

Article

Unsupervised Machine Learning for Multi-Phenotype Anxiety Discovery: A Large-Scale Longitudinal Study of Workplace Physiology

Nicu Ahmadi , Farzan Sasangohar , Ben Zoghi and Tracy Hammond 

Department of Engineering, Texas A&M University, College Station, TX 77843, USA; sasangohar@tamu.edu (F.S.); zoghi@tamu.edu (B.Z.); hammond@tamu.edu (T.H.)

* Correspondence: nikiahmadi@tamu.edu

Abstract

Anxiety disorders affect more than 301 million people worldwide and remain the most prevalent class of mental health conditions. Detection methods have traditionally relied on subjective self-reports, which fail to capture the dynamic physiological signatures of anxiety in daily life. In this longitudinal study, continuous wearable monitoring was conducted with 67 working professionals over nine months, yielding 44,443 h of multimodal data segmented into 890,494 five-minute windows. An unsupervised approach was applied, and ten distinct autonomic states were identified through spectral clustering without the use of emotion labels during clustering. Among these, a rare dysregulated state was discovered, characterized by extreme peripheral cooling ($d = -6.77$), tachycardia ($d = +4.41$), and heightened electrodermal activity, showing a 31-fold enrichment for self-reported anxiety ($p = 0.0022$). Beyond this extreme signature, anxiety was expressed heterogeneously across four physiological phenotypes ranging from acute dysregulation to mild perturbation. External validation with the WESAD dataset confirmed generalizability, while the absence of the rare state in laboratory stress paradigms highlighted its specificity to ecological anxiety. By integrating unsupervised discovery with post hoc self-report validation, this work advances our theoretical understanding of autonomic affect and provides a methodological foundation for scalable workplace mental health monitoring.

Keywords: anxiety detection; unsupervised learning; wearable sensors; physiological computing; affective state recognition; workplace stress; spectral clustering; electrodermal activity; heart rate variability



Academic Editor: Silvia Liberata Ullo

Received: 2 July 2026

Revised: 16 July 2026

Accepted: 20 July 2026

Published: 22 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Anxiety and stress in professional environments represent a significant challenge for workplace well-being and productivity, affecting millions of workers globally and contributing to burnout, reduced performance, and diminished quality of life. Beyond individual suffering, the economic costs of workplace anxiety and related disorders are substantial, with global estimates exceeding US\$1 trillion annually in lost productivity [1,2]. Addressing anxiety in occupational contexts requires reliable, scalable methods for detection and monitoring.

Current workplace assessments rely primarily on retrospective self-report surveys and questionnaires. While useful, these tools are constrained by recall bias, reporting burden, and limited temporal resolution. While Ecological Momentary Assessment (EMA) partially

mitigates recall bias through in-the-moment reports, it still requires active user input and cannot continuously monitor physiological dynamics.

Affective computing offers a complementary approach by developing computational systems that recognize, interpret, and respond to human emotions [3,4]. Physiological signals such as Heart Rate (HR), Heart Rate Variability (HRV), Electrodermal Activity (EDA), skin temperature, and metabolic indicators provide continuous and involuntary windows into autonomic nervous system activity. These multimodal biosignals support real-time affect recognition and address key limitations of self-report methods [5,6].

Despite their promise, most prior anxiety detection studies have relied on small, laboratory-based datasets collected under constrained conditions. Widely used benchmarks such as Wearable Stress and Affect Detection (WESAD) [7], the Database for Anxious States based on a Psychological Stimulation (DASPS) [8], and the Continuously Annotated Signals of Emotion (CASE) [9] predominantly involve students or healthy volunteers exposed to short-term stressors like the Trier Social Stress Test (TSST) [10]. While invaluable for methodological development, these datasets lack ecological validity and fail to capture the contextually embedded nature of anxiety in professional settings.

This study addresses these gaps by conducting the first large-scale, longitudinal analysis of Science, Technology, Engineering, and Mathematics (STEM) professionals using wearable physiological monitoring. Over a nine-month period, 44,443 h of multimodal recordings were collected from 67 participants. Processing yielded 890,494 five-minute windows, which were analyzed using unsupervised machine learning (ML) techniques without requiring emotion labels during clustering. Spectral clustering revealed ten distinct physiological states, including a rare dysregulated state with 31-fold anxiety enrichment, demonstrating that anxiety can emerge as a distinct physiological signature at scale. The focus on STEM professionals is particularly relevant, as this population faces unique stressors, including technical problem-solving pressures, continuous skill adaptation, and project deadline intensities.

