

Research Article

State Anxiety Biomarker Discovery: Electrooculography and Electrodermal Activity in Stress Monitoring

Jadelynn Dao¹, Ruixiao Liu¹, Sarah Solomon², Samuel Solomon¹

1. Andrew and Peggy Cherng Department of Medical Engineering, Division of Engineering and Applied Science, California Institute of Technology, United States; 2. Adult Psychiatry Residency, Dartmouth–Hitchcock Medical Center, Lebanon, United States

Anxiety has become a significant health concern affecting mental and physical well-being, with state anxiety—a transient emotional response—linked to adverse cardiovascular and long-term health outcomes. This research explores the potential of non-invasive wearable technology to enhance the real-time monitoring of physiological responses associated with state anxiety. Using electrooculography (EOG) and electrodermal activity (EDA), we have reviewed novel biomarkers that reveal nuanced emotional and stress responses. Our study presents two datasets: 1) EOG signal blink identification dataset BLINKEO, containing both true blink events and motion artifacts, and 2) EOG and EDA signals dataset EMOCOLD, capturing physiological responses from a Cold Pressor Test (CPT). From analyzing blink rate variability, skin conductance peaks, and associated arousal metrics, we identified multiple new anxiety-specific biomarkers. SHapley Additive exPlanations (SHAP) were used to interpret and refine our model, enabling a robust understanding of the biomarkers that correlate strongly with state anxiety. These results suggest that a combined analysis of EOG and EDA data offers significant improvements in detecting real-time anxiety markers, underscoring the potential of wearables in personalized health monitoring and mental health intervention strategies. This work contributes to the development of context-sensitive models for anxiety assessment, promoting more effective applications of wearable technology in healthcare.

Corresponding author: Samuel Solomon, ssolomon@caltech.edu

Introduction

Despite being a short-term response, state anxiety (s-anxiety) has emerged as a significant factor impacting long-term health outcomes. Researchers have linked sustained s-anxiety with adverse cardiovascular effects^[1], underscoring its profound effects on mental and physical health. Approximately 23.1% of American adults experience some form of diagnosable mental disorder^[2] and 74% of American adults reported experiencing stress-related health issues within a given month^[3], illustrating the widespread impact of anxiety-induced stress. Reliable biomarkers are essential for capturing the complexities of s-anxiety, enabling more precise and effective models.

Non-invasive wearable technology has the potential to transform health monitoring by continuously capturing physiological data through real-time sensor measurements^[4]. These devices collect a broad array of metrics, yielding critical insights into the body's responses to anxiety. The ability to seamlessly collect large amounts of health-related data opens new ways to study and build an understanding of the onset and progression of anxiety, enabling more effective interventions and advancing our knowledge of human health. Identifying reliable biomarkers of s-anxiety offers a promising pathway to real-time health monitoring using wearable biosensors that can detect subtle physiological changes not immediately obvious in raw signal data.

The cold pressor test (CPT) is a widely utilized experimental method for studying anxiety responses in controlled settings. Participants immerse their hand in ice-cold water (0–4°C), eliciting a sympathetic nervous system response. This test reliably induces physiological markers of anxiety^{[5][6]}, such as increased heart rate and sweat production. Other techniques, such as public speaking simulations and mental arithmetic tasks, also provoke anxiety and can be used to identify reliable biomarkers.

Physiological responses to s-anxiety and arousal have been extensively documented, revealing clear links between emotional states and indicators such as blink rate variability^[7] and stress-induced sweating^[8]. The two-factor model of emotion, developed by Schachter and Singer, suggests that emotions arise from physiological arousal and subsequent cognitive interpretation. This model underscores that physiological responses are interpreted within a contextual framework, which are further hidden in indirect biomarkers for specific emotional experiences. For instance, fatigue, which affects the blink conditions, can intensify physiological arousal, directly impacting how the brain interprets anxious states. Such contextual cues are crucial for understanding s-anxiety in real-world settings, but they are often filtered out or controlled for in existing studies. Wearable devices offer a way to contextualize these arousal states

dynamically. Through advanced human-machine interfaces, wearables can monitor how individuals respond to their environments, integrating data on physical responses to build a richer understanding of s-anxiety. This approach allows wearables to add depth to biomarker data, interpreting physiological responses in respect to real-time contextual cues and providing a more comprehensive view of emotional states.

Research shows that blink rates tend to increase under difficult mental tasks or anxiety-provoking situations^{[9][10]}, reflecting activation of the autonomic nervous system.

Electrooculography (EOG) captures electrical signals produced by eye movements, allowing for the detection of blink-related patterns. But EOG signals are often filtered out in stress studies to improve clarity of other signals^[11], potentially overlooking valuable information related to emotional arousal. Studies suggest that specific components of EOG signals can be analyzed to extract physiological markers of s-anxiety, highlighting the need for further research into EOG biomarkers. Moreover, fatigue—closely associated with emotional arousal—provides an additional avenue for understanding anxiety through EOG features^{[12][13]}. Studies examining EOG signals in the context of drowsiness reveal correlations between blink frequency, blink duration, and stages of fatigue^[10], highlighting a non-invasive method for tracking emotional arousal over time. Given the interplay between fatigue and anxiety, this relationship prompted our investigation into how fatigue-related features within EOG signals may serve as indirect indicators of anxiety, offering new opportunities for nuanced and comprehensive stress monitoring.

Similarly, stress has a pronounced effect on sweat production. Emotional sweating, triggered by the sympathetic nervous system, occurs in response to psychological stressors rather than temperature changes^{[14][15]}. Electrodermal activity (EDA) is a method that measures changes in skin conductance. Under emotional arousal and stress, body sweats and skin conductance increases. Previous studies often rely on basic features like median values^[16] or the phasic component of the EDA signal, focusing on nonspecific skin conductance responses (SCR) to correlate with self-reported s-anxiety^[17] scores. In such studies, peaks in the phasic signal exceeding 0.01 μS were counted as responses, and the frequency of these nonspecific SCRs per minute served as the primary measure for physiological s-anxiety. EDA primarily reflects the *magnitude* of emotional arousal without distinguishing between positive and negative affective states^[18]. In other words, a high SCR could result from excitement or stress, making it challenging to interpret EDA data as a standalone indicator of anxiety. This underscores the importance

of using EDA in combination with other physiological markers^[19], such as heart rate variability or blink rate, to gain a more comprehensive picture of an individual's emotional and physiological state. A more methodical exploration of signal characteristics found in EDA and EOG signals reveal nuanced physiological markers that strongly correlate with s-anxiety.

Currently, no widely accepted biomarkers reliably assess anxiety across diverse contexts, highlighting the need for continued exploration. Researchers have tested markers like heart rate variability, skin conductance, and blink rate, but results often vary due to individual differences and contextual influences. While many studies report that depressed patients exhibit reduced EDA responses, indicating diminished autonomic nervous system activity, some research presents conflicting findings. These discrepancies are attributed to variations in study designs, methodologies, and the influence of factors such as antidepressant treatment on EDA measurements^[20].

While machine learning models have shown promise in detecting anxiety, their black-box nature limits interpretability, making it difficult to validate findings across diverse populations^[21]. By introducing additional context-sensitive biomarkers, we aim to enhance the reliability and transparency of anxiety assessments, making models more applicable to real-world scenarios.

