stoplight and steering task and participants are familiar with the symbols and respective inputs.

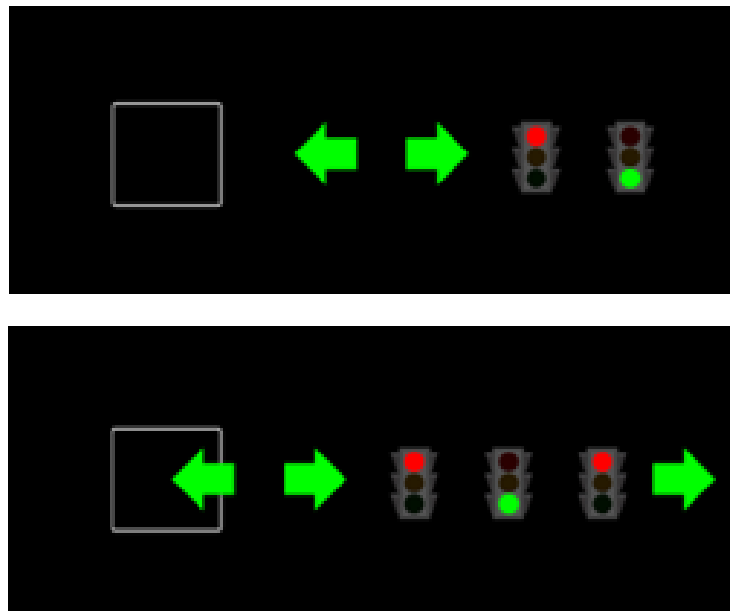


Figure 4. Slider task example screen.

Reaction time and response accuracies to the presented symbols (both the center screen symbols and target “E” stimuli) are recorded; however, accuracies (in percentages) are included in the analyses. Workload is manipulated by the speed at which the symbols slide from appearance at the left edge of the screen to the white response box (Table 4). The task is performed for 10 minutes.

This space is intentionally blank.

Table 4. Parameter Specifications for Each Slider Task Workload Level

Workload Level	Number of Symbols Presented	Speed of Symbol From Appearance to Response Box (seconds)	“E” Task Workload Level (Table 3)
Level 1	317 – 318	6.8	Level 1
Level 2	628 – 629	3.4	Level 2
Level 3	1054 – 1055	3.0	Level 2
Level 4	1493 – 1494	1.4	Level 3

Procedure

Participants provided written informed consent prior to study enrollment or conduct of any study procedures. First, participants completed a set of questionnaires (demographics, BDI-II, STAI, ASRS) on a tablet, then were instrumented with the B-Alert system. Next, participants were “fit” to the driving simulator (e.g., calibration of pedals and steering wheel, adjustment of monitor height) prior to completing the baseline EEG recording. Baseline EEG was recorded while participants, seated at the driving simulator, looked at the center monitor and read the same nonsense text for 10 minutes. Participants then completed the driving simulator tasks. Additionally, the TLX and KSS were completed at the end of each workload level within the tasks. Markers were manually inputted to mark the start and end of each task within the EEG recordings. This procedure is a subset of the entire study.

Statistical Approach

Physiological data were processed, as described above, prior to analyses. All data were inspected for impossible values and technical errors in recording. The analyses are divided into two sections: effects of workload level, and impact of individual differences on performance. First, to assess the effects of workload level, which also served as a manipulation check for the tasks, the EEG, ECG, KSS, TLX, and performance outcomes were analyzed using within-subjects analyses of variance (ANOVAs). Next, individual difference variables (Table 5) were evaluated for inclusion in a model predicting performance. Correlational matrices were first calculated to estimate relationships between the variables (Pearson’s r for relationships between continuous variables, Spearman’s ρ for relationships between binary and continuous variables). In order to mitigate the influence of multicollinearity, a principal components analysis (PCA) was then used to reduce the variables into factors. Finally, a generalized estimating equations (GEE) (Liang & Zeger, 1986) was conducted to evaluate the impact of individual differences on performance. GEEs, an extension of generalized linear models, account for the correlation between repeated measurements within subjects by specifying a working correlation structure. Importantly, GEEs do not assume a linear relationship between predictors and outcomes, nor do they require individual differences to be modeled solely as random effects. Instead, GEEs focus on estimating population-averaged effects while adjusting for within-subjects’ correlations, making this analysis approach well-suited to address the study objective.

Table 5. Individual Differences and Dependent Measures Used in Analyses

Construct	Survey/Instrument Used	Dependent Variable
Sleepiness ratings	KSS	Self-reported level of sleepiness
Reported workload levels	TLX	Subscale scores: Temporal demand Physical demand Performance Mental demand Frustration Effort
Depression symptoms	BDI-II	Total score
Anxiety symptoms	STAI	State score (event-dependent anxiety) Trait score (persistent demonstrations of anxiety)
Cognitive workload	Theta to beta ratio from EEG data	Workload index: Ratio of theta to beta frequency band PSD for frontal regions
Heart rate variability	ECG data	Heart rate variability
Age	Demographics	Age in years
Gender	Demographics	Male Female
Education level	Demographics	Years of education
ADHD symptoms	ASRS	Hyperactivity score Inattentiveness score
Task performance	Slider and stoplight and steering tasks	Accuracy measured in percentage

Results

Effects of Workload Level

EEG.