The contributions of this study are threefold. First, the sensing and analysis dimension of affective computing is advanced through the introduction of an unsupervised framework capable of discovering autonomic phenotypes from large-scale wearable data. Second, the psychology and behavior domain is extended by characterizing how anxiety manifests physiologically in naturalistic workplace contexts, beyond the constraints of laboratory stress paradigms. Finally, the application and interaction track is informed by identifying pathways for affect-adaptive technologies in occupational health, human–computer interaction, and workplace well-being.

2. Background Studies

The study of anxiety through physiological signals has developed primarily within clinical and laboratory research. While working professionals—particularly in STEM fields—remain largely unexamined, the foundational understanding from these domains provides essential context for interpreting physiological patterns in occupational settings.

2.1. Anxiety as a Psychophysiological Construct

Anxiety manifests as persistent states of fear or worry, often accompanied by somatic symptoms such as elevated HR, chest tightness, headaches, and gastrointestinal discomfort [11–13]. In the psychological literature, a distinction is drawn between *state anxiety*—a transient affective reaction to immediate stressors—and *trait anxiety*, which reflects a stable predisposition toward heightened arousal across contexts [14,15]. Both constructs are clinically relevant, though state anxiety is more directly observable in dynamic physiological patterns. Although anxiety subtypes (e.g., generalized anxiety, social anxiety,

panic) are clinically meaningful, subtype boundaries are typically defined by cognitive, behavioral, and contextual criteria rather than physiology alone. In wearable-based affective computing, subtype-specific labeling is uncommon in ecological and longitudinal studies due to the absence of clinical diagnostic assessments and the sparsity and timing uncertainty of in-the-wild self-reports. Consistent with findings from recent systematic reviews, most physiological anxiety research therefore operationalizes anxiety as a dimensional, state-level construct rather than attempting diagnostic subtype differentiation [16]. Accordingly, the clusters identified in this work should be interpreted as autonomic phenotypes statistically enriched for anxiety reports, without assuming one-to-one correspondence to specific anxiety disorders.

Anxiety encompasses multiple subtypes, including Generalized Anxiety Disorder (GAD), phobias, panic disorder, and social anxiety disorder. Each subtype is characterized by distinct triggers but shares physiological correlates mediated by the autonomic nervous system. Chronic anxiety also increases vulnerability to comorbidities such as depression and cardiovascular disease [17,18]. While primarily studied in clinical populations, these psychophysiological mechanisms likely extend to subclinical workplace anxiety, though this remains empirically unexamined.

2.2. Physiological Biomarkers of Anxiety

Physiological signals provide a noninvasive window into the autonomic mechanisms underlying anxiety. EDA is among the most widely studied biomarkers, reflecting sympathetic modulation of sweat gland activity [3,4]. Increased skin conductance level and heightened frequency of skin conductance responses have been repeatedly associated with acute anxiety episodes. Cardiovascular indices also play a central role: elevated HR and reduced HRV, particularly in time-domain metrics, reliably accompany anxious arousal [19,20]. Peripheral temperature decreases due to sympathetic vasoconstriction, and respiratory patterns often shift toward rapid, shallow breathing [21,22]. Together, these markers capture the multidimensional activation characteristic of anxiety states.

2.3. Evolution of Detection Methods

Early approaches to anxiety recognition frequently relied on thresholding single biomarkers, such as EDA amplitudes or HR [23]. While these heuristics achieved moderate performance in constrained laboratory tasks, they lacked robustness in naturalistic environments where artifacts, environmental variability, and overlapping affective states complicate interpretation. To overcome these limitations, ML approaches capable of integrating multimodal signals were introduced.

Classical ML models such as support vector machines and random forests demonstrated improved detection performance by leveraging engineered features from EDA, HRV, temperature, and respiration [24–26]. More recent studies have employed deep learning architectures to capture spatiotemporal dependencies in raw signals [27–29]. These methods show promise but remain constrained by the need for dense, high-quality labels—a significant barrier in ecological contexts where anxiety episodes are rare and self-report through EMA is biased [5,6]. These approaches have also not been validated in working populations, where anxiety patterns may differ substantially from laboratory-induced stress.