In our research, we leverage EOG and EDA data to develop a comprehensive, real-time model of s-anxiety. We have compiled two distinct datasets for this purpose. The first dataset, *BLINKEO*, consists of EOG signal features from samples characterized by peak-like patterns, annotated to differentiate natural blink events from extraneous noise and wire movement artifacts. The second dataset, *EMOCOLD*, contains time-series EOG and EDA signals along with demographic data and stress responses elicited by the Cold Pressor Test (CPT). Using interpretability techniques such as SHAP (SHapley Additive exPlanations), we identify and quantify specific biomarkers within the EOG and EDA data, with a focus on blink rate variability and sweat-related stress indicators. Our approach goes beyond simple anomaly detection by uncovering nuanced, anxiety-specific physiological markers informed by the two-factor model of emotion. This research contributes to a more detailed understanding of stress mechanisms, with the potential to improve mental health interventions and enable personalized, context-specific stress management strategies with wearable technology.

Methods

Description of Question

This research aims to identify reliable, interpretable biomarkers of s-anxiety through EOG and EDA data for real-time stress monitoring.

Blink Identification EOG Dataset (BLINKEO) Data Collection

To create the *BLINKEO Blink Identification Dataset*, EOG data was collected and analyzed to differentiate natural blinks from noise or wire movements. To detect vertical eye movements using EOG, one electrode was positioned above the eye and another below it, aligning on the vertical axis. This configuration captures the corneo-retinal potential changes associated with upward and downward eye movements. 65 trials involving repeated blinking under controlled conditions where no extraneous movement occurred. Additionally, 19 trials lasting between 30 seconds and 2 minutes were conducted under conditions with no blinking, but with deliberate wire movements. These trials provided a baseline for accurately distinguishing noise artifacts from genuine blink events.

Trial Lengths and Identified Peaks

Label	No. Sessions	Total Time (s)	No. Peaks Detected	No. Peaks After Filtering
Blink	65	12103.14	6792	4734
Wire Movement	19	2007.75	5704	203

Table 1. Characteristics of blink and wire movement trials in the Blink Identification Dataset. This table summarizes the number of independent sessions, cumulative recording time, and peak detection results before and after literature- supported blink peaks filtering for both blink and wire movement events.

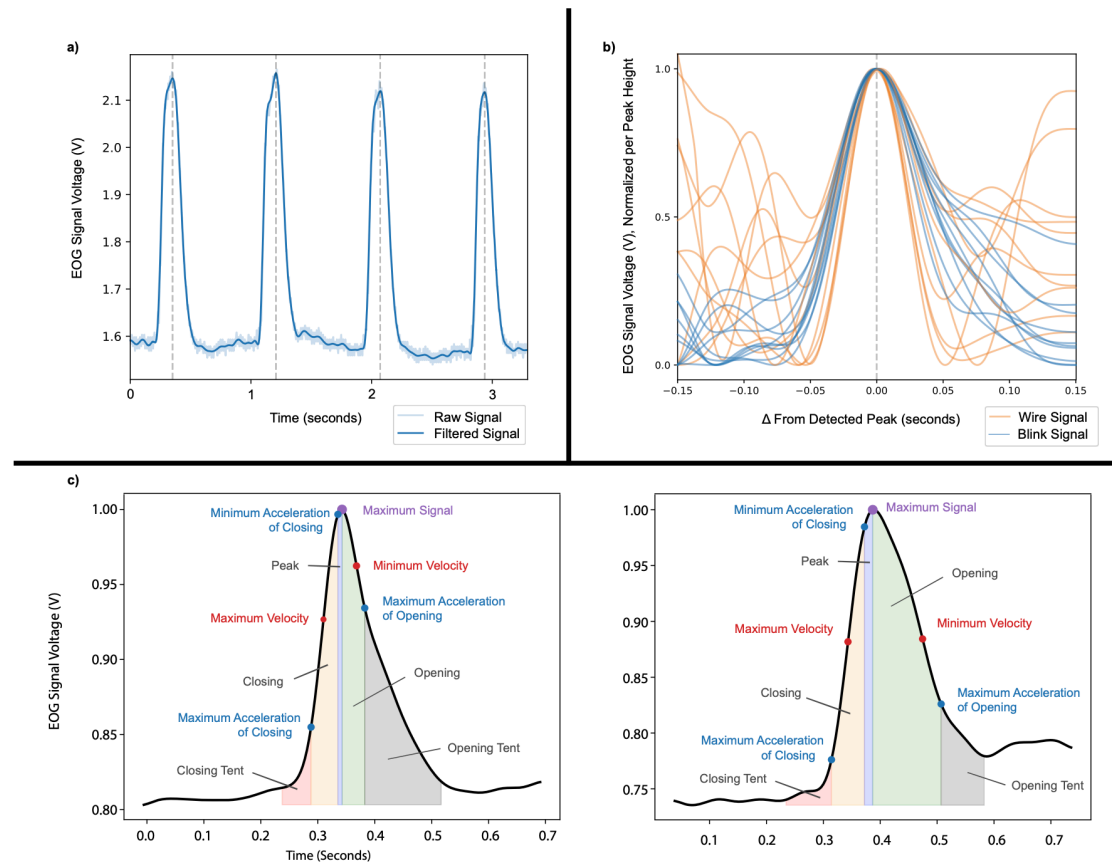


Figure 1. 1a. Blink Peak Examples from the BLINKEO dataset. The red dotted lines indicate the center of the peak, extracted by the peak detection method outlined in this section. 1b. Blink examples (blue) plotted against wire examples (green), as filtered EOG Voltage signals, normalized per peak between 0 and 1. Peaks are time-aligned by time, in seconds, from the center of peak. Wire signals typically have higher variability. 1c. A singular blink peak. The purple dot marks the peak of the blink event, while the outer edges of the red and grey shaded sections represent the boundaries used for feature extraction. These boundaries are determined by identifying the nearest minimum on each side of the peak, providing a precise range for analyzing blink characteristics.

To preprocess the EOG data, motion artifacts were identified and removed, to make the data suitable for downstream features. A fifth-order low-pass Butterworth filter was applied to isolate low-frequency components indicative of meaningful physiological signals. This was followed by a Savitzky-Golay filter for additional smoothing, which preserved essential features while reducing minor signal fluctuations.

Peak detection was performed using the Scipy Signal *find_peaks* function, identifying peaks with a prominence exceeding 0.1 with a peak width greater than 0.04 seconds^[22]. (blinks typically last between

0.1 and 0.4 seconds^[23], averaging around 0.25 seconds). To focus on blink-like events, we applied criteria based on established blink characteristics: a maximum peak width of 0.5 seconds and a minimum peak height of 0.05 volts. We compared the signal quality after this initial peak detection with that obtained using conventional blink filtering methods. Traditional filtering techniques frequently overlook subtle blink patterns or introduce artifacts during data cleaning, potentially compromising accuracy. In contrast, a learned-feature approach refines this process by reducing noise and enhancing the precision of true blink identification within the dataset. Figure 1a illustrates examples of detected blink peaks from the BLINKEO dataset, with red dotted lines marking the center of each peak. This figure demonstrates the effectiveness of the peak detection method described in this section, highlighting its ability to accurately locate and extract the central point of each blink event during blink trials.

However, wire movements can also produce peak-like shapes, which poses challenges for this filtering method. While effective in controlled or low-noise environments, the filter is easily triggered by noisy conditions, where artifacts such as wire movements may mimic blink patterns. Figure 1b presents time series segments of both blink and wire movement examples that have been classified as blinks under the current filtering approach, overlaid for comparison. The figure shows that wire movements exhibit greater variability in the regions surrounding the peak, as well as in the overall shape of the peak itself. Current approaches are unable to distinguish between true blinks and wire artifacts, underscoring the limitations of the method in noisier environments.

