Biomedical Signal Processing and Control 74 (2022) 103529

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Archetypal physiological responses to prolonged wakefulness

Matthew S. Daley [a,*], Krystina Diaz [a], Hugo F. Posada-Quintero [b], Youngsun Kong [b], Ki Chon [b], Jeffrey B. Bolkhovsky [a]

[a] Warfighter Performance, Naval Submarine Medical Research Laboratory, Groton, CT 06340, USA
[b] Department of Biomedical Engineering, University of Connecticut, Storrs, CT 06269, USA

ARTICLE INFO

Keywords:
Wearables
Physiological monitoring
Machine Learning
Cluster analysis
SAFTE
Mixture model
Heart Rate Variability
Electrodermal analysis
Eye tracking
Galvanic skin response

ABSTRACT

Acute sleep deprivation is a common workplace and lifestyle hazard known to affect physiological and cognitive performance. Previous research emphasizes the effect that sleep deprivation has on cognitive performance and often correlates changes in physiological parameters to measured performance. One potential problem with this approach is that there is some ambiguity about the true physiological response to sleep deprivation as the cognitive task employed to elicit some performance measurements may introduce confounds. This work aims to supplement previous works by analyzing physiological metrics from the domains of heart rate variability, electrodermal activity, and eye tracking across 25 h of continual wakefulness collected during a baseline measurement period that did not include a concurrent cognitive task to identify metrics sensitive, but not necessarily specific to, time awake. Of the 78 physiological metrics computed, 31 were shown to be significantly affected by time awake according to a Friedman test ($\alpha = 0.05$). The patterns of the 31 significant metrics were input into an unsupervised clustering algorithm, yielding five archetypal patterns of behavior over 25 h of wakefulness. Of the five archetypal patterns observed, two were well explained by the Sleep, Activity, Fatigue, and Task Effectiveness model of sleep deprivation and effectiveness.

1. Introduction

Acute sleep deprivation, defined as a bout of prolonged wakefulness, is well-established to decrease cardiovascular health [1–3], decrement cognitive performance [4–6], increase risk the of motor vehicle accidents [7–9], and contribute to workplace mishaps [10,11]. Good sleep practices would ideally prevent these negative effects [12,13]; however, acute sleep deprivation remains a persistent issue, especially considering the nature of occupations for which good sleep practices are difficult to sustain (e.g., shift workers). In these non-ideal circumstances, efforts to prevent these negative outcomes are aimed at developing the ability to discern the early signs of acute sleep deprivation and intervene. Non-invasive monitoring provides an easy-to-deploy toolset that is capable of detecting signs of acute sleep deprivation. For example, in HRV, the standard deviation (SD) of interbeat intervals has been shown to increase significantly during 24 h of prolonged wakefulness [14,15], and in EDA the time-varying metric of sympathetic activity (TVSymp) has correlated with cognitive performance during prolonged wakefulness [16,6]. These relationships between these metrics and cognitive performance are useful for estimating the onset of performance decrement. The extraction of these physiological metrics is difficult to standardize and manage, however.

Building on our prior work of using non-invasive physiological monitoring to predict latent aspects of cognitive performance, the Naval Submarine Medical Research Laboratory (NSMRL) has developed and compiled a physiological signal processing suite for use in MATLAB [17]. At present, this suite processes interbeat interval waveforms (from electrocardiogram or photoplethysmography), EDA signals, and gaze tracking inputs, with minimal preprocessing, to produce evidence-based HRV [18,19], EDA [20–22], and ET [23–25] metrics, respectively. Allowing for easy extraction and management of 78 physiological metrics (see Appendix A for a breakdown of these extracted metrics). This processing suite was developed with the intention of maintaining consistency and transparency with the computation of physiologically-derived metrics.

Of particular interest is the relationship among these physiological metrics, prolonged wakefulness, and cognitive task performance. Previous works that have investigated this relationship commonly use prolonged wakefulness to drive performance impairment [26,27,16,28,29]. Results have indicated that metrics of heart rate

* Corresponding author.
E-mail address: Matthew.S.Daley.CTR@mail.mil (M.S. Daley).

https://doi.org/10.1016/j.bspc.2022.103529
Received 16 September 2021; Received in revised form 8 December 2021; Accepted 21 January 2022
Available online 10 February 2022

variability [14], eye-tracking [29,26,28], and electrodermal activity [16] can be used to estimate cognitive task decrement during sleep deprivation. A common criticism of these works is the ambiguity as to which physiological metrics are sensitive to prolonged wakefulness and which are sensitive to other latent factors related to performance decrement (e.g., distraction, boredom). Therefore, the ability to infer what a physiological metric contributes to a prediction becomes obfuscated. For these reasons, the authors believe it is beneficial to explore the within-subject variability of physiological metrics during prolonged wakefulness in the absence of a concurrent performance-based task. We hypothesize that 1) some metrics of HRV, EDA, and ET will be sensitive to prolonged wakefulness; 2) sensitive metrics will exhibit common patterns during wakefulness that can be sorted into a limited number of archetypal clusters; and 3) existing mathematical models of effectiveness during sleep deprivation can be used discern drivers of these patterns (e.g. time awake, circadian, or both). The objective of this work is to examine metrics of HRV, EDA, and ET that are commonly used to predict cognitive performance and determine what, if any, aspects of prolonged wakefulness make a metric a potent predictor. A long-term goal of this work is to understand the sensitivity of these metrics to wakefulness in order to better craft models aimed at estimating circadian processes or cognitive performance in general.

2. Methods

Data from the physiological dataset analyzed in this paper have been previously published; the methods entailed using metrics of face and eye tracking to classify immediate performance as "normal" or "impaired" (Daley et al., 2020) based on PVT performance. Previously published data was collected during the PVT, the data presented here was collected during the baseline.

2.1. Subjects

Twenty healthy participants (13 male, 7 female; 19–32 years old) were recruited from the University of Connecticut (UConn). Subjects were compensated for their participation in the study. The study was approved by the Institutional Review Board (IRB) of UConn as the IRB of record for NSMRL (protocol No. H14-328) in compliance with all applicable Federal regulations governing the protection of human subjects. All subjects gave informed written consent in accordance with the Declaration of Helsinki.