Of the 53 participants, 31 had complete sets of EEG data (baseline, three workload levels for the stoplight and steering task, and four workload levels for the slider task). Thus, these 31 participants were included in the analyses. The effects of workload level in each task were estimated for PSD across all frequency bands within the frontal areas (delta, theta, alpha, beta). Repeated measures ANOVAs evaluating the main effects of workload level (3 levels for stoplight and steering task and 4 levels for slider task) and frequency band (delta, theta, alpha, beta) as well as interaction effects were conducted. Results suggest increases in task workload elevated PSDs, for both tasks, (stoplight and steering task: $F(2, 60) = 11.25, p = 7.11 * 10^{-5}$; slider task: $F(3, 90) = 39.55, p = 2.15 * 10^{-16}$) (Figure 5). Moreover, a significant interaction was observed between task workload and frequency band (stoplight and steering task: $F(6, 180) = 9.97, p = 1.69 * 10^{-9}$; slider task: $F(9, 270) = 19.15, p = 1.15 * 10^{-24}$).

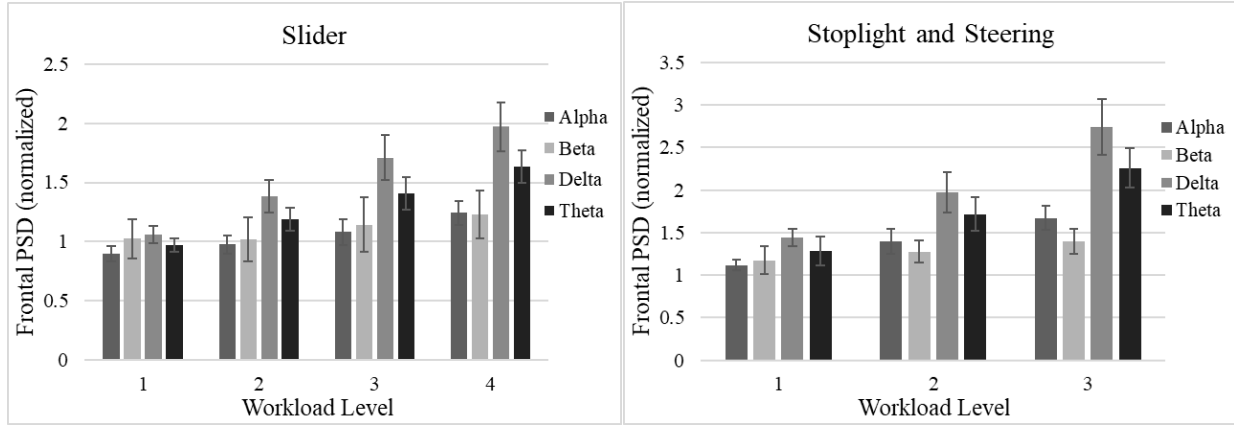


Figure 5. The normalized group PSD results from the frontal regions. Error bars represent standard error of the mean.

Additionally, a workload index was derived from the EEG data, specifically, the ratio of theta to alpha band PSD from the frontal regions, which has been shown to be a reliable indicator of cognitive workload (Raufi & Longo, 2022). A repeated measures ANOVA evaluated the effect of workload level on the workload index for each task. The analysis shows the increasing task workload significantly impacts participants' cognitive workload (Figure 6) (stoplight and steering task: $F(2,60) = 14.73, p = 6 * 10^{-5}$; slider task, $F(3, 90) = 30.84, p = 8.32 * 10^{-14}$). For the stoplight and steering task, the post-hoc analysis found significant differences in cognitive workload between workload level 1 vs. level 3 ($t(30) = -4.40, p = 3.78 * 10^{-3}$ (corrected)), and level 2 vs. level 3 ($t(30) = -3.5, p = 4.42 * 10^{-2}$ (corrected)), with insignificant difference between level 1 vs. level 2 ($t(30) = -2.58, p = 0.045$). These findings suggest that workload level 3 is substantially more cognitively demanding than levels 1 and 2.

For the slider task, the post-hoc analysis revealed significant differences in cognitive workload at a $p = 0.01$ level across all comparisons, except between workload levels 2 and 3 ($t(30) = -2.87, p = 0.044$ (corrected)) and between level 3 and level 4 ($t(30) = -3.33, p = 0.014$ (corrected)). These results suggest that while workload levels 2, 3, and 4 are all more cognitively demanding than level 1, cognitive workload increased gradually from level 2 to level 4.

This space is intentionally blank.