2.4. The Ecological Validity Gap

A persistent limitation in the literature is the reliance on small, laboratory-based datasets. The WESAD dataset includes only 15 participants exposed to TSST [7], while DASPS involves 23 participants undergoing anxiety cue recitation [8]. Other corpora, such as CASE, also rely on scripted stressors and short collection windows [9]. Although

these datasets have catalyzed methodological progress, they capture induced stress rather than naturally occurring anxiety and predominantly involve student volunteers. Most physiological benchmarks labeled as “anxiety” rely on externally triggered, time-bounded stress paradigms (e.g., TSST), which primarily elicit acute arousal rather than persistent anxiety-related states. In contrast, workplace anxiety is often an internally driven, future-oriented state that can persist or recur without a single identifiable trigger, making it difficult to capture authentically under laboratory constraints. This motivates an ecological framing in which anxiety is treated as a naturally emerging occupational phenomenon and evaluated via its enrichment over long-term autonomic state structure rather than assumed from stress induction alone. This distinction aligns with recent synthesis work showing that stress–anxiety conflation and labeling shortcuts are a dominant threat to construct validity in physiological anxiety detection research [16]. Complementary evidence from dense experience sampling in naturalistic settings further shows that affective experience varies substantially across contexts and time, supporting an ecological framing in which anxiety is not assumed to be discrete or time-bounded [30].

As Alberdi et al. [31] emphasize, contextual variability and ecological noise pose substantial obstacles to reliable deployment. Setz et al. [23] likewise note that discriminating anxiety from confounding influences such as physical activity remains challenging. Critically, no published datasets include working professionals in their natural occupational environments, leaving a fundamental gap in understanding how anxiety manifests during actual work activities as opposed to artificial laboratory stressors.

2.5. Methodological Imperatives for Workplace Monitoring

The trajectory of affective computing research reflects a shift toward multimodal integration, longitudinal monitoring, and unsupervised modeling. By leveraging physiological signals that are largely involuntary, researchers aim to overcome the biases of self-report and expand applicability to real-world contexts. While supervised ML models dominate the literature, the challenges of labeling at scale underscore the importance of unsupervised discovery methods that can identify latent affective states without predefined annotations. These methodological requirements become even more critical when studying previously unexamined populations such as STEM professionals, where anxiety patterns and triggers may fundamentally differ from those observed in student volunteers.

3. Methods

3.1. Study Design and Participants

This longitudinal study monitored 67 working professionals (55 males and 12 females; average age of 38.8) over nine months (August 2024–May 2025). All participants held full-time STEM positions, were enrolled in Texas A&M University’s Master of Engineering Technical Management (METM) program, and had at least four years of professional technical experience. The METM is a primarily remote program, with participants spending only the first week together on campus before continuing coursework virtually while working full-time. Informed consent was obtained under IRB protocol #2023-0831D. Unlike conventional laboratory cohorts, this group reflected authentic workplace contexts, encountering real-world stressors such as project deadlines, technical problem solving, and performance accountability.

3.2. Data Collection and Processing

Participants continuously wore Empatica EmbracePlus smartwatches. Table 1 summarizes the signals collected. EMA was administered using Plutchik’s eight-emotion wheel [32], and its eight primary (first-order) emotions are joy, trust, fear, surprise, sadness,

disgust, anger, and anticipation. Anxiety is not among Plutchik’s primary emotions; it appears as a dyadic secondary emotion (a combination of fear and anticipation). As a structural consequence of the instrument, participants could select *anxious* only at the secondary level and never as a primary emotion. Emotion labels were used solely for post hoc validation and were not involved in clustering. The prevalence of fear alongside anxiety in the co-occurrence analyses reflects this dyadic structure of Plutchik’s wheel rather than an independent physiological finding.

Table 1. Physiological signals captured by Empatica EmbracePlus.

Signal	Modality	Sampling Rate
Electrodermal Activity (EDA)	Skin conductance	4 Hz
Blood Volume Pulse (BVP)	Photoplethysmography	64 Hz
Pulse Rate	Derived from BVP	1 Hz
Skin Temperature	Peripheral thermistor	1 Hz
Accelerometry	Tri-axial motion	64 Hz
Metabolic Equivalent (MET)	Derived activity index	1 Hz