2.2. Protocol

This study used a repeated measures design, with 13 blocks spaced by 2 h over 24 h. The day prior to the experimental protocol, each participant arrived at the experimental facility and completed practice sessions of the task battery; consisting of a psychomotor vigilance task (PVT), a series of speaking tasks, and a Multi-Attribute Task Battery II administration (MAT-B II). Practice sessions continued until a performance plateau was reached. Within 2 h of waking on the day of the experiment, participants arrived at the facility with the expectation that they would remain onsite for the duration of the 24-hour protocol. For each block, participants donned a set of Tobii Pro Eye Tracking glasses [30] and calibrated them according to the manufacturer's instructions. Participants also donned Biopac Bionomadix ECG[2] and EDA[3] sensors, with the help of a research assistant. Participants were seated in front of a webcam that would record video for facial action coding system analysis [31]. Eye tracking, ECG, EDA, and facial action data were recorded and time-synchronized using iMotions[1] ("iMotions 7.0"). No filtering was applied during recording. At the start of each test block a minimum of 4 min of baseline measurements—determined to be the minimal time-length to collect at least 8 full cycles of the lowest frequency metrics—were collected while the subject was seated at the data collection computer and not engaged in any task. To assess the effect of prolonged wakefulness on physiological metrics, only the baseline data for each time block is used in the analyses of this work.

2.3. Physiological metrics

NSMRL has developed a suite of custom MATLAB code for the analysis of ET, HRV, and EDA. Facial action data were excluded from this analysis as they were thoroughly analyzed in Kong et al. [32] and the focus of this analysis was intended for wearable technologies. 78 total evidence-based physiological metrics were computed from baseline data. Brief descriptions of the signal processing used to compute these metrics are provided in the following sub-sections, and full diagrams detailing the analysis of each modality can be found in Appendix A. Overall, the suite is intended to be a user-friendly and consistent tool for performing physiological signal processing to obtain evidence-based metrics for a wide array of applications.

2.4. Metrics of eye tracking

Computed metrics of ET were divided into three domains: eyelid-closure-based, gaze fixation-based, and gaze velocity-based (See Fig. 2 in Appendix for analysis diagram). ET data was collected via Tobii pro eyeglasses [30] and recorded using iMotions with a sampling rate of 50 Hz.

Closures, defined as instances wherein the ET signal was lost due to obstruction by the eyelid, are in units of milliseconds. The mean and SD of all eye closure durations were computed as metrics, as well as the frequency of eye closures, and the maximum closure duration. Closures were further divided into blinks (closures less than 100 ms), extended closures (closures between 100 and 500 ms), and superextended closures (closures longer than 500 ms) [23,24]. For each subdivision, the mean and SD of the durations, and frequencies were computed as metrics.

Fixations, defined as instances wherein the ET signal exhibited a stable location and velocity did not exceed 100 degrees per second [33], are in units of milliseconds. The mean and SD of all fixation durations were computed as metrics, as well as the frequency of fixations, and the maximum fixation duration. Fixations were further divided into express fixations (fixations less than 150 ms), normal fixations (fixations between 150 and 900 ms), and overlong fixations(fixations longer than 900 ms) [25]. For each subdivision, the mean and SD of the durations, and the frequency were computed as metrics.

Velocity, defined as the differentiation of the original gaze signal, is in units of "degrees per second." The mean and SD of all eye velocity were computed as metrics. Following this, eye velocity was further divided into: intrafixation velocity, defined as eye velocity during fixation, and interfixation velocity, defined as eye velocity between fixations. For both subdivisions, the mean and SDs were computed as metrics.

2.5. Metrics of heart rate variability via electrocardiogram

Computed metrics of HRV were divided into two domains: time- and frequency-based (See Fig. 3 in Appendix for analysis diagram), and metrics were derived from Camm et al. [18] and Sassi et al. [19]. Electrocardiogram (ECG) data was collected via Biopac Bionomadix[2] hardware and recorded using iMotions with a sampling rate of 500 Hz. The peaks of QRS complexes in the ECG signal were identified using the Pan-Tompkins QRS peak identification algorithm [34]. The location of peaks was used to create the time-series data of the interbeat intervals (also known as 'NN' or 'RR' intervals). The interbeat intervals were resampled to 5 Hz.

In the time domain analysis, mean, SD, median, maximum, minimum, and range of the interbeat intervals were computed as metrics. Next, the derivative of the interbeat intervals was used to determine successive differences. The SD, raw count longer than 50 ms, proportion longer than 50 ms, and the root-mean-square of successive differences were computed as metrics. Lastly, interbeat intervals were converted to heart rate; the mean and SD were computed as metrics.

In the frequency domain, fast Fourier transform (FFT) of the interbeat interval time-series was taken to compute the power spectral density (PSD). The power sums between the ranges of 0 to 0.40 Hz (0 Hz to 0.04 Hz, 0.04 to 0.15 Hz, and 0.15 to 0.40 Hz) were computed as metrics. The ratio of power between the 0.04 to 0.15 Hz and the 0.15 to 0.40 Hz ranges was computed as a metric. The peak frequencies in the aforementioned ranges were computed as metrics. Lastly, the PSD was normalized to total power, and the relative powers in the ranges 0 to 0.04 Hz, 0.04 to 0.15 Hz, and 0.15 to 0.40 Hz were computed as metrics.

2.6. Metrics of electrodermal activity

Computed metrics of EDA were divided into three domains: time domain, frequency domain–time invariant, and frequency domain–time variant (See Fig. 4 in Appendix for analysis diagram) [20]. EDA data was collected via Biopac Bionomadix[3] hardware and recorded using iMotions with a sampling rate of 500 Hz.

In the time domain, mean and SD of the raw EDA signal were computed as metrics [20]. Skin conductance response (SCR) events were identified by detecting peaks in the signal. Mean and SDs of the peaks were computed as metrics [22,20]. The rise time of each peak was computed as the time between the onset of the SCR and the peak amplitude of the SCR [22,20].The mean and SDs of the rise time and area under the curve during rise times were computed as metrics [20]. Mean and SD of the recovery times, defined as the time between the peak of an SCR to 50% the amplitude of the peak, were computed as metrics [22]. The area under the curve between the onset and 50% recovery were computed as the "total area" of an SCR. The mean and SD of the total area were computed as metrics. The general rate of SCR events metric was computed as the number of SCRs divided by the total time. The mean and SDs of the instantaneous peak-to-peak rate, defined as time differences between successive SCRs, were computed as metrics.

In the time-invariant frequency domain, the FFT of the EDA signal was used to compute the PSD of the EDA signal. The power sum and peak frequencies in the following ranges were computed as metrics: 0 to 0.16 Hz, 0.16 to 2.1 Hz, and 0.045 to 0.25 Hz [21,22].

In the time-variant frequency domain, the EDA signal was decomposed into time–frequency components via variable frequency complex demodulation [35]. The power between 0.08 and 0.24 Hz were summed together to yield a bandpass filtered signal. The mean and SD of this filtered signal were computed as metrics.

2.7. Statistical analysis

Each physiological metric was tested for significant differences over hours awake via Friedman test [36] ($\alpha = 0.05, twotailed$). The Friedman test was selected in place of a traditional repeated measures ANOVA due to preliminary analysis showing that 67 of the 78 metrics computed were not normally distributed according to the Anderson-Darling test. Due to the exploratory nature of this analysis, the Benjamini-Hochberg procedure was used as the alpha correction as it is less conservative than the typical Bonferroni correction and thus more appropriate for comparing a large number of metrics, albeit at the cost of increased type I errors [37].