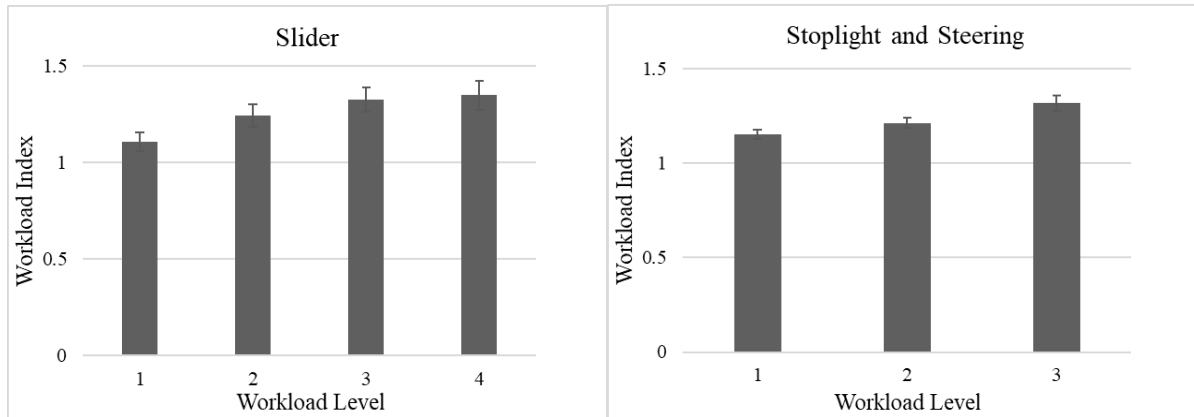


Figure 6. Mean workload index values by workload level for each task.

ECG.

Of the 53 participants, 27 completed all ECG recording sessions. Among these, 21 participants had data categorized as barely acceptable or above (see Methods). However, due to data quality issues, not all 21 participants had useable data for each workload level. For example, in the stoplight and steering task, usable data was obtained from 13 participants at workload level 1, 15 participants at workload level 2, and 12 participants at workload level 3. Given the imbalanced data across workload levels, statistical analysis was not performed, and descriptives are presented.

The outcome variable derived from the ECG data was HRV, which measures the variation in time between heartbeats and reflects the regularity of the beats. In general, a decrease in HRV indicates increased stress which matches the pattern of observed data (Figure 7). Specifically, HRV decreases as task workload level increases, with this pattern being more pronounced for the slider than the stoplight and steering task, suggesting that higher workload levels may elevate participants' cognitive workload. This observation is consistent with the EEG results presented above.

This space is intentionally blank.

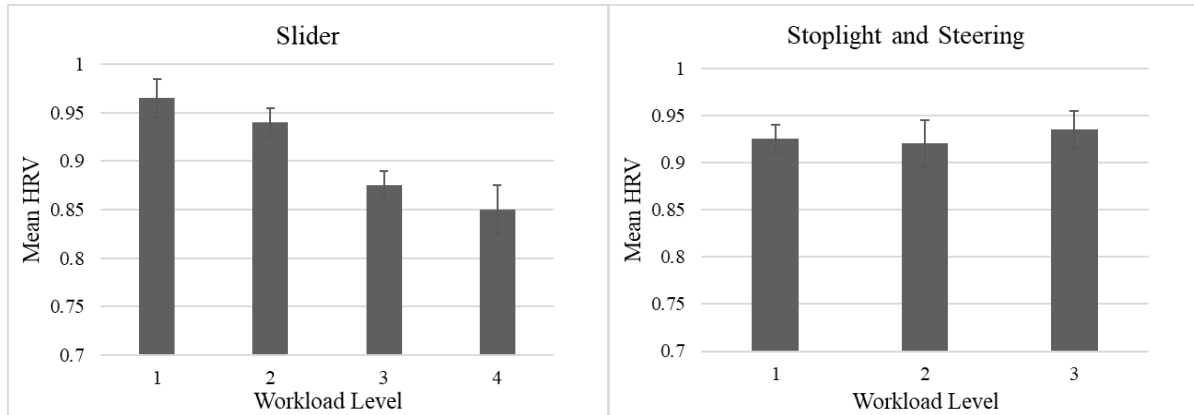


Figure 7. Mean HRV by workload level for both tasks. Error bars represent standard error of the mean.

KSS and TLX.

Twenty-two participants completed the KSS and TLX survey for both tasks. Workload level in the stoplight and steering task did not significantly impact the KSS scores ($F(2, 34) = 3.8, p = 0.055$), while workload level in the slider task had a marginally significant effect on the KSS scores ($F(3, 39) = 6.316, p = 0.019$). These findings suggest that participants were able to maintain their alertness across different levels of workload during both tasks.

In contrast, the workload level in the stoplight and steering task did not affect the total TLX score ($F(2, 34) = 9.174, p = 0.030$) whereas the workload level in the slider task significantly influenced the total TLX score ($F(3, 39) = 58.540, p = 3.44 \times 10^{-10}$). These results indicate that participants required more effort to perform the slider than the stoplight and steering task as the workload level increased.