For each detected peak, baseline values were calculated to provide a reference point for the signal's amplitude. This involved locating the nearest minimum values on either side of the peak by performing binary search with a window size of up to 0.5 seconds in the left and right direction from the peak observed (see pseudocode in Supplementary 1). It recursively narrows down the search range to locate a local minimum, while avoiding minor fluctuations.

After establishing the baseline points, we extracted a comprehensive set of amplitude- independent features for each peak. These features include blink duration and various acceleration and velocity metrics, as utilized in prior EOG feature extraction and peak signal analysis studies^{[24][25]}. A total of 32 peak-related features and label are stored as examples in the dataset, with labels distinguishing natural blinks from noise artifacts.

Figure 1c illustrates examples of EOG signals from two independent singular blink events, with distinct sections of the peak highlighted for clarity. The purple dot at the peak center represents the highest voltage point, detected by the peak detection algorithm. Red dots indicate local maxima in velocity, while

blue dots show local acceleration points. Shaded regions in different colors represent key sections of the blink, such as the rising and falling phases, as well as acceleration and deceleration phases. This segmentation captures various aspects of the blink shape, this detailed segmentation provides valuable insights into the blink dynamics, enabling the extraction of relevant blink-related features.

We establish bounds for each feature by discretizing its range into 50 intervals. This discretization splits the feature's values into small, equally spaced segments, enabling a systematic exploration of possible lower and upper bounds that optimize model accuracy. The process begins by identifying the minimum and maximum values of each feature. The range between these values is divided by the bin count (50), yielding an incremental "step size," or delta value, for testing. This delta value determines how much the threshold will shift at each iteration when exploring the bounds. To identify the best lower bound, the algorithm starts from the minimum value and iteratively adds the delta value (e.g., 0.2) to the threshold, testing each increment by culling data points below it and evaluating the model's accuracy with the adjusted dataset. The lower bound with the highest accuracy is selected as the optimal starting point for that feature.

The search then proceeds to find an optimal upper bound, beginning with the maximum value and reducing it by increments of the delta value until reaching the previously identified lower bound. This decremental approach ensures the upper bound remains above the lower bound. Each new threshold is applied to the dataset, and the accuracy is recorded. The upper bound yielding the best accuracy becomes the final threshold for that feature.

The individually optimized lower and upper bounds for each feature are compiled into a list, representing the complete culling thresholds that maximize model performance across the dataset. By discretizing each feature's range into 50 intervals, the *individual search* method ensures a thorough yet efficient exploration of potential thresholds.

Emotion, EOG, and EDA Monitoring in Cold Pressor Conditions (EMOCOLD) Data Collection

The data collection process employed wearable sensors to record EDA and EOG signals from participants during controlled stress trials. For EOG, electrodes were placed above and below one eye to measure vertical eye movements. The EDA sensor was positioned on the forehead to monitor skin conductance responses. Sixteen participants (N=16) between ages 26–31 took part in the study, and demographic

information, including race and gender, was collected and is summarized in Table 1. Each trial lasted about 10-15 minutes was divided into three phases: baseline, CPT (Cold Pressor Test), and recovery. The length of the trial and the data used for feature analysis is as detailed in Table 2.

EOG signals were recorded using a three-electrode configuration designed to capture vertical eye movements, particularly blink activity. Electrodes were positioned as follows: one above the eye, one below the eye, and a reference electrode in the middle of the forehead. This setup effectively captured vertical eye movement signals, with the reference electrode providing signal stability and reducing noise.

For EDA, a single electrode was placed on the forehead to measure changes in skin conductance associated with sympathetic nervous system activation. The forehead was chosen for its accessibility and stable conductance properties, making it suitable for detecting stress-related physiological changes in skin conductance.

Participants wore the device throughout the Cold Pressor Test (CPT) trials, which were conducted to simulate acute stress events. The trials included both physical and environmental stressors. In the cold-water trials, participants immersed their hand in cold water, inducing a sympathetic nervous response. This provided a controlled means of eliciting stress responses. Additionally, exercise trials involved brief physical exertion to stimulate physiological stress, allowing for a comparison of stress responses across different types of stimuli.

The design of these trials facilitated the collection of time-series data, capturing participants' physiological reactions to both physical exertion and environmental stressors, thereby providing a comprehensive view of their autonomic responses under varying stress conditions.

Table 2. Characteristics of trials in the EMOCOLD dataset. *a-b.* Demographic details of the study participants, including *a.* race and *b.* assigned sex. *c-d.* *c.* Summary of trial durations across different experimental phases. *d.* Summary of the duration of time EDA and EOG features are collected from, across different experimental phases. For each phase—Baseline, Cold Pressor Test (CPT), and Recovery—both tables list the minimum, 25th percentile, median, 75th percentile, and maximum duration (in seconds).

Race	Count
Asian	11
Hispanic or Latino	2
White	1
Middle Eastern or North African	1
Black or African American	1

a.

Assigned Sex	Count
Male	11
Female	5

b.

Experiment	Length (Seconds)				
	Min	25th Percentile	Median	75th Percentile	Max
Baseline	245.6	274.0	281.7	310.0	414.8
CPT	261.9	278.4	290.4	306.4	358.0
Recovery	238.6	252.8	261.3	278.1	311.2

c. Trial

Experiment	Length (Seconds)				
	Min	25th Percentile	Median	75th Percentile	Max
Baseline	167.5	172.1	177.0	182.3	194.0
CPT	160.6	165.0	177.2	184.1	188.2
Recovery	157.1	168.4	172.1	180.3	191.9

d. Feature Collection

At each stage of the experiment—baseline, Cold Pressor Test (CPT), and recovery— participants completed an excerpt of the Positive and Negative Affect Schedule (PANAS) and the State-Trait Anxiety Inventory (STAI-State) to assess their emotional responses. The PANAS measures both positive emotions (e.g., Inspired, Attentive) and negative emotions (e.g., Upset, Nervous) on a 5-point scale, capturing general mood states. The STAI-State survey, consisting of items such as “I feel tense” and “I feel worried”, assesses immediate anxiety levels on a 4-point scale, making it particularly useful for tracking s-anxiety in response to acute stress. The survey recorded at each stage is detailed in the supplementary section (Supplementary 2). Administering these surveys at each stage allowed us to correlate physiological data from EDA and EOG signals with subjective emotional responses, providing a comprehensive view of how participants’ mood and anxiety levels evolved across stress phases.

Electrodermal Activity (EDA) Signal Segmentation

The tonic and phasic components of skin conductance reveal different aspects of autonomic arousal, with the tonic level representing a stable baseline and the phasic response capturing transient, stimulus-driven changes. Tonic signals vary significantly across individuals due to factors like skin type and hydration, making them challenging to analyze consistently in relation to specific stress events. Phasic responses, however, reflect rapid fluctuations in skin conductance directly tied to acute stress or anxiety-inducing stimuli, characterized by quick rises and gradual declines.

Phasic signals were divided into rise and fall phases to capture the dynamics of the skin conductance response, which is indicative of sympathetic nervous system activation. Specifically, peaks were detected

by identifying rapid increases in skin conductance (rise phases) followed by gradual decreases (fall phases). To preprocess the EDA data and extract the phasic signal, motion artifacts were identified and removed, to make the data suitable for downstream features. A first-order low-pass Butterworth filter was applied to isolate low-frequency components indicative of meaningful physiological signals.

This signal was divided into windows of 1 second in length. Each section was analyzed to determine key features, such as mean value, signal range, and standard deviation. 15 features were extracted from these windows, and the full list of features and their definitions can be found in Supplementary 3. These features are critical for quantifying the intensity and duration of autonomic arousal events, providing valuable insights into stress response dynamics. The segmentation process allowed for the extraction of detailed temporal characteristics of each skin conductance event, facilitating a comprehensive analysis of physiological arousal under stress.