2.8. Unsupervised nonparametric clustering

Metrics that met the criteria for significance were sorted via unsupervised GM modeling clustering to discern archetypal patterns across the 25 h of wakefulness. GM modeling was employed due to the necessity to identify subgroups with similar patterns of behavior with no supervised label [38]. This method of unsupervised clustering is accomplished via the expectation–maximization (EM) algorithm, which iteratively adjusts the parameters of the gaussian distributions until convergence to a maximal likelihood is achieved [39]. Similar to the Friedman test, the relative rankings of the metric data are used in place of the raw values. GM models using 1 to 10 clusters were fit to the data. Since this method of clustering is unsupervised, it is expected that the inclusion of more clusters will result in a higher overall likelihood, therefore, Akaike information criterion (AIC), a relative score of a models likelihood penalized by the number of parameters, was used to determine the optimal number of clusters to describe the patterns seen in the data [40].

2.9. Comparison to SAFTE model components

The Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE) model is a predictive tool for modeling effectiveness during bouts of sleep deprivation [41]. This is accomplished by combining equations intended to model time-awake and circadian processes. The patterns derived from the GM step were compared to the SAFTE model to determine potential driving factors. Specifically, the generalized task effectiveness equation of SAFTE was used to derive generalized linear, circadian, and combined linear-circardian, representing the full SAFTE model, explanatory models [42,41]. Each cluster was fit to the set of explanatory models. Model fit was evaluated using the R^2 metric from the resulting regression and the quality of the candidate models were evaluated using AIC weightings.

3. Results

3.1. Friedman test results

Of the 78 physiological metrics evaluated, 31 showed significant differences as a result of hours awake, after alpha correction. Of those 31, 16 (of 32) belonged to ET, 11 (of 23) belonged to HRV, and 4 (of 23) belonged to EDA. All significant metrics are presented in Table 1, and clusters describing behavior during time awake are depicted in Fig. 1. Table 2 displays the comparisons of clusters to components of the SAFTE model.

Table 1
Significant metrics as determined by the Friedman tests with Benjamini-Hochberg corrections. For each metric the modality, domain, metricID (corresponding to the diagrams in Appendix A), a description, chi-square, p-value, and associated cluster are listed.

Modality	Domain	Metric ID	Metric Description	χ^2	p-value	Cluster ID
HRV	Time Domain	medNN	Median of N—N Intervals	54.7	2.0E-07	D
HRV	Freq. Domain	sumES	Sum of Spectral Power	53.8	2.9E-07	C
EYE	Velocity	sigmaAll	SD of All Velocity	52.5	5.1E-07	C
HRV	Time Domain	sigmaNN	SD of N—N Intervals	51	9.5E-07	B
HRV	Time Domain	muNN	Mean of N—N intervals	49.9	1.5E-06	D
EYE	Velocity	muAll	Mean of All Eye Velocity	46.3	6.1E-06	B
HRV	Freq. Domain	vlfES	Very Low Frequency Power	44.6	1.2E-05	E
HRV	Time Domain	muHR	Mean Heartrate	42	3.3E-05	E
EYE	Velocity	sigmaIntrafixation	SD of Intrafixation Velocity	39.9	7.5E-05	B
EDA	Time Domain	muSCL	Mean Skin Conductance Level	38.4	1.3E-04	A
EYE	Fixation	muAll	Mean Duration of All Eye Fixations	35.6	3.8E-04	A
EYE	Closure	muAll	Mean Duration of All Eye Closures	34.9	4.9E-04	C
EYE	Closure	freqSuperExtended	Frequency of Superextended Eye Closures	33.7	7.6E-04	C
EYE	Closure	muBlink	Mean Duration of Blink Eye Closures	32.8	0.0010	C
EYE	Fixation	maxAll	Maximum Eye Fixation Duration	32.8	0.0010	A
HRV	Time Domain	maxNN	Maximum N—N Intervals	32.2	0.0013	C
HRV	Time Domain	rangeNN	Range of N—N Intervals	31.8	0.0015	B
HRV	Freq. Domain	lfES	Low Frequency Power	31.4	0.0017	D
HRV	Time Domain	sigmaHR	SD of Heartrate	31.2	0.0018	E
EDA	Freq. Domain - Time Invariant	valHF	High Frequency Power	30.8	0.0021	A
EYE	Closure	freqExtended	Frequency of Extended Eye Closures	30.3	0.0025	D
EYE	Closure	freqAll	Frequency of All Eye Closures	29.9	0.0029	C
EDA	Time Domain	allSCR	Mean of All Skin Conductance Response Rates	28.6	0.0046	A
EYE	Closure	maxAll	Maximum Eye Closure Duration	28.1	0.0054	D
EYE	Fixation	sigmaAll	SD of All Eye Fixation Durations	27.7	0.0061	A
EYE	Velocity	muIntrafixation	Mean of Intrafixation Eye Velocity	27.7	0.0062	B
HRV	Freq. Domain	hfES	High Frequency Power	27.4	0.0067	E
EDA	Freq. Domain - Time Varying	muTVSymp	Mean of Time-Varying Sympathetic Index	27.3	0.0069	E
EYE	Closure	sigmaAll	SD of All Eye Closure Durations	26	0.0109	D
EYE	Closure	muSuperextended	Mean Duration of Superextended Eye Closures	25	0.0147	B
EYE	Fixation	sigmaOverlong	SD of Overlong Fixation Durations	24.3	0.0187	A

3.2. Unsupervised clustering analysis

The best-fit model, carrying approximately 99.85% of the cumulative AIC weight, grouped the data into 5 clusters (Fig. 1). Of the 31 available metrics, cluster A is composed of 7, cluster B is composed of 6, cluster C is composed of 7, cluster D is composed of 6, and cluster E is composed of 5. The affiliations between metrics and clusters are listed in Table 1 under "Cluster."

3.3. SAFTE model comparison

Table 2 lists the R^2 and relative AIC weighting for the fit between each cluster and the each explanatory model derived from the SAFTE model. Relative model quality is measured by the AIC weighting with higher percentages representing higher quality models. Model fit is measured by the R^2 value between the cluster and the explanatory model. The full SAFTE model was the highest quality explanatory model for clusters A (R^2 = 0.92; AIC Weight = 89.03%), B (R^2 = 0.85; AIC Weight = 53.53%) and C (R^2 = 0.92; AIC Weight = 96.26%). The linear component was the highest quality explanatory model for cluster D (R^2 = 0.75; AIC Weight = 83.31%). The circadian component was the highest quality explanatory model for cluster E (R^2 = 0.60; AIC Weight = 80.24%).