This space is intentionally blank.

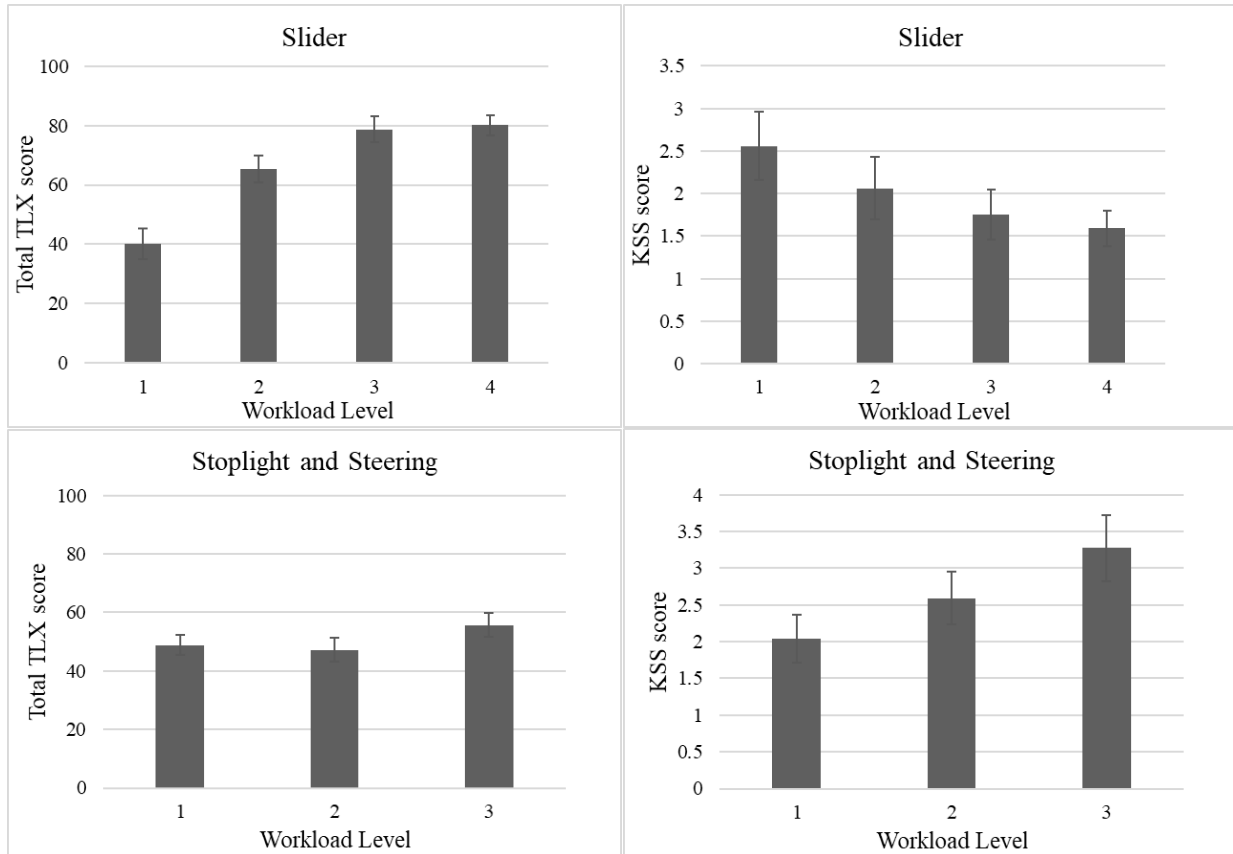


Figure 8. Mean KSS and TLX scores for both tasks. Error bars represent standard error of the mean.

Task performance.

Performance analysis was not completed on the stoplight and steering task given the lack of differences in perceived workload between difficulty levels observed with the TLX data. Forty-five participants completed the slider task and were included in the analysis. Performance (percentage of accurate responses to center screen stimuli) degraded with increasing workload levels, $F(3, 132) = 612.33, p = 3.10 \times 10^{-77}$.

This space is intentionally blank.

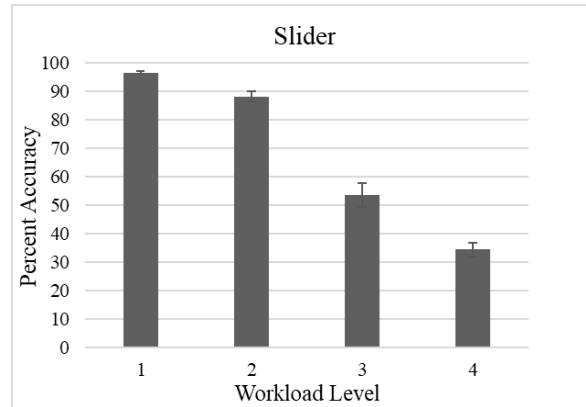


Figure 9. Mean performance level (percent accuracy) on slider task. Error bars represent standard error of the mean.