Electrooculography (EOG) Signal Segmentation

In analyzing Electrooculography (EOG) signals, we segmented the data to isolate individual blink peaks, which are essential for understanding blink dynamics in response to stress.

From these peaks, we extracted 33 features, including blink duration, amplitude, frequency, and various acceleration and velocity metrics. A comprehensive list of these features and their definitions is provided in Supplementary 4.

Discussion

Blink Identification Results

Building upon the non-intentional blink signal processing outlined by previous research^{[26][27]}, a feature bounding analysis aligned closely with the article's approach of differentiating blink events based on slope and derivative features. By using blink duration alone as a feature, we achieved a classification accuracy of 87.46% and an F1 score of 79.99% in distinguishing blinks from wire movements (see Supplementary 5). This suggests that feature extraction can yield strong performance metrics. Even without deep learning techniques, finding the right markers of blink peaks can reach the same efficacy of the article's outlined slope-based signal differentiation.

In our approach, we systematically evaluate all possible combinations of five selected features to optimize classification performance for distinguishing blink events from wire movements. For each feature

combination, we apply a breadth-first search (BFS) traversal to explore and fine-tune the upper and lower bounds of each feature, seeking the configuration that maximizes classification accuracy.

The BFS traversal begins with initializing the bounds for each feature to cover its entire observed range, ensuring that no data points are culled at the outset. Each feature range is discretized into 15 bins, allowing for incremental adjustments to the bounds with a step size (delta) calculated as the range divided by the number of bins. These initial bounds are stored as a “node” in the BFS queue, representing a unique culling configuration.

During each iteration of the BFS traversal, we dequeue a culling configuration and calculate its classification accuracy and F1 score using a performance function. If the configuration achieves a higher accuracy than previously recorded, it becomes the current optimal configuration. The BFS traversal then generates neighboring configurations by slightly tightening the bounds for each feature—either increasing the lower bound or decreasing the upper bound by the computed delta. Each of these neighboring configurations, if unvisited, is added to the queue for further exploration.

This BFS traversal continues until all relevant bound configurations for the current feature combination are evaluated. The outcome is an empirically derived set of feature bounds that maximizes classification performance for each combination of features. By applying this process across all combinations of the selected five features, we ensure a comprehensive search of the parameter space, yielding an optimal culling pipeline tailored for precise blink detection. This method demonstrates the robustness of combining BFS with multi-feature analysis to achieve a high-performing, data-driven classification model.

In our approach, we select combinations of five high-quality features and use a breadth-first search (BFS) traversal to optimize their combined bounds for maximal classification performance. For each combination, BFS systematically explores adjustments to the upper and lower bounds of each feature, identifying the optimal configuration that yields the highest accuracy and F1 score.

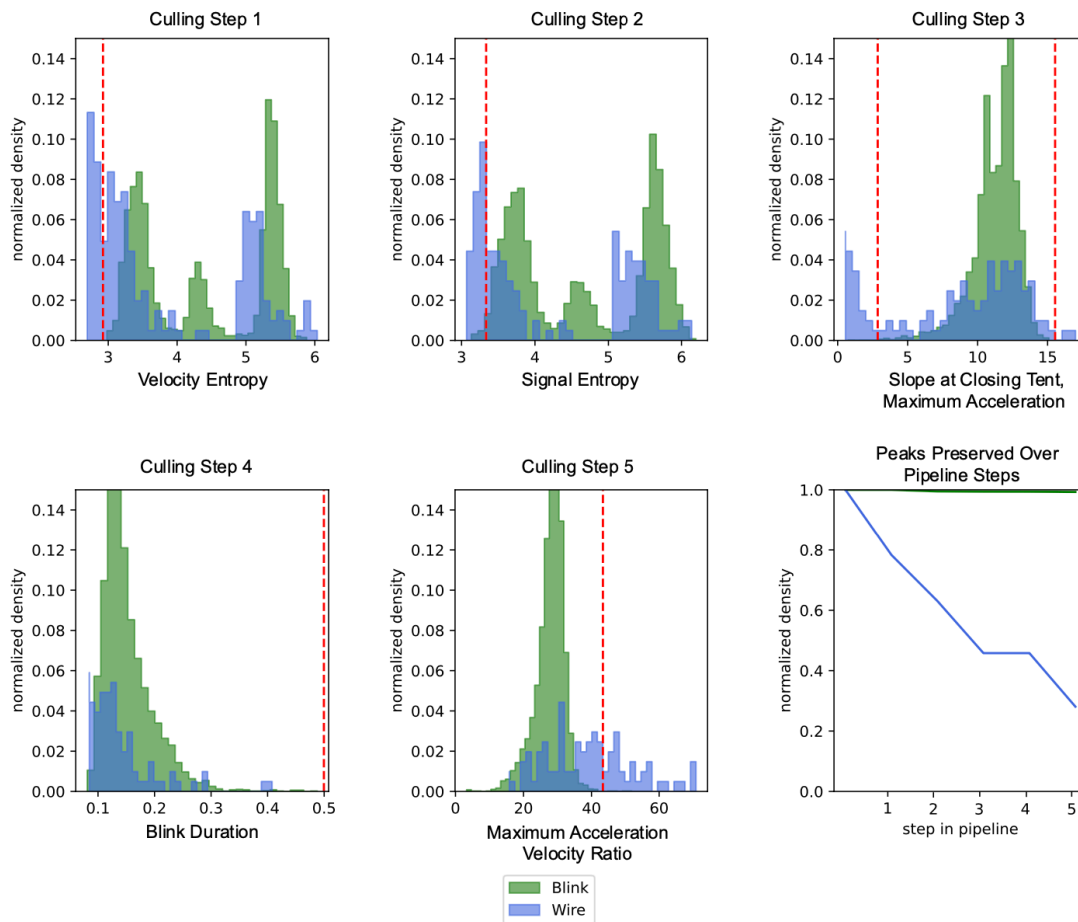


Figure 2. Optimal Culling Steps for Differentiating Blink Events from Wire Movement Artifacts in EOG Data.

This figure presents the sequential culling steps optimized to achieve the highest accuracy and F1 score in distinguishing blink events (green) from wire artifacts (blue) in EOG data. Each subplot demonstrates a unique culling step, applying specific feature thresholds to progressively refine the data. The final subplot, "Peaks Preserved Over Culling Pipeline," illustrates the proportion of retained peaks at each stage for both blink and wire signals, showcasing the efficacy of each step in isolating genuine blink events.

The optimal feature combination achieved an accuracy of 98.17% and an F1 score of 87.34%, utilizing five key features that capture distinctive characteristics of blink dynamics. These features include: *Velocity Entropy*, the entropy of the first derivative of the signal, which measures the variability and complexity of the blink motion; *Signal Entropy*, the entropy of the signal itself, providing a broader assessment of the overall blink pattern; *Slope at Closing Tent, Maximum Acceleration*, the maximum acceleration during the closing phase of a blink, which isolates the rapid deceleration typical of blink closure; *Blink Duration*,

representing the total time span of the blink event; and *Maximum Acceleration Velocity Ratio*, the ratio between the maximum acceleration and maximum velocity during the closing phase, which captures the relationship between these peak dynamics, indicative of voluntary eye closure. The results of each feature bounding step, against the BLINKEO labelled examples, is shown in Figure 2.