4. Discussion

In this paper, 78 evidence-based physiological metrics belonging to three modalities—ET, HRV, and EDA—were analyzed for their sensitivity to prolonged wakefulness. Thirty-one metrics were found to have significant differences during prolonged wakefulness based on the results of the Friedman test ($\alpha = 0.05$), and these were further analyzed via unsupervised clustering using GM modeling. The purpose of the latter analysis was to organize the 31 significant metrics into a smaller number of archetypal patterns based on the relative rankings of their values (i.e., ranking of lowest to highest values over time), before then comparing these pattern clusters the widely used SAFTE model for fit.

Of the 31 significant physiological metrics, many overlap with metrics used to build predictive performance models, including models built using this same dataset [26]. For example, many of the current metrics were included, among others (e.g., facial tracking metrics), in building sets of classification models for determining whether an individual's performance was 'normal' or 'impaired' during this same bout of acute sleep deprivation. These metrics include, the mean duration of all eye closures (EYE), SD of N—N intervals (HRV), and mean skin conductance level (EDA) [26]. One critique of those models is the ambiguity as to whether the models could be considered sensitive to performance in general, prolonged wakefulness, or a combination of the two. The information provided by the current analyses enables us to parse apart which metrics included in the classification models were indeed sensitive to time awake from which metrics were not, suggesting sensitivity to the performance decrement and elucidating the differences within metric fluctuations across a period of acute sleep deprivation. However, a clear limitation in the presented work is that there is no control for factors such as boredom influencing the changes in physiological metrics during prolonged wakefulness.

Figure 1 shows the results of the cluster analysis, which yielded 5

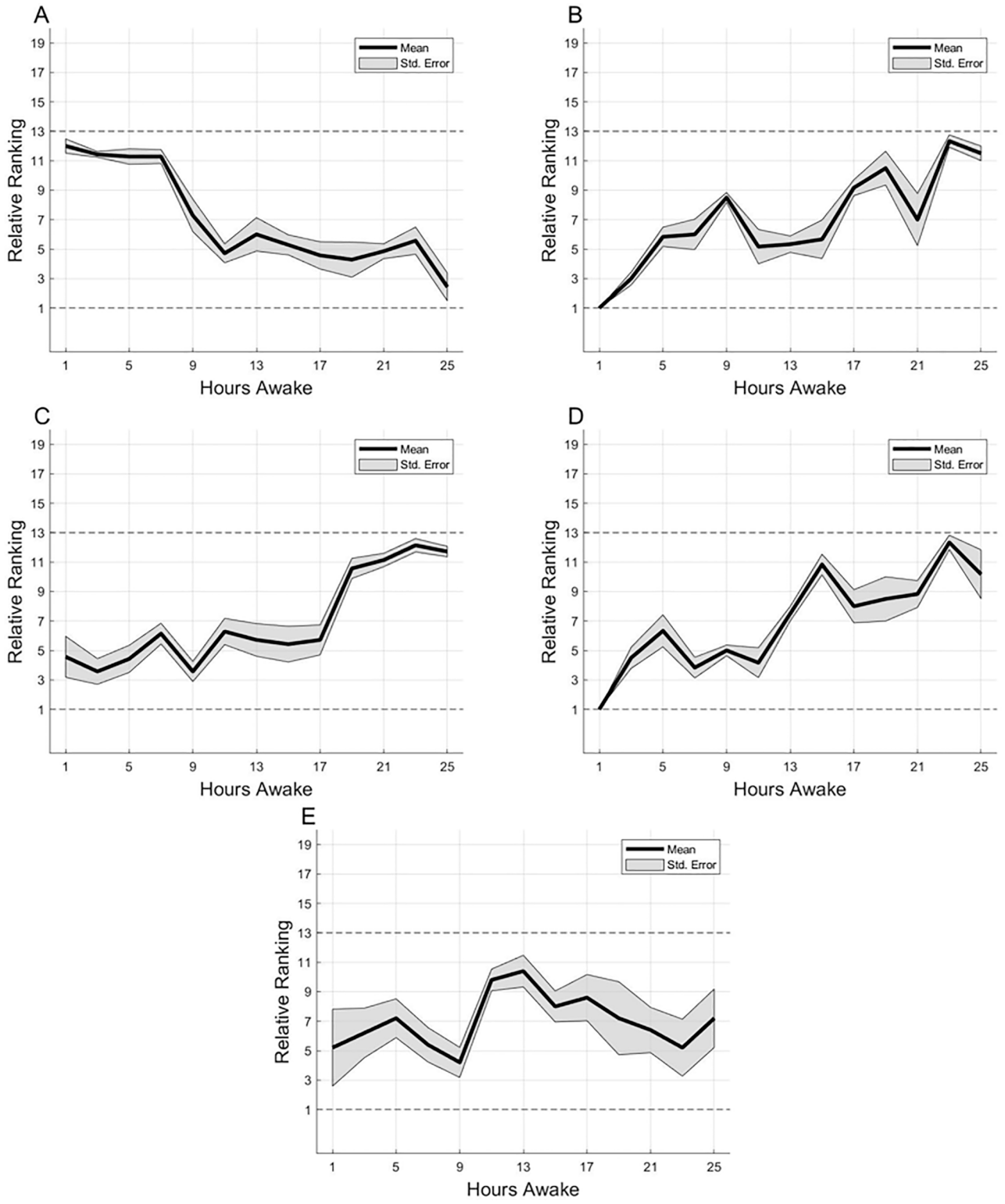

Fig. 1. Clusters of physiological metric patterns over 25 h of sleep deprivation. The solid black line represents the mean relative ranking of indices within the cluster, the grey shaded area represents the standard error, and the dashed black lines demarcate the minimum and maximum values possible for the relative rankings. The affiliations between metrics and clusters are listed in Table 1 under "Cluster."

Table 2
R^2 and Relative AIC weightings (%) of explanatory SAFTE component models fit to clusters.