Individual Differences and Dimensionality Reduction

Correlation analyses.

Correlational relationships between individual differences variables and mean slider task performance (percent accuracy) were calculated (Table 6). The variables included were the variables measured only once as the repeated variables (KSS, TLX scores) were already evaluated in relation to the task workload levels. Analyses were also conducted between the individual differences variables and performance at each workload level to identify inconsistent patterns, none of which were noted. Correlation analyses were conducted to determine the relationships between predictor variables (individual differences), specifically, total BDI, inattentiveness and hyperactivity (from the ADHD measure), and state and trait anxiety (Table 7). With regard to demographic variables, age and education level were correlated ($r(45) = 0.65, p < 0.001$).

Table 6. Correlational Analyses for Individual Differences Variables and Slider Task Performance

Variable	Statistic	Correlation Value	<i>p</i>
Hyperactivity	<i>r</i>	-0.25	0.34
Inattentiveness	<i>r</i>	-0.09	0.74
Beck Depression Inventory	<i>r</i>	0.03	0.46
State-anxiety	<i>r</i>	0.04	0.87
Trait-anxiety	<i>r</i>	0.03	0.89
Gender	<i>rho</i>	-0.10	0.70
Education level	<i>r</i>	-0.54*	0.02*
Age	<i>r</i>	-0.40	0.11

*Indicates significant result.

Table 7. Correlational Analyses Between Individual Differences Variables

		Hyperactivity	Inattentiveness	Beck Depression Inventory	State- Anxiety	Trait- Anxiety
Hyperactivity	<i>r</i>	1	.758**	.305*	.451**	.465**
Inattentiveness	<i>r</i>	.758**	1	.376**	.525**	.541**
Beck Depression Inventory	<i>r</i>	.305*	.376**	1	.711**	.780**
State-anxiety	<i>r</i>	.451**	.525**	.711**	1	.861**
Trait-anxiety	<i>r</i>	.465**	.541**	.780**	.861**	1

**Correlation is significant at the 0.01 level (2-tailed).

*Correlation is significant at the 0.05 level (2-tailed).

Principal components analysis.

PCA and a Varimax rotation were performed on total BDI, inattentiveness and hyperactivity (both derived from the ASRS), and state- and trait-anxiety scores (both derived from the STAI). Suitability of the data for factor analysis was assessed and supported using the Kaiser-Meyer-Olkin measure of sampling adequacy, with a value of 0.76, and the Bartlett's Test of Sphericity ($\chi^2(10) = 206.93, p < 0.001$). Two PCA components, collectively explaining 87.23% of the total variance, were extracted with eigenvalues that exceed 1 (Table 8). The first component, named "depression and anxiety symptoms," included the BDI total score, and state- and trait-anxiety scores. The second component, named "ADHD symptoms," included the inattentiveness and hyperactivity scores (Table 9).

Table 8. Principal Components Analysis: Total Variance Explained

Component	Eigenvalue	Variance (%)	Cumulative Variance (%)
1	3.31	66.23	66.23
2	1.05	20.99	87.22
3	0.28	5.53	92.75
4	0.25	4.94	97.69
5	0.12	2.31	100.00

This space is intentionally blank.

Table 9. Principal Components Analysis: Varimax Rotated Factor Loadings

Variable	Component	
	1	2
Hyperactivity score	0.18	0.92
Inattentiveness score	0.28	0.89
Total BDI score	0.91	0.10
State-anxiety score	0.87	0.33
Trait-anxiety score	0.90	0.32

Task Performance and Cognitive Workload Models

A GEE was used to evaluate factors predicting performance on the slider task. Performance was defined as percent accuracy. The variables included in the model as predictors were the two PCA components (“depression and anxiety symptoms” and “ADHD symptoms”), KSS score, TLX total score, EEG workload index, workload level, and education level. Data from 17 participants with complete datasets were included. Given the small number of full datasets for inclusion, the variables of age, gender, and ethnicity were excluded from the analysis as they did not correlate with task performance. Multicollinearity was assessed and between predictors correlations exceeded $r = 0.5$. The results yielded, unsurprisingly, an effect of workload level (previously demonstrated), reported workload perception, as well as the EEG workload index (Table 10). One individual difference variable, education level, was significantly negatively associated with performance.

This space is intentionally blank.

Table 10. Performance Model Parameter Estimates

Predictor	Coefficient	Standard Error	Wald χ^2	<i>p</i>
(Intercept)	66.99	8.86	57.11	< 0.0001
Depression and anxiety symptoms	-1.72	2.11	0.67	0.41
ADHD symptoms	0.29	1.20	0.06	0.81
EEG workload index	-9.06	4.12	4.84	0.03
TLX score	-0.16	0.07	5.04	0.02
Workload level 1	54.78	4.19	171.12	< 0.0001
Workload level 2	51.05	2.50	417.39	< 0.0001
Workload level 3	19.24	3.06	39.54	< 0.0001
KSS	-1.06	0.82	1.67	0.19
Education level	-2.15	0.94	5.28	0.02

Note. Workload level 4 served as a reference category.