These features together form a comprehensive representation of blink characteristics, enabling differentiation of blinks from other signal types in the culling pipeline. This highlights how strategically selected bounds on multiple features, when combined, result in high classification performance without relying on complex algorithms.

EMOCOLD Analysis

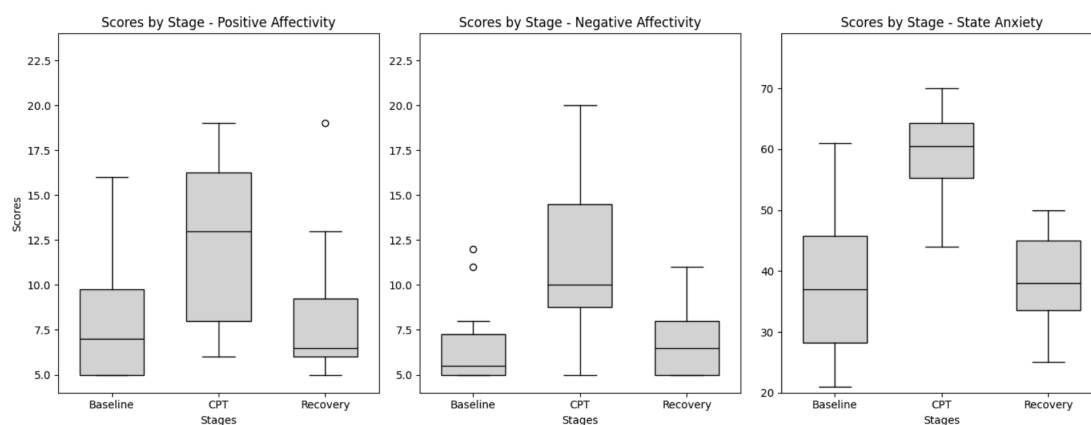


Figure 3. User-reported survey responses during each stage of the trial. During CPT, participants showed higher levels of positive affectivity, negative affectivity, and stage anxiety. Elevated levels recovered to baseline responses when participants took their hand out of the cold-water bath during the recovery phase.

The EMOCOLD dataset analysis highlights significant physiological and emotional responses to acute stress induced by the Cold Pressor Test (CPT). Figure 3 shows participants' aggregated self-reported survey scores for positive affectivity, negative affectivity, and s-anxiety across the three trial stages: Baseline, CPT (Cold Pressor Test), and Recovery. In figure 3, for each stage, survey responses were summarized and visualized using box plots, which display the distribution of scores. Positive Affectivity and Negative Affectivity are scored on a scale of 5-25, and State Anxiety is scored on a scale of 20-80.

Participants reported increased positive and negative affectivity, as well as elevated s- anxiety during CPT, which returned to baseline during recovery. This dual affective response suggests heightened

arousal may include both alertness and discomfort. The recovery phase indicates effective autonomic regulation, as emotional states normalized once the stressor was removed. These findings validate CPT as a method for inducing short-term anxiety.

SHAP (SHapley Additive exPlanations) Analysis

SHAP (SHapley Additive exPlanations) analysis is a method used to explain the output of machine learning models by breaking down the prediction into contributions from each feature. SHAP values are based on Shapley values from cooperative game theory, which attribute the impact of each feature on the model's output by treating each feature as a "player" in a game and calculating its contribution to the final prediction.

In this study, SHAP analysis was performed on combinations of 5 features, highlighting the significance of how certain biomarkers, used together, reveal more prominent interactions and effects on model predictions. This approach underscores that certain biomarkers, while potentially less impactful individually, can demonstrate substantial importance when analyzed as part of a group. By evaluating these interactions, we understand how combinations of features can provide insights into the model's behavior that single-feature analyses might overlook.

SHAP evaluates the quality of a set of features by considering their collective contribution to the model's predictions, measured through the mean absolute SHAP values across the dataset. A high-quality set of features is one where the combination of features consistently demonstrates substantial importance, as indicated by elevated mean absolute SHAP values. This benchmark reflects not only the magnitude of individual contributions but also the degree to which the features, as a group, interact to enhance the predictive power of the model.



Figure 4. 4a. SHAP Analyses for optimal combinations of five EOG Features (top row) and five EDA Features (bottom row), for Positive Affectivity (left column), Negative Affectivity (middle column), and State Anxiety (right column). **4b.** SHAP Analysis of Feature Combinations. This analysis explores the quality of distinguishing different affectivity levels using different sets of features. This is an example of five EOG features and their impact on the Negative Affectivity score.

Substituting one key feature with another can reveal new interdependencies among remaining features, thereby enhancing the model's interpretability.

The SHAP value maps provide insights into how various EOG features used in combination, and EDA features used in combination, contribute to predictions for positive affectivity negative affectivity, and s-anxiety. Each SHAP sub-plot illustrates the impact of individual features on model outputs, with higher SHAP values (toward the right) signifying a positive contribution to the prediction, and lower SHAP values (toward the left) indicating a negative contribution. Figure 4a highlights the SHAP analysis identifying the combination of features that best polarize model predictions across the affective states.

Among the EOG features analyzed, the *Opening Phase Energy*, the integral of the opening phase of the peak signal, and *Opening Signal Range*, the amplitude of the opening phase of the peak signal, consistently appeared in optimal feature combinations across all three outputs, suggesting their robustness as predictors. Additionally, the *Signal Height* feature exhibited a particularly strong influence

on predictions for negative affectivity and state anxiety, underscoring its significance in these contexts. Among the EDA features analyzed, Hjorth parameters and the *Signal Standard Deviation* emerged as important predictors across the different affective states. These findings highlight the importance of analyzing feature interactions to reveal critical combinations that drive model performance, offering deeper insights into the physiological signals underpinning emotional and stress-related states.

The SHAP analyses in Figure 4b illustrate the importance of considering features in combination when identifying the most relevant biomarkers. By selecting sets of five features, we aim to identify a group of biomarkers that not only are individually relevant but also work effectively together. In Figure 4b, the inclusion of the feature *Opening Phase Energy* contributes significantly to the model's performance, yielding a well-defined distinction in SHAP values. When *Opening Phase Energy* is removed from the features considered, model performance decreases, and features such as *Blink Full-Close Duration* appear to show more distinction.

Conclusion

This study highlights the potential of electrooculography (EOG) and electrodermal activity (EDA) as powerful tools for identifying nuanced physiological biomarkers associated with state anxiety. Through the development and analysis of the BLINKEO and EMOCOLD datasets, we demonstrated the effectiveness of combining advanced feature extraction techniques with interpretability methods such as SHAP analysis to uncover anxiety-specific markers. Our results emphasize the importance of not just identifying individual biomarkers but also understanding their context-dependent interactions and collective contributions to predictive models.

By systematically evaluating combinations of features, we mitigated challenges often faced in the literature, where biomarkers show inconsistent or non-significant correlations with anxiety due to situational variability. For instance, while blink rate and skin conductance metrics have been previously explored, our analysis reveals that their predictive utility depends heavily on contextual factors, such as the type and intensity of the stressor. For example, biomarkers like blink duration and skin conductance peaks performed well under controlled Cold Pressor Test (CPT) conditions but may not generalize to other stress-inducing scenarios like public speaking. This underscores the need for adaptive, context-sensitive models that account for the situational variability of physiological responses.

A key contribution of this work is the identification of feature combinations that consistently provide reliable predictions. For EOG data, features like blink duration, peak height, and the opening integral

were shown to be robust predictors across various emotional states. Similarly, for EDA data, features such as the mean signal, permutation entropy, and Hjorth activity emerged as significant contributors. By leveraging SHAP analysis, we identified not only which features are most relevant but also when and how they interact to enhance model performance. This approach offers a more comprehensive understanding of physiological responses compared to studies focusing solely on single- feature analyses.