	Linear Component		Circadian Component		Combined	
	R^2	AIC	R^2	AIC	R^2	AIC
Cluster A	0.80	10.95%	0.61	0.02%	0.92	89.03%
Cluster B	0.71	46.44%	0.31	0.02%	0.85	53.53%
Cluster C	0.76	3.66%	0.68	0.08%	0.92	96.26%
Cluster D	0.75	83.31%	0.37	0.03%	0.83	16.67%
Cluster E	0.04	1.99%	0.60	80.24%	0.63	17.77%

clusters of behavior during 25 h of prolonged wakefulness. To discern the driving factors of the obtained clusters of these metrics, clusters were fitted to explanatory models derived from the SAFTE model of cognitive effectiveness and fatigue. Each component model was fitted to the five individual clusters and evaluated for fit and quality. Clusters A & C have a highly correlated fit with the combined components model ($R^2 = 0.92$ for both models). The combined component model was similarly determined to be the highest quality candidate model to explain clusters A & C (AIC weights = 89.03% & 96.26%, respectively). Cluster B increased over time, similar to cluster C, and exhibited a high correlation to the combined components model ($R^2 = 0.85$), however the relative AIC weights were near evenly distributed between the linear and combined component explanatory models (46.44% and 52.52%, respectively). This suggests that the metrics that compose clusters A, B, & C are strongly influenced by latent circadian processes during prolonged wakefulness. The linear component was determined to be the highest quality explanatory model for cluster D (AIC weights = 83.31%; $R^2 = 0.75$). This would indicate that time awake is the most potent driver of the behavior seen in cluster D, however visual inspection does show that cluster D contain some undulating behavior that is perhaps not well captured by the full SAFTE model. The metrics that compose cluster D are likely affected by prolonged wakefulness and latent circadian processes similar to clusters A, B, & C, however the weaker fit between explanatory model and cluster makes it more difficult to draw a definitive conclusion. Lastly, the circadian explanatory model was determined to be the highest quality fit for cluster E (AIC Weight = 80.25%), however this is accompanied by a relatively weak regression ($R^2 = 0.60$). Unfortunately, the large undulation exhibited in cluster E appears to have a low enough frequency such that it is curtailed by the 25 h data collection period. While this makes it unreasonable to draw a conclusion about the drivers of the metrics in cluster E, it does highlight the need for future work investigating these metrics over longer periods of prolonged wakefulness. A clear limitation for drawing conclusions of the effect of sleep deprivation on physiological patterns is the absence of data during similar time intervals wherein participants were allowed to sleep. Future works supporting these efforts should seek to compare how these physiological metrics behave during both normal sleep and sleep deprivation.

Performance prediction via non-invasive monitoring is a quickly growing field [43,44]. Self-monitoring through personal devices to improve sleep quality [45,46], physical performance [47–49], mental performance [50], and professional productivity [51] are becoming mainstays in how people are using their technologies. ET, HRV, and EDA are among the most common modalities for building performance-predictive models based on physiology and indeed are some of the most prevalent in commercially-available personal monitoring devices.

In this work we have presented a novel use of unsupervised clustering to identify archetypal patterns of physiological metrics during prolonged wakefulness across multiple modalities. This was done to allow patterns of behavior to be identified in an emergent fashion that relies solely on the strength of their similarities and is free from outside bias. Additionally, we have analyzed the fit of these archetypal patterns to components comprising the SAFTE model to identify potential latent drivers of the observed patterns of physiological behaviors (i.e. Time, Circadian Rhythm, or both). The major contribution of this work is the contribution to the current body of knowledge on the effects of prolonged wakefulness on these physiological metrics. Specifically, grouping metrics across differing modalities based on similar patterns and drivers of those patterns. The expected impact of this work is to highlight which metrics are likely sensitive to circadian influences and hours awake, which can better inform model development. For example, in developing a system that estimates circadian alignment based on non-invasive physiological metrics, the metrics in clusters A & C are likely strong candidates, so a developer could choose to compute the mean duration of eye fixations if they have an eye tracker or the mean skin conductance level if they have a electrodermal activity system. Overall, the results presented in this paper contribute to better understanding the impact of prolonged wakefulness and circadian processes on these physiological metrics.

5. Disclaimer

The views expressed in this article reflect the results of research conducted by the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the United States Government.

I am a military Service member or employee of the U.S. Government. This work was prepared as part of my official duties. Title 17, U.S.C., §105 provides that copyright protection under this title is not available for any work of the U.S. Government. Title 17, U.S.C., §101 defines a U. S. Government work as a work prepared by a military Service member or employee of the U.S. Government as part of that person's official duties.

Funding

This work was supported/funded by work unit number F1103 – Joint Program Committee 5 (JPC-5)/Military Operational Medicine Research Program (MOMRP).

CRediT authorship contribution statement

Matthew S. Daley: Writing - original draft, Writing - review & editing, Formal analysis, Methodology, Software. **Krystina Diaz:** Writing - review & editing, Resources. **Hugo F. Posada-Quintero:** Data curation, Investigation. **Youngsun Kong** Data curation, Investigation. **Ki Chon** Conceptualization, Investigation, Supervision. **Jeffrey B. Bolkhovsky** Conceptualization, Investigation, Funding acquistion, Supervision, Validation.

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.

Appendix:. Signal processing diagrams

.

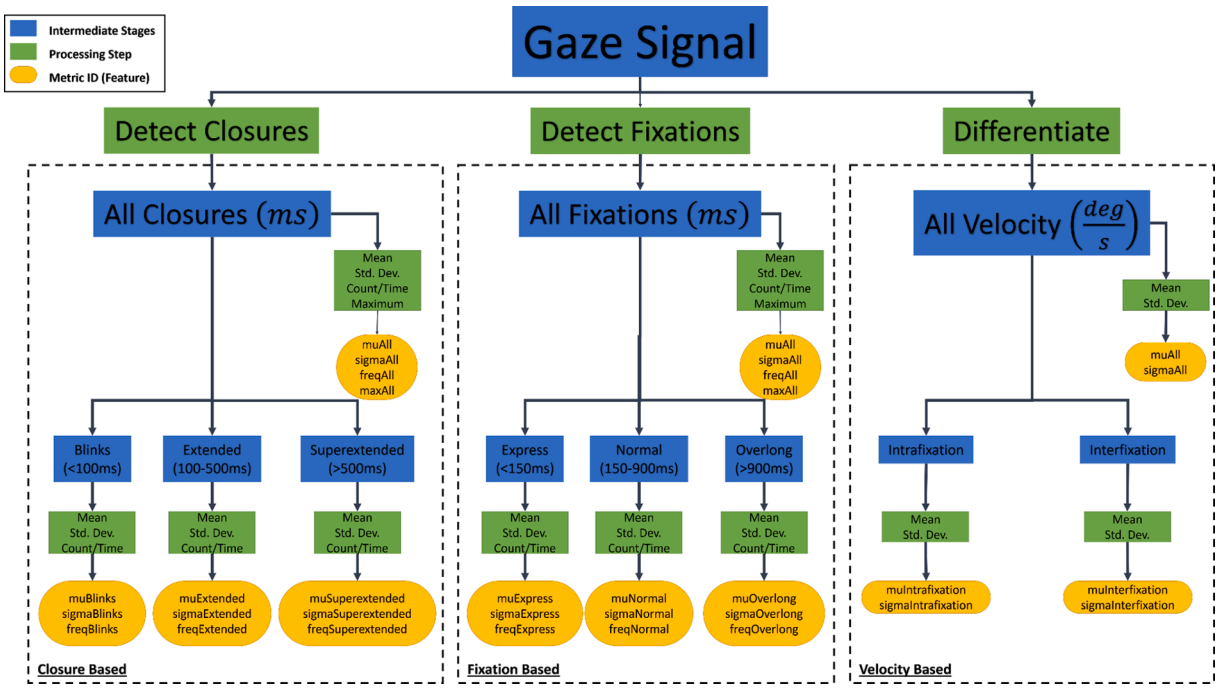

Fig. 2. Eye tracking signal processing diagram. Blue boxes represent intermediate steps of the data. Green boxes represent processing steps (e.g. "mean" or "standard deviation"). Yellow ovals represent the discrete metrics/features yielded from this process. This diagram is intended to be a visual accompaniment of the "Metrics of Eye Tracking" section under Methods.