In order to further explore the education level variable and its relationship to performance, correlational analyses were conducted, specifically with mean values of the TLX subscale scores. The results showed a significant, positive relationship between frustration and education level, such that as education level increases, so do reported frustration levels. Frustration did not correlate with the other TLX subscale components.

This space is intentionally blank.

Table 11. Correlations Between Education Level and TLX Subscale Scores Averaged Across Workload Levels on the Slider task. Pearson's r is Reported in Each Cell

	Education Level	Effort	Frustration	Mental Workload	Performance	Physical Workload	Temporal Demands
Education level	1	0.41	0.62*	0.13	0.21	0.36	0.23
Effort	-	1	0.16	0.82*	0.48	0.76*	0.65*
Frustration	-	-	1	0.14	0.11	0.30	0.29
Mental workload	-	-	-	1	0.23	0.82*	0.89*
Performance	-	-	-	-	1	0.16	0.05
Physical workload	-	-	-	-	-	1	0.75*
Temporal demands	-	-	-	-	-	-	1

*Indicates significance at $p < 0.01$.

Discussion

OSM is one of the first steps toward adaptive automation. OSM will incorporate physiological data to predict performance and, subsequently, risks to mission success. A number of factors confound the relationships between cognitive states, physiological responses, and performance levels. In order to develop a model that predicts performance deficits with an acceptable degree of accuracy, consideration must be given to internal and external factors that influence physiological outcomes as well as interpretation of changes. This study focused on individual differences previously shown to correlate with workload perception and/or performance. The findings demonstrated little support for the individual measures included but did highlight the reliability of some variables. The tasks and workload level manipulations in this study were demonstrated as successful across all variables.

After accounting for the effect of task difficulty on performance, the EEG workload index remained a significant predictor of task performance, showing a negative association with task performance. This indicates that as workload increased, participants' performance decreased. Similarly, total TLX score also showed an inverse relationship with task performance. These findings are consistent with cognitive load theory (e.g., Wickens, 1984), which hypothesizes that higher mental workloads impair task performance and supports the inclusion of EEG, specifically the theta-to-beta ratio, in OSM (Raufi & Longo, 2022). Unfortunately, the low quality of the ECG data precluded its inclusion in any meaningful analyses.

Of the individual differences variables measured, education level was the only one to significantly impact task performance and the relationship was negative. Further exploration of education level and the subscale scores on the TLX showed a positive relationship between education level and frustration, suggesting higher education levels corresponded to higher reported levels of frustration. It may be that these frustration levels impacted the level of task performance, thus, lending toward education level as a significant predictor of task performance. Research on workload and psychological stress in real-world settings has shown that those of

higher education levels manage task complexity and workload better than those of lower education levels but conversely are more negatively impacted by psychological stress (Schoger, 2023). In this study, workload was primarily manipulated by throughput of stimuli presentation (increases in stimuli quantity and speed of presentation) which does not necessarily translate to the type of workload complexity and psychological stress studied in real-world settings. It is more likely that, in this study, education level is redundant with the frustration component of the total TLX scores. Further exploration is needed to confirm the replicability of this finding and its true nature with regard to task performance. This finding also highlights the importance of clear definitions in terms of workload manipulations and stressors.

The findings of this study align with the previously published results regarding individual differences, cognitive workload, and task performance. Specifically, a re-analysis of data from four separate studies with common data elements found that anxiety and depression symptoms as well as abstract reasoning capability, correlated with cognitive workload. However, these relationships were not strong enough to influence the utility of physiological measurements to detect state changes or performance deficits (Kelley et al., 2023).

The primary limitation of this study was the resultant sample size. Data loss was underestimated to a large degree, yielding a less than desirable sample size for this type of analysis approach. The research team at Clemson University continues to collect data and thus a re-analysis will be conducted when additional data becomes available. An additional limitation is the lack of medical history information given the impact that medications and medical conditions have on physiological data. Further analyses on the dataset are warranted to explore longitudinal changes over the course of the task. These types of analyses will assist in identification of thresholds to predict when workload (or an element of workload such as frustration) rises to the point where performance deficits are imminent and detrimental to overall mission success.

Conclusion

The primary objective of this study was to evaluate which individual difference variables are most critical in predicting performance. The key findings from this study include the support for use of EEG in predicting changes in cognitive state and performance deficits, the lack of support for many of the hypothesized confounding individual differences, and the potential impact of variations in education level on workload perception. In the short-term, this information translates to refinements to model development that will be able to better predict how individuals will respond to stressors imposed. In the long-term, this study contributes to the overall development of OSM algorithms in order to improve predictive validity and reliability across aviators.