Our findings bridge a critical gap in the literature by offering a systematic approach to addressing the variability and context-dependence of physiological biomarkers. This research advances the field by providing a framework for building more robust, interpretable, and context-sensitive models for anxiety assessment. The ability to dynamically adapt to different stress scenarios makes these biomarkers more applicable to real-world settings, paving the way for more personalized and effective mental health interventions.

Future work should focus on validating these findings across diverse populations and stress-inducing contexts to further enhance the generalizability of these biomarkers. Additionally, integrating these models into wearable technology has the potential to revolutionize mental health monitoring, providing real-time, personalized insights that could transform how we understand and manage anxiety. By addressing the challenges of situational variability and leveraging the strengths of combined biomarker analyses, this study contributes significantly to the growing field of wearable health technology and its applications in mental health.

Statements and Declarations

Funding

This work was funded by the Translational Research Institute for Space Health through NASA NNX16AO69A, Office of Naval Research grant nos. N00014-21-1-2483 and N00014- 21-1-2845, Army Research Office grant no. W911NF-23-1-0041, National Institutes of Health grant nos. R01HL155815 and R21DK13266, National Science Foundation grant no. 2145802, National Academy of Medicine Catalyst Award and High Impact Pilot Research Award no. T31IP1666 from the Tobacco-Related Disease Research Program and Heritage Medical Research Institute (all to W.G.). T.K.H. acknowledges the support from National Institutes of Health grant nos. T32HL144449 and T32EB027629.

Human and Animal Research Statements

This work involved human subjects or animals in its research. The author(s) confirm(s) that all human/animal subject research procedures and protocols are exempt from review board approval.

Supplementary

Supplementary 1. Find Nearby Minimum Function

Function: findNearbyMinimum(x, p, w, m)

- **Inputs:**

- x : Signal data array
- p : Current index in x (initially the peak index)
- w : Binary search window size (sign determines search direction; positive for right, negative for left)
- m : Maximum points to search

- **Output:**

- p_{\min} : Index of the nearest local minimum within the search range

1. **Base Case:**

If $|w| < 1$ or $m = 0$:

- Set $s = x[\max(0, p - 1) : \min(p + 2, |x|)]$
- Find the relative minimum within s using $p_{\text{local}} = \arg \min(s)$
- Return $p_{\min} = p - p_{\text{local}} + \arg \min(s)$

2. **Initialize Search Variables:**

- $p_{\max} = p$
- $x_{\max} = x[p]$
- Set $d = \frac{w}{|w|}$ (search direction: +1 for right, -1 for left)

3. **Binary Search for Minimum:**

- **For** i from $\max(0, p)$ to $\min(p + d \times m, |x|)$ with steps of w :
 - **If** $x[i] \geq x_{\max}$ and $p \neq i$:
 - * Return findNearbyMinimum($x, i - w, \frac{w}{4}, m - d \times |i - w - p|$)
 - **Else:**
 - * Set $p_{\max} = i$
 - * Set $x_{\max} = x[i]$

4. **Reduce Binary Search Window:**

- Return findNearbyMinimum($x, p_{\max}, \frac{w}{2}, m - 1$)

Supplementary 2. EMOCOLD Participant Survey Questionnaire

The survey comprises 30 items assessing Positive Affectivity, Negative Affectivity, and State Anxiety. The first 10 items are rated on a 1-5 scale, while the remaining 20 items use a 1-4 scale. The table below details each item, its corresponding scale, and the affectivity or anxiety dimension it evaluates.

Survey Item	Scale	Dimension Evaluated
Upset	1-5	Negative Affectivity
Hostile	1-5	Negative Affectivity
Alert	1-5	Positive Affectivity
Ashamed	1-5	Negative Affectivity
Inspired	1-5	Positive Affectivity
Nervous	1-5	Negative Affectivity
Determined	1-5	Positive Affectivity
Attentive	1-5	Positive Affectivity
Active	1-5	Positive Affectivity
Afraid	1-5	Negative Affectivity
I feel calm	1-4	State Anxiety (Positive)
I feel secure	1-4	State Anxiety (Positive)
I am tense	1-4	State Anxiety (Negative)
I feel strained	1-4	State Anxiety (Negative)
I feel at ease	1-4	State Anxiety (Positive)
I feel upset	1-4	State Anxiety (Negative)
I am presently worrying over possible misfortunes	1-4	State Anxiety (Negative)
I feel satisfied	1-4	State Anxiety (Positive)
I feel frightened	1-4	State Anxiety (Negative)
I feel comfortable	1-4	State Anxiety (Positive)
I feel self-confident	1-4	State Anxiety (Positive)
I feel nervous	1-4	State Anxiety (Negative)
I am jittery	1-4	State Anxiety (Negative)
I feel indecisive	1-4	State Anxiety (Negative)
I am relaxed	1-4	State Anxiety (Positive)

I feel content	1-4	State Anxiety (Positive)
I am worried	1-4	State Anxiety (Negative)
I feel confused	1-4	State Anxiety (Negative)
I feel steady	1-4	State Anxiety (Positive)

Scale Interpretations:

- 1-5 Scale (Items 1-10):
 - 1: Very slightly or not at all
 - 2: A little
 - 3: Moderately
 - 4: Quite a bit
 - 5: Extremely
- 1-4 Scale (Items 11-30):
 - 1: Not at all
 - 2: Somewhat
 - 3: Moderately so
 - 4: Very much so

Supplementary 3. EDA Features Extracted

name	definition
Signal Mean	The average voltage within a specific window of the electrodermal activity (EDA) signal.
Signal Standard Deviation	The standard deviation of the EDA signal within a window, indicating variability in the signal.
Signal Range	The range of values (difference between maximum and minimum) within the EDA signal window.
Velocity Mean	The average value of the first derivative of the EDA signal within the window, capturing the average rate of change.
Velocity Standard Deviation	The standard deviation of the first derivative, measuring the variability in the rate of change of the EDA signal.
Petrosian Fractal Dimension	The Petrosian fractal dimension of the EDA signal, used to quantify the complexity of the signal's structure.
Higuchi Fractal Dimension	The Higuchi fractal dimension of the EDA signal, another measure of complexity, particularly suited for time series data.
DFA	Detrended Fluctuation Analysis of the EDA signal, which quantifies self-similarity and long-range correlations within the signal.
Katz Fractal Dimension	The Katz fractal dimension of the EDA signal, providing a measure of waveform complexity and irregularity.
Hjorth Activity (1st Hjorth Parameter)	A measure of the signal's power or variance, part of the Hjorth parameters used in time-domain analysis.
Hjorth Mobility (2nd Hjorth Parameter)	The square root of the variance of the first derivative of the signal divided by the variance of the signal itself, measuring the signal's frequency characteristics.
Hjorth Complexity (3rd Hjorth Parameter)	A measure derived from Hjorth parameters, capturing the signal's waveform complexity.
Variance of Rate of Change	The variance of the first derivative of the EDA signal, indicating variability in the signal's rate of change.

name	definition
Spectral Entropy	The entropy of the EDA signal's power spectrum, representing the disorder or unpredictability within the frequency domain.
Permutation Entropy	The permutation entropy of the EDA signal, a nonlinear measure of signal complexity that is sensitive to dynamic changes in the signal's structure.