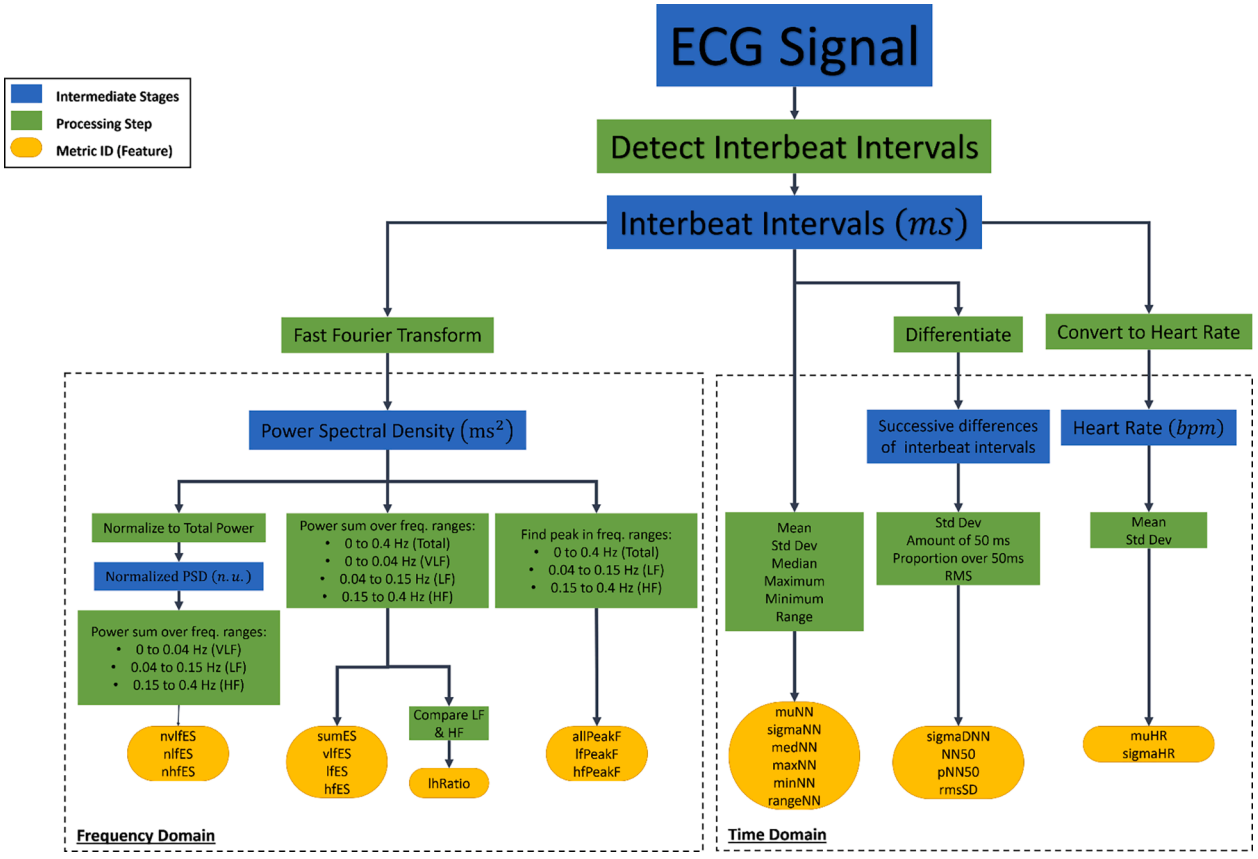

Fig. 3. ECG signal processing diagram. Blue boxes represent intermediate steps of the data. Green boxes represent processing steps (e.g. "mean" or "standard deviation"). Yellow ovals represent the discrete metrics/features yielded from this process. This diagram is intended to be a visual accompaniment of the "Metrics of Heart Rate Variability vis Electrocardiogram" section under Methods.