The exclusive use of wrist-worn sensors imposed inherent constraints on signal acquisition. While electrocardiography would provide superior heart rate variability metrics and respiratory inductance plethysmography would enable direct breathing pattern assessment—both established anxiety biomarkers—such modalities require chest straps or adhesive electrodes incompatible with continuous nine-month workplace monitoring. The study design prioritized ecological validity and compliance over signal fidelity, accepting the limitations of wrist-based photoplethysmography for cardiac assessment. This trade-off reflects a fundamental challenge in affective computing: laboratory-grade sensors enable precise measurement but constrain real-world deployment, whereas consumer wearables sacrifice signal quality for scalability and continuous wear. Reports were provided as daily logs during an intensive residency week, weekly reflections throughout the nine-month program, and real-time web submissions. All reports were timestamped and aligned post hoc to physiological windows using a ± 3 -minute temporal tolerance, enabling validation of anxiety-related states without assuming continuous or retrospective labeling. All EMA entries were self-initiated: real-time reports could be submitted at any moment, while daily and weekly reflections could be completed at any time within the corresponding day or week.

Multi-emotion responses were decomposed into single labels, yielding 4474 entries from 2196 primary reports, to enable temporally precise alignment with physiological windows and to support state-level validation; this step was taken for analytical tractability rather than to impose a single-emotion assumption.

Raw data (123,961 AVRO files) were converted to structured CSVs, producing approximately 44,443 h of multimodal recordings. Signal-specific preprocessing was performed using standardized routines from the NeuroKit2 toolbox. For EDA, a four-stage pipeline was implemented, including sanitization, Butterworth low-pass filtering at 0.5 Hz, tonic-phasic decomposition, and peak detection using `scipy.signal.find_peaks` with physiologically informed constraints. The pulse rate (PR) was derived from BVP using Empatica’s onboard detection, with quality checks for artifacts. Skin temperature was normalized per participant to reduce baseline differences. Accelerometry and derived MET were used for artifact screening and activity estimation.

Signals were segmented into five-minute windows with one-minute stride (80% overlap). The five-minute duration was chosen to balance temporal resolution with physiological stability, ensuring sufficient samples for reliable PRV and EDA measures while maintaining sensitivity to short-term affective fluctuations. This window size has been

widely adopted in psychophysiological research for comparable reasons [33]. Within each window, mean and standard deviation values were computed for each modality, yielding a multimodal feature representation. Participant-wise z-normalization was applied to emphasize within-person dynamics. The complete preprocessing and feature aggregation pipeline is summarized in Figure 1.