References

- American Psychiatric Association. (2000). *DSM-IV: Diagnostic and statistical manual of mental disorders* (4th Ed., Text Revision).
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory-II*. Psychological Corporation.
- Borghini, G., Aricó, P., Astolfi, L., Toppi, J., Cincotti, F., Mattia, D., Cherubino, P. Vecchiato, G., Maglione, A. G., Graziani, I., & Babiloni, F. (2013). Frontal EEG theta changes assess the training improvements of novices in flight simulation task. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society's Annual Conference*, 6619–6622. <https://doi.org/10.1109/EMBC.2013.6611073>
- Bulea, T. C., Kim, J., Damiano, D. L., Stanley, C. J., & Park, H. S. (2015). Prefrontal posterior parietal and sensorimotor network activity underlying speed control during walking. *Frontiers in Human Neuroscience*, 9, 247.
- Causse, M., Dehais, F., & Pastor, J. (2011). Executive functions and pilot characteristics predict flight simulator performance in general aviation pilots. *The International Journal of Aviation Psychology*, 21(3), 217–234.
- Charles, R. L., & Nixon, J. (2019). Measuring mental workload using physiological measures: A systematic review. *Applied Ergonomics*, 74, 221–232.
- Feng, C., Wanya, X., Yang, K., Zhuang, D., & Wu, X. (2018). A comprehensive prediction and evaluation method of pilot workload. *Technology and Health Care*, 26, S65–S78. <https://doi.org/10.3233/THC-174201>
- Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24, 381–395.
- Goodenough, R. R., Brooks, J. O., Crisler, M. C., & Rosopa, P. J. (2012). Developing a driving simulator based functional object detection task. *Occupational Therapy in Health Care*, 26(4), 240–256. <https://doi.org/10.3109/07380577.2012.721575>
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 139–183.
- Hidalgo-Muñoz, A. R., Béquet, A. J., Astier-Juvenon, M., Pépin, G., Fort, A., Jallais, C., Tattegrain, H., & Gabaude C. (2019). Respiration and heart rate modulation due to competing cognitive tasks while driving. *Frontiers in Human Neuroscience*, 12(525). <https://doi.org/10.3389/fnhum.2018.00525>
- Hussain, I., Young, S., & Park, S. (2021). Driving-induced neurological biomarkers in an advanced driver-assistance system. *Sensors*, 21(21), 6985. <https://doi.org/10.3390/s21216985>

- Kaida M., Takahashi T., Åkerstedt A., Nakata Y., Otsuka T., Haratani K., & Fukasawa. (2006). Validation of the Karolinska Sleepiness Scale against performance and EEG variables. *Clinical Neurophysiology*, 117, 1574–81.
- Kelley, A. M., McAtee, A., Duffy, M., & Feltman, K. A. (2023). *Evaluation of inter-subject variability in physiological metrics and workload perception: Implications for operator state monitoring* (USAARL-TECH-TR--2023-016). U.S. Army Aeromedical Research Laboratory.
- Kessler, R. C., Adler, L., Ames, M., Demler, O., Faraone, S., Hiripi, Howes, M. J., Jin, R., Secnik, K., Spencer, T., Ustun, T. B., & Walter, E. E. (2005). The World Health Organization adult ADHD Self-Report Scale (ASRS): A short screening scale for use in the general population. *Psychological Medicine*, 35, 245–256.
- Kurimori, S., & Kakizaki, T. (1995). Evaluation of work stress using psychological and physiological measures of mental activity in a paced calculating task. *Industrial Health*, 33, 7–22.
- Liang, K. Y. & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear model. *Biometrika*, 73, 13–22.
- Lowe, E., Brooks, J., Bacon, J., Baskins, A., Belle, L., Bendigeri, V., Blackburn, C., Collins, M., Crisler, M., Drouin, N., Evans, D., Hirth, V., Jackson, L., Jenkins, C., Melnrick, K., Mossey, M., Neese, M., Pierce, S., Rosopa, P., ... Zhang S. (2018). Using a driving simulator to teach drivers how to use a left foot accelerator. *News Brake*, 42(2), 15–19.
- Luu, T. P., Nakagome, S., He, Y., & Contreras-vidal, J. (2017). Real-time EEG-based brain-computer interface to a virtual avatar enhances cortical involvement in human treadmill walking. *Scientific Reports*, 7, 8895
- Ma, X., Monfared, R., Grant, R., & Goh, Y. M. (2024). Determining cognitive workload using physiological measurements: Pupillometry and heart-rate variability. *Sensors*, 24(6), 2010.
- Markowski, D, Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., Scholzel, C., & Chen, S. A. (2021). Neurokit2: A Python toolbox for neurophysiological signal processing. *Behavior Research Methods*, 53,1689–1696.
- Mullen, T., Kothe, C. A., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., & Jung, T. P. (2013). Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG. *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Osaka)*, 2184–2187.
- Mullen, T. R., Kothe, C. A., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., Jung, T. P., & Cauwenberghs, G. (2015). Real-time neuroimaging and cognitive monitoring using wearable dry EEG. *IEEE Transactions on Biomedical Engineering*, 62, 2553–2567.