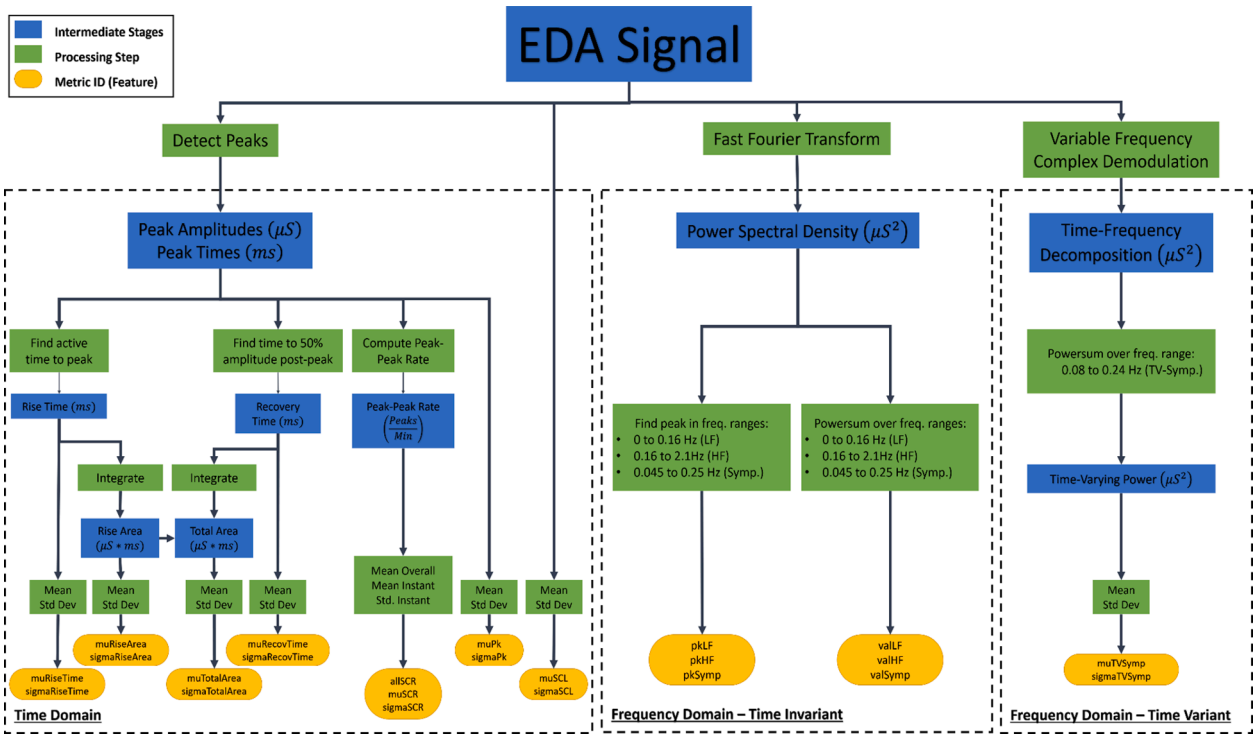

Fig. 4. EDA signal processing diagram. Blue boxes represent intermediate steps of the data. Green boxes represent processing steps (e.g. "mean" or "standard deviation"). Yellow ovals represent the discrete metrics/features yielded from this process. This diagram is intended to be a visual accompaniment of the "Metrics of Electrodermal Activity" section under Methods.

References

[1] N. Covassin, P. Singh, Sleep duration and cardiovascular disease risk: epidemiologic and experimental evidence, Sleep Med. Clinics 11 (1) (2016) 81–89.

[2] D.J. Gottlieb, S. Redline, F.J. Nieto, C.M. Baldwin, A.B. Newman, H.E. Resnick, N. M. Punjabi, Association of usual sleep duration with hypertension: the sleep heart health study, Sleep 29 (8) (2006) 1009–1014.

[3] S.F. Quan, B.V. Howard, C. Iber, J.P. Kiley, F.J. Nieto, G.T. O'Connor, P.W. Wahl, The sleep heart health study: design, rationale, and methods, Sleep 20 (12) (1997) 1077–1085.

[4] P. Alhola, P. Polo-Kantola, Sleep deprivation: impact on cognitive performance, Neuropsychiatr. Dis. Treat. 3 (5) (2007) 553–567.

[5] R.G. Angus, R.J. Heslegrave, Effects of sleep loss on sustained cognitive performance during a command and control simulation, Behav. Res. Methods Instruments Comput. 17 (1) (1985) 55–67.

[6] H.F. Posada-Quintero, J.B. Bolkhovsky, M. Qin, K.H. Chon, Human performance deterioration due to prolonged wakefulness can be accurately detected using time-varying spectral analysis of electrodermal activity, Hum. Factors 60 (7) (2018) 1035–1047.

[7] A.I. Pack, A.M. Pack, E. Rodgman, A. Cucchiara, D.F. Dinges, C.W. Schwab, Characteristics of crashes attributed to the driver having fallen asleep, Accid. Anal. Prev. 27 (6) (1995) 769–775.

[8] J.C. Stutts, J.W. Wilkins, J. Scott Osberg, B.V. Vaughn, Driver risk factors for sleep-related crashes, Accid. Anal. Prev. 35 (3) (2003) 321–331.

[9] B.C. Tefft, Acute sleep deprivation and culpable motor vehicle crash involvement, Sleep 41 (10) (2018) zsy144.

[10] D.R. Gold, S. Rogacz, N. Bock, T.D. Tosteson, T.M. Baum, F.E. Speizer, C. A. Czeisler, Rotating shift work, sleep, and accidents related to sleepiness in hospital nurses, Am. J. Public Health 82 (7) (1992) 1011–1014.

[11] K. Uehli, A.J. Mehta, D. Miedinger, K. Hug, C. Schindler, E. Holsboer-Trachsler, Jörg.D. Leuppi, N. Künzli, Sleep problems and work injuries: a systematic review and meta-analysis, Sleep Med. Rev. 18 (1) (2014) 61–73.

[12] M.A. Grandner, L. Hale, M. Moore, N.P. Patel, Mortality associated with short sleep duration: the evidence, the possible mechanisms, and the future, Sleep Med. Rev. 14 (3) (2010) 191–203.

[13] M. Hirshkowitz, K. Whiton, S.M. Albert, C. Alessi, O. Bruni, L. DonCarlos, N. Hazen, J. Herman, E.S. Katz, L. Kheirandish-Gozal, D.N. Neubauer, A. E. O'Donnell, M. Ohayon, J. Peever, R. Rawding, R.C. Sachdeva, B. Setters, M. V. Vitiello, J.C. Ware, P.J. Adams Hillard, National Sleep Foundation's sleep time duration recommendations: methodology and results summary, Sleep Health 1 (1) (2015) 40–43.

[14] E.C.P. Chua, W.Q. Tan, S.C. Yeo, P. Lau, I. Lee, I.H. Mien, J.J. Gooley, Heart rate variability can be used to estimate sleepiness-related decrements in psychomotor vigilance during total sleep deprivation, Sleep 35 (3) (2012) 325–334.

[15] A.U. Viola, C. Simon, J. Ehrhart, B. Geny, F. Piquard, A. Muzet, G. Brandenberger, Sleep processes exert a predominant influence on the 24-h profile of heart rate variability, J. Biol. Rhythms 17 (6) (2002) 539–547.

[16] H.F. Posada-Quintero, J.B. Bolkhovsky, N. Reljin, K.H. Chon, Sleep deprivation in young and healthy subjects is more sensitively identified by higher frequencies of electrodermal activity than by skin conductance level evaluated in the time domain, Front. Physiol. 8 (2017) 409.

[17] Mathworks, *MATLAB Release 2019a,* pp. Natick, MA: Mathworks Inc., 2019.

[18] A.J. Camm, M. Malik, J.T. Bigger, G. Breithardt, S. Cerutti, R.J. Cohen, D.H. Singer, "Heart rate variability: Standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology," 1996.

[19] R. Sassi, S. Cerutti, F. Lombardi, M. Malik, H.V. Huikuri, C.K. Peng, R. Macfadyen, Advances in heart rate variability signal analysis: joint position statement by the e-Cardiology ESC Working Group and the European Heart Rhythm Association co-endorsed by the Asia Pacific Heart Rhythm Society, Ep Europace 17 (9) (2015) 1341–1353.

[20] J.J. Braithwaite, D.G. Watson, R. Jones, M. Rowe, A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments, Psychophysiology 49 (1) (2013) 1017–1034.

[21] H.F. Posada-Quintero, J.P. Florian, A.D. Orjuela-Cañón, T. Aljama-Corrales, S. Charleston-Villalobos, K.H. Chon, Power spectral density analysis of electrodermal activity for sympathetic function assessment, Ann. Biomed. Eng. 44 (10) (2016) 3124–3135.