Supplementary 4. EOG Features Extracted

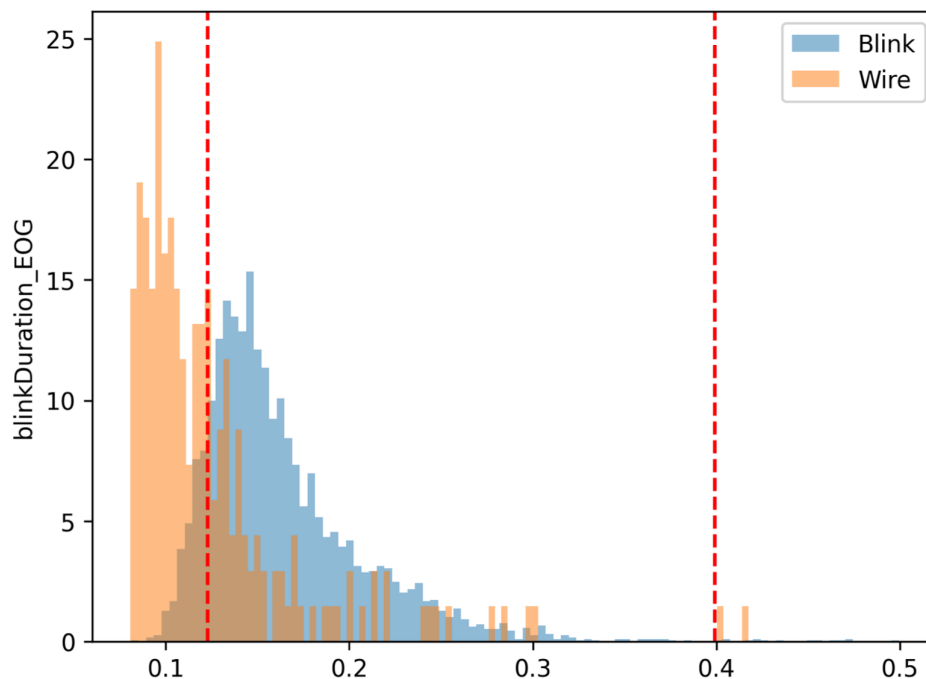
name	definition
Signal Height	The maximum amplitude of the EOG signal during an eye movement or blink peak.
X-Axis Deviation	Measures the horizontal difference between the x- coordinate of the signal's peak and the x-coordinate of the intersection of the upslope and downslope tangents, indicating asymmetry along the horizontal axis.
Y-Axis Deviation	Measures the vertical difference between the y- coordinate of the signal's peak and the y-coordinate of the intersection of the upslope and downslope tangents, reflecting differences in the steepness or curvature of the peak's slopes.
Symmetry Ratio	The ratio between tent deviations in X and Y axes, providing insight into blink symmetry or angle.
Closing Signal Range	The amplitude of the EOG signal during the full segment of the eye-closing phase. This is the range between the start of the peak rise and the point where the first derivative approaches zero after the closing segment.
Opening Signal Range	The amplitude of the EOG signal during the full segment of the eye-opening phase. This is the end of the closing segment to the baseline return, where the first derivative crosses zero after the peak.
Closing Duration	The interval between maximum velocity and amplitude during the blink, measuring blink speed. This is the second derivative peaks (indicating maximum velocity) before the first derivative returns to zero at the amplitude peak.
Closing Dynamics Ratio	The ratio of peak velocity to peak amplitude of the blink. This is the maximum velocity divided by Signal Height.
Blink Duration	The total duration of the blink event in the EOG signal.
Closing Tent Duration	The time taken to close the eye during the tenting phase of the blink. This is the time from the start of the upward slope until the first derivative peaks, indicating maximum closing speed.
Opening Tent Duration	This corresponds to the eye- opening duration during tenting. This is the time from the highest point of the blink (local maximum) to the return to baseline, indicated by a zero in the first derivative.

name	definition
Closing Tent Duration by Proportion of Blink	The fraction of the blink duration taken by the eye- closing phase. This is the closing time tent duration divided by the total blink duration.
Opening Tent Duration by Proportion of Blink	The fraction of the blink duration taken by the eye- opening phase. This is the opening time tent duration divided by the total blink duration.
Blink Half-Close Duration	The time interval during which the eye is half-closed within the blink event. This is when the signal amplitude reaches half of Signal Height during both the closing and opening phases.
Blink Full-Close Duration	This corresponds to the duration for which the eyes are fully closed, or near the peak. This is the time from the end of the closing phase (peak of the signal) until the start of the opening phase, using where the first derivative approaches zero.
Full-Close Duration by Percentage of Blink	This corresponds to the percentage of the blink time during which the eyes are fully closed. Divide Blink Full- Close Duration by Blink Duration and multiply by 100.
Opening Acceleration to Peak Duration	This corresponds to a condensed duration measure of the blink, specifically calculated from the start of significant acceleration, the left acceleration maximum, to the end of the blink peak.
Velocity Recovery Duration	This metric corresponds to the duration of the blink segment where the signal returns to the same amplitude level as the initial velocity peak. This is the time interval from the left-side velocity peak to the point after the blink peak where the signal reaches a similar amplitude as the initial velocity peak.
Closing Tent Duration	This corresponds to the time from blink start to the maximum peak velocity. This is from the start of the blink event to the point where the second derivative reaches its maximum.
Maximum Velocity to Peak Duration	This corresponds to the time between peak velocity and peak amplitude. This is the interval by measuring from the highest second derivative point (max velocity) to the first derivative zero crossing (peak amplitude).
Slope of Closing Tent	This corresponds to the slope of the closing phase up to maximum velocity. This is the change in amplitude over time from blink start to the point of maximum velocity.

name	definition
Slope of Opening Tent	This corresponds to the slope of the opening phase after minimum velocity. This is the minimum velocity point to baseline, tracking the decrease in amplitude over time.
Slope at Closing Tent, Maximum Acceleration	This corresponds to the slope of the closing phase at the maximum acceleration of the closing tent, at the beginning of the signal's peak phase.
Blink Phase Velocity Ratio	This corresponds to the ratio of velocities in closing and opening phases. This is calculated by dividing the closing velocity by the opening velocity.
Initial Blink Energy	This corresponds to the integral of the blink signal across the initial 5% of the blink, representing early energy.
Closing Phase Energy	This corresponds to the integral of the EOG signal over the eye-closing phase. Integrates the signal from the start of the blink to the peak.
Opening Phase Energy	This corresponds to the integral of the EOG signal over the eye-opening phase. Integrates from the peak to the return to baseline.
Closing Phase Slope Energy	This corresponds to the integral of the closing slope over time. Integrates the signal's slope up to the peak velocity during closing.
Closing Phase Velocity Energy	This corresponds to the integral of the EOG signal from the left-side velocity peak up to the blink amplitude peak. It captures the energy or accumulated signal from the point where the closing velocity is at its maximum to the blink's amplitude peak.
Opening Phase Velocity Energy	This corresponds to the integral of the EOG signal from the blink amplitude peak to the right-side velocity minimum. It represents the energy decay in the signal after the blink peak, spanning from the peak to where the opening velocity reaches its minimum.
Signal Average	This corresponds to the average value around the blink peak.
Acceleration Standard Deviation	This is the standard deviation of acceleration across the full blink.
Velocity Entropy	This corresponds to the entropy of blink velocity, indicating irregularities.
Acceleration Entropy	This corresponds to the entropy of blink acceleration, capturing the randomness in acceleration changes.
Maximum Acceleration Velocity	This is the ratio of maximum acceleration to maximum velocity in the opening phase of the peak.