- Nkurikiyeyezu, K., Yokokubo, A., & Lopez, G. (2019). Importance of individual differences in physiological-based stress recognition models. In *Proceedings of the 2019 15th International Conference on Intelligent Environments*, 37–43.
- Raufi, B., & Longo, L. (2022). An Evaluation of the EEG alpha-to-theta and theta-to-alpha band ratios as indexes of mental workload. *Frontiers in Neuroinformatics*, 16, 861967.
- Sassaroli, A., Zheng, F., Hirshfield, L. M., Girouard, A., Solovey, E. T., Jacob, R. J. K., & Fantini, S. (2008). Discrimination of mental workload levels in human subjects with functional near-infrared spectroscopy. *Journal of Innovative Optical Health Sciences*, 01(02). <https://doi.org/10.1142/S1793545808000224>
- Schoger, L. I. (2023). Coping with work-related stressors: does education reduce work-related stress? *Journal of Public Health*, 1–12.
- Schulz, C. M., Schneider, E., Fritz, L., Vockeroth, J., Hapfelmeier, A., Wasmaier, M., Kochs, E. F., & Schneider, G. (2011). Eye tracking for assessment of workload: A pilot study in an anaesthesia simulator environment. *British Journal of Anaesthesia*, 106(1), 44–50.
- Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. TEA Ediciones.
- Stuiver, A., Brookhuis, K. A., de Waard, D., & Mulder, B. (2014). Short-term cardiovascular measures for driver support: Increasing sensitivity for detecting changes in mental workload. *International Journal of Psychophysiology*, 92, 35–41. <http://dx.doi.org/10.1016/j.ijpsycho.2014.01.010>
- Tarabay, R., & Abou-Zeid, M. (2018). Assessing the effects of auditory-vocal distraction on driving performance and physiological measures using a driving simulator. *Transportation Research Part F*, 58, 351–364. <https://doi.org/10.1016/j.trf.2018.06.026>
- Teo, G., Matthews, G., Reinerman-Jones, L., & Barber, D. (2020). Adaptive aiding with an individualized workload model based on psychophysiological measures. *Human-Intelligent Systems Integration*. <https://doi.org/10.1007/s42454-019-00005-8>
- Unni, A., Ihme, K., Jipp, M., & Rieger, J. W. (2017). Assessing the driver's current level of working memory load with high density functional near-infrared spectroscopy: A realistic driving simulator study. *Frontiers in Human Neuroscience*, 11(167). <https://doi.org/10.3389/fnhum.2017.00167>
- Van Orden, K. F., Limbert, W., Makeig, S., & Jung, T. P. (2001). Eye activity correlates of workload during a visuospatial memory task. *Human Factors*, 43, 111–121.
- Warm, J. S., & Parasuraman, R. (2007). Cerebral hemodynamics and vigilance. In R. Parasuraman & M. Rizzo (Eds.), *Neuroergonomics: The brain at work* (pp. 146-158). Oxford University Press.

- Wheelock, M. D., Harnett, N. G., Wood, K. H., Orem, T. R., Granger, D. A., Mrug, S., & Knight, D. C. (2016). Prefrontal cortex activity is associated with biobehavioral components of the stress response. *Frontiers in Human Neuroscience*, 10(583). <https://doi.org/10.3389/fnhum.2016.00583>
- Wickens, C. D. (1984). Multiple resources in attention. In R. Parasuraman and R. Davies, (Eds.), *Varieties of attention*. Academic Press.
- Wu, Y., Miwa, T., & Uchida, M. (2017). Using physiological signals to measure operator's mental workload in shipping – an engine room simulator study. *Journal of Marine Engineering & Technology*, 16(2), 61–69. <https://doi.org/10.1080/20464177.2016.1275496>
- Xiaoli, F., Chaoyi, Z., Xin, Z., Hong, L., & Wei, Z. (2020). Assessment of mental workload based on multi-physiological signals. *Technology and Health Care*, 28(S1), 67–80.
- Zhao Z., & Zhang Y. (2018). SQI quality evaluation mechanism of single-lead ECG signal based on simple heuristic fusion and fuzzy comprehensive evaluation. *Frontiers in Physiology*, 9, 727. <https://doi.org/10.3389/fphys.2018.00727>

U.S. ARMY AEROMEDICAL RESEARCH LABORATORY



FORT NOVOSEL, ALABAMA

Optimizing

**HUMAN PROTECTION
AND PERFORMANCE**
since 1962

All of USAARL's science and technical informational documents are available for download from the Defense Technical Information Center.

<https://discover.dtic.mil/results/?q=USAARL>



U.S. ARMY



FUTURES COMMAND



MRDC