[22] W. Boucsein, D.C. Fowles, S. Grimnes, G. Ben-Shakhar, W.T. Roth, M.E. Dawson, D. L. Filion, Society for psychophysiological research ad hoc committee on electrodermal measures. Publication recommendations for electrodermal measurements, Psychophysiology 49 (8) (2012) 1017–1034.

[23] G.W. Ousler 3rd, M.B. Abelson, P.R. Johnston, J. Rodriguez, K. Lane, L.M. Smith, Blink patterns and lid-contact times in dry-eye and normal subjects, Clin. Ophthalmol. (Auckland, NZ) 8 (2014) 869.

[24] J.D. Rodriguez, K.J. Lane, G.W. Ousler, E. Angjeli, L.M. Smith, M.B. Abelson, Blink: characteristics, controls, and relation to dry eyes, Curr. Eye Res. 43 (1) (2018) 52–66.

[25] R. Schleicher, N. Galley, S. Briest, L. Galley, Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? Ergonomics 51 (7) (2008) 982–1010.

[26] M.S. Daley, D. Gever, H.F. Posada-Quintero, Y. Kong, K. Chon, J.B. Bolkhovsky, Machine learning models for the classification of sleep deprivation induced performance impairment during a psychomotor vigilance task using indices of eye and face tracking, Front. Artif. Intell. 3 (2020) 17.

[27] J.C. Liu, S. Verhulst, S.A. Massar, M.W. Chee, Sleep deprived and sweating it out: The effects of total sleep deprivation on skin conductance reactivity to psychosocial stress, Sleep 38 (1) (2015) 155–159.

[28] K.J. Heaton, A.L. Maule, J. Maruta, E.M. Kryskow, J. Ghajar, Attention and visual tracking degradation during acute sleep deprivation in a military sample, Aviat. Space Environ. Med. 85 (5) (2014) 497–503.

[29] L.M. Rowland, M.L. Thomas, D.R. Thorne, H.C. Sing, J.L. Krichmar, H.Q. Davis, G. Belenky, Oculomotor responses during partial and total sleep deprivation, Aviat. Space Environ. Med. 76 (7) (2005) C104–C113.
[30] Tobii Pro, *Tobii Pro Glasses 2 Best Practices Guide.,* pp. Reston, VA: Tobii Pro., 2019.
[31] L. Kulke, D. Feyerabend, A. Schacht, A Comparison of the affectiva iMotions facial expression analysis software with EMG for identifying facial expressions of emotion, Front. Psychol. 11 (2020) 329.
[32] Y. Kong, H.F. Posada-Quintero, M.S. Daley, K.H. Chon, J. Bolkhovsky, Facial features and head movements obtained with a webcam correlate with performance deterioration during prolonged wakefulness, Attent. Percept. Psychophys. 83 (1) (2021) 525–540.
[33] J. Otero-Millan, X.G. Troncoso, S.L. Macknik, I. Serrano-Pedraza, S. Martinez-Conde, Saccades and microsaccades during visual fixation, exploration, and search: foundations for a common saccadic generator, *J. Vis.* 8(14) 2008 12-21.
[34] J. Pan, W.J. Tompkins, A real-time QRS detection algorithm, IEEE Trans. Biomed. Eng. (3) (1985) 230–236.
[35] H.F. Posada-Quintero, J.P. Florian, Á.D. Orjuela-Cañón, K.H. Chon, Highly sensitive index of sympathetic activity based on time-frequency spectral analysis of electrodermal activity, Am. J. Physiol.-Regulatory Integr. Comparative Physiol. 311 (3) (2016) R582–R591.
[36] M. Friedman, The use of ranks to avoid the assumption of normality implicit in the analysis of variance, J. Am. Stat. Assoc. 32 (200) (1937) 675–701.
[37] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. Royal Stat. Soc. Ser. B (Methodological) 57 (1) (1995) 289–300.
[38] G.J. McLachlan, S.U. Chang, Mixture modeling for cluster analysis, Stat. Methods Med. Res. 15 (5) (2004) 347–361.
[39] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, J. Roy. Stat. Soc.: Ser. B (Methodol.) 39 (1) (1977) 1–22.
[40] H. Akaike, Information theory and an extension of the maximum likelihood principle. *Selected papers of Hirotogu Akaike,* pp. 199-213, 1998.
[41] S.R. Hursh, D.P. Redmond, M.L. Johnson, D.R. Thorne, G. Belenky, T.J. Balkin, D. R. Eddy, Fatigue models for applied research in warfighting, Aviat. Space Environ. Med. 75 (3) (2004) A44–A53.
[42] S.R. Hursh, Modeling Sleep and Performance within the Integrated Unit Simulation System (IUSS), *US Army Soldier Systems Command, Technical Report TR-98 L, 26.,* 1998.
[43] S. Park, S. Jayaraman, Wearables: Fundamentals, advancements, and a roadmap for the future, *Wearable Sens.* 2021 3-27.
[44] J.M. Peake, G. Kerr, J.P. Sullivan, A critical review of consumer wearables, mobile applications, and equipment for providing biofeedback, monitoring stress, and sleep in physically active populations, Front. Physiol. 9 (2018) 743.
[45] Z. Chen, M. Lin, F. Chen, N.D. Lane, G. Cardone, R. Wang, A.T. Campbell, Unobtrusive sleep monitoring using smartphones, in: In 2013 7th International Conference on Pervasive ComputIng Technologies for Healthcare and Workshops, 2013, pp. 145–152.
[46] S. Haghayegh, S. Khoshnevis, M.H. Smolensky, K.R. Diller, R.J. Castriotta, Performance assessment of new-generation Fitbit technology in deriving sleep parameters and stages, Chronobiol. Int. 37 (1) (2020) 47–59.
[47] G. Dobreff, P. Revisnyei, G. Schuth, G. Szigeti, L. Toka and A. Pašić, "Physical Performance Optimization in Football," in *International Workshop on Machine Learning and Data Mining for Sports Analytics*, Springer, 2020.
[48] N. Harms, The Impact of WHOOP Technology on Sleep, Recovery, and Performance in NAIA Baseball Players, 2018.
[49] R.T. Li, S.R. Kling, M.J. Salata, S.A. Cupp, J. Sheehan, J.E. Voos, Wearable performance devices in sports medicine, Sports Health 8 (1) (2016) 74–78.
[50] B.A. Hickey, T. Chalmers, P. Newton, C.-T. Lin, D. Sibbritt, C.S. McLachlan, R. Clifton-Bligh, J. Morley, S. Lal, Smart devices and wearable technologies to detect and monitor mental health conditions and stress: a systematic review, Sensors 21 (10) (2021) 3461, https://doi.org/10.3390/s21103461.
[51] J. Khakurel, H. Melkas, J. Porras, Tapping into the wearable device revolution in the work environment: a systematic review, Inf. Technol. People 31 (3) (2018) 791–818.