name	definition
Ratio	

Supplementary 5. Blink Duration Feature Culling



References

1. [△]Steptoe A, Kivimäki M. "Stress and cardiovascular disease." *Nat. Rev. Cardiol.* 9 (6): 360–370, Jun. 2012. doi: 10.1038/nrcardio.2012.45.
2. [△]"Key Substance Use and Mental Health Indicators in the United States: Results from the 2022 National Survey on Drug Use and Health".
3. [△]"november-2023-topline-data.pdf." Accessed: Nov. 12, 2024. [Online]. Available: <https://www.apa.org/news/press/releases/stress/2023/november-2023-topline-data.pdf>
4. [△]"A review on flexible wearables – Recent developments in non-invasive continuous health monitoring - ScienceDirect." Accessed: Nov. 12, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0924424723008427>

5. [△]Freeman R, Chapleau MW. "Chapter 7 - Testing the autonomic nervous system." in *Handbook of Clinical Neurology*, vol. 115, Said G, Krarup C, Eds., in *Peripheral Nerve Disorders*, vol. 115., Elsevier, 2013, pp. 115–136. doi:10.1016/B978-0-444-52902-2.00007-2.
6. [△]Bullock T, et al. "Habituation of the stress response multiplex to repeated cold pressor exposure." *Front. Physiol.* 13, Jan. 2023. doi:10.3389/fphys.2022.752900.
7. [△]Bentivoglio AR, Bressman SB, Cassetta E, Carretta D, Tonali P, Albanese A. "Analysis of blink rate patterns in normal subjects." *Mov. Disord. Off. J. Mov. Disord. Soc.* 12 (6): 1028–1034, Nov. 1997. doi:10.1002/mds.870120629.
8. [△]Boucsein W. "Principles of Electrodermal Phenomena." in *Electrodermal Activity*, Boucsein W, Ed., Boston, MA: Springer US, 2012, pp. 1–86. doi:10.1007/978-1-4614-1126-0_1.
9. [△]"Recognition of Blinks Activity Patterns during Stress Conditions Using CNN and Markovian Analysis." Accessed: Nov. 12, 2024. [Online]. Available: <https://www.mdpi.com/2624-6120/2/1/6>
10. ^a, ^b"FULLTEXT01.pdf." Accessed: Nov. 12, 2024. [Online]. Available: <https://www.diva-portal.org/smash/get/diva2:673960/FULLTEXT01.pdf>Test
11. [△]Mannan MMN, Kamran MA, Kang S, Jeong MY. "Effect of EOG Signal Filtering on the Removal of Ocular Artifacts and EEG-Based Brain-Computer Interface: A Comprehensive Study." *Complexity* 2018 (1): 4853741, 2018. doi:10.1155/2018/4853741.
12. [△]Zhu X, Zheng WL, Lu BL, Chen X, Chen S, Wang C. "EOG-based drowsiness detection using convolutional neural networks." in *2014 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2014, pp. 128–134. doi:10.1109/IJCNN.2014.6889642.
13. [△]Ma JX, Shi LC, Lu BL. "An EOG-based Vigilance Estimation Method Applied for Driver Fatigue Detection." *Neurosci. Biomed. Eng.* 2 (1): 41–51, Jun. 2014.
14. [△]Harker M. "Psychological sweating: a systematic review focused on aetiology and cutaneous response." *Skin in Pharmacol. Physiol.* 26 (2): 92–100, 2013. doi:10.1159/000346930.
15. [△]Xu C, et al. "A physicochemical-sensing electronic skin for stress response monitoring." *Nat. Electron.* 7 (2): 168–179, Feb. 2024. doi:10.1038/s41928-023-01116-6.
16. [△]Sebastião R. "Classification of Anxiety Based on EDA and HR." in *IoT Technologies for HealthCare*, Goleva R, Garcia NRC, Pires IM, Eds., Cham: Springer International Publishing, 2021, pp. 112–123. doi:10.1007/978-3-030-69963-5_8.
17. [△]Strohmaier AR, Schiepe-Tiska A, Reiss KM. "A comparison of self-reports and electrodermal activity as indicators of mathematics state anxiety." *Frontline Learn. Res.* 8 (1): 16–32, Feb. 2020. doi:10.14786/flr.v8i1.427.

18. [△]Horvers A, Tombeng N, Bosse T, Lazonder AW, Molenaar I. "Detecting Emotions through Electrodermal Activity in Learning Contexts: A Systematic Review." *Sensors* 21 (23): 7869, Nov. 2021. doi:10.3390/s21237869.
19. [△]"IEEE Xplore Full-Text PDF:" Accessed: Nov. 17, 2024. [Online]. Available: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9508589>
20. [△]Sarchiapone M, et al. "The association between electrodermal activity (EDA), depression and suicidal behaviour: A systematic review and narrative synthesis." *BMC Psychiatry* 18 (1): 22, Jan. 2018. doi:10.1186/s12888-017-1551-4.
21. [△]Shatte ABR, Hutchinson DM, Teague SJ. "Machine learning in mental health: a scoping review of methods and applications." *Psychol. Med.* 49 (9): 1426–1448, Jul. 2019. doi:10.1017/S0033291719000151.
22. [△]Kleifges K, Bigdely-Shamlo N, Kerick SE, Robbins KA. "BLINKER: Automated Extraction of Ocular Indices from EEG Enabling Large-Scale Analysis." *Front. Neurosci.* 11, Feb. 2017. doi:10.3389/fnins.2017.00012.
23. [△]"Sensation and Perception: An Integrated Approach – Harvey Richard Schi4man – Google Books." Accessed: Nov. 12, 2024. [Online]. Available: https://books.google.com/books/about/Sensation_and_Perception.html?id=NKUQAQAIAAJ
24. [△]Reda R, Tantawi M, shedeed H, Tolba MF. "Analyzing Electrooculography (EOG) for Eye Movement Detection." in *The International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2019)*, Hassanien AE, Azar AT, Gaber T, Bhatnagar R, Tolba MF, Eds., Cham: Springer International Publishing, 2020, pp. 179–189. doi:10.1007/978-3-030-14118-9_18.
25. [△]López A, Ferrero F. "Biomedical Signal Processing and Artificial Intelligence in EOG Signals." in *Advances in Non-Invasive Biomedical Signal Sensing and Processing with Machine Learning*, Qaisar SM, Nisar H, Subasi A, Eds., Cham: Springer International Publishing, 2023, pp. 185–206. doi:10.1007/978-3-031-23239-8_8.
26. [△]Zheng WL, Lu BL. "A multimodal approach to estimating vigilance using EEG and forehead EOG." *J. Neural Eng.* 14 (2): 026017, Feb. 2017. doi:10.1088/1741-2552/aa5a98.
27. [△]"Classifying blinking and winking EOG signals using statistical analysis and LSTM algorithm | Journal of Electrical Systems and Information Technology | Full Text." Accessed: Nov. 19, 2024. [Online]. Available: <http://jesit.springeropen.com/articles/10.1186/s43067-023-00112-2>

Declarations

Funding: This work was funded by the Translational Research Institute for Space Health through NASA NNX16AO69A, Office of Naval Research grant nos. N00014-21-1-2483 and N00014-21-1-2845, Army

Research Office grant no. W911NF-23-1-0041, National Institutes of Health grant nos. R01HL155815 and R21DK13266, National Science Foundation grant no. 2145802, National Academy of Medicine Catalyst Award and High Impact Pilot Research Award no. T31IP1666 from the Tobacco-Related Disease Research Program and Heritage Medical Research Institute (all to W.G.). T.K.H. acknowledges the support from National Institutes of Health grant nos. T32HL144449 and T32EB027629.

Potential competing interests: No potential competing interests to declare.