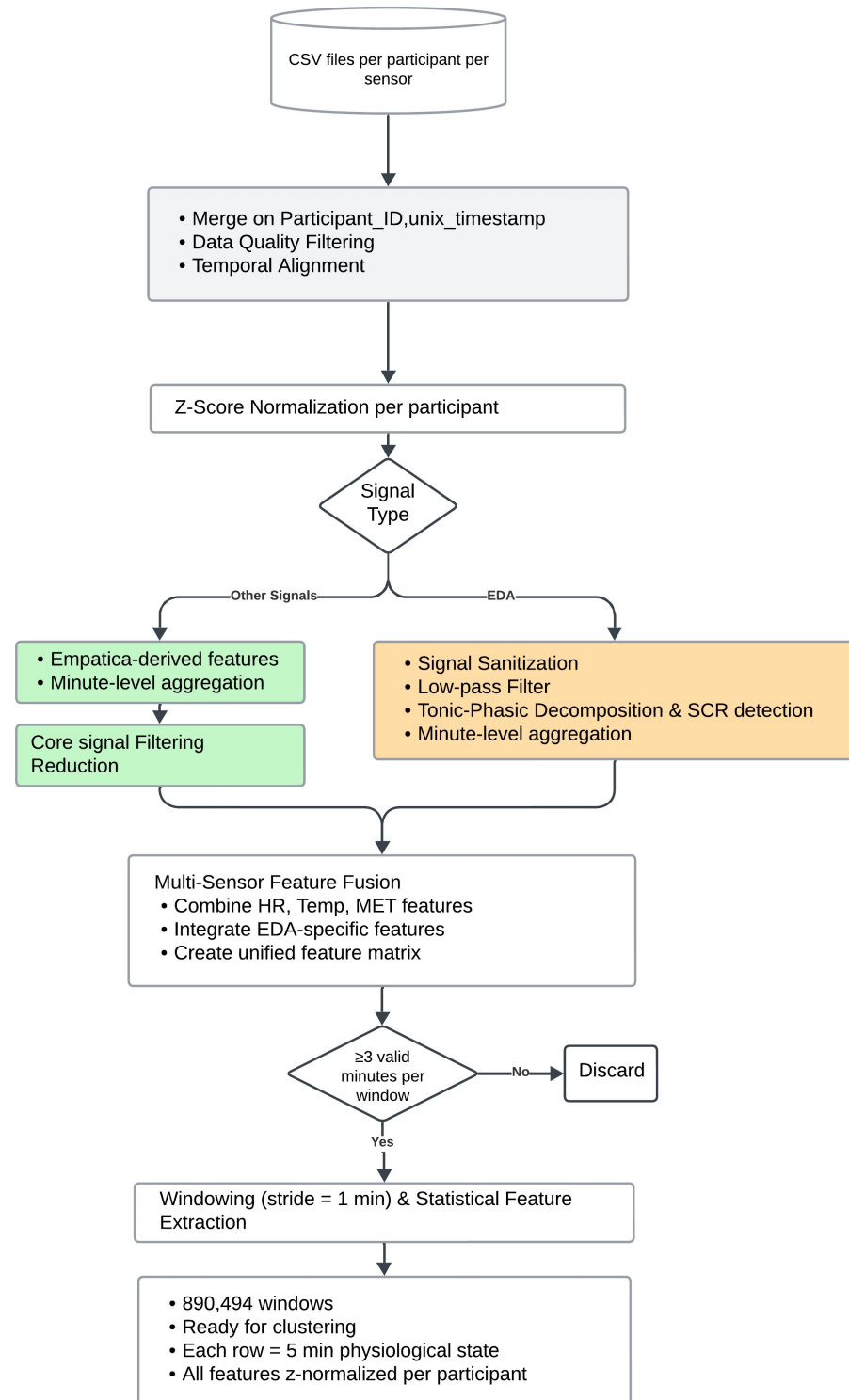


Figure 1. Preprocessing workflow for multimodal physiological signals, including raw-to-structured conversion, NeuroKit2-based filtering and decomposition.

3.3. Feature Selection and Clustering Approach

Six features were retained for clustering: mean PR, standard deviation of PR, mean phasic EDA, standard deviation of phasic EDA, mean skin temperature, and mean metabolic equivalent (MET). MET was intentionally retained as a clustering feature to explicitly represent physical activity and exercise-related physiology within the state space, rather than treating activity as a confounder to be removed. This design allows exercise-driven patterns to emerge as distinct physiological configurations if dominant. In addition, accelerometry-informed artifact screening was applied during preprocessing to suppress transient motion-related signal distortions, ensuring that feature statistics reflected physiological dynamics rather than transient motion artifacts.

The retained feature set was selected through systematic ablation rather than automated feature selection. An initial pool of 20 candidate features was derived from EDA, PR, temperature, and MET, including statistical moments (mean, standard deviation, skewness, and kurtosis) and simple temporal summaries computed within each five-minute window. Feature configurations of increasing dimensionality (4, 6, and 12 features) were evaluated with respect to cluster separability, stability of the resulting state structure, and preservation of rare physiological states rather than optimization of any single clustering metric. This process favored a parsimonious representation that balanced physiological interpretability with robustness under ecological noise and long-term wearable recording.

Although the full feature pool was retained for post hoc characterization, the six-feature configuration was selected for clustering based on the ablation results using internal clustering quality metrics only, prior to and independent of any post hoc emotion overlay. Silhouette analysis conducted in the two-dimensional uniform manifold approximation and projection (UMAP) embedding space indicated improved separability for the six-feature configuration ($s = 0.339$) compared to embeddings derived from four ($s = 0.215$) or twelve ($s = 0.284$) raw features. The three tested configurations and their internal validity scores are summarized in Table 2.

Table 2. Feature-configuration ablation. Silhouette scores computed in the two-dimensional UMAP embedding space for spectral clustering with $k = 10$. The six-feature configuration was selected based on internal clustering quality metrics, prior to any post hoc emotion overlay.

Configuration	Silhouette (s)	Features Included
4 features	0.215	Mean of PR, EDA phasic, temperature, MET
6 features (selected)	0.339	4-feature set + SD of PR and EDA phasic
12 features	0.284	Mean, SD, and skewness of all four signals

Clustering itself was performed in the original feature space, with UMAP used only for low-dimensional analysis and validation, as detailed in the validation framework.

Spectral clustering was employed to discover physiological states. A similarity graph was constructed as follows:

$$A_{ij} = \begin{cases} \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right), & \mathbf{x}_j \in \mathcal{N}_k(\mathbf{x}_i) \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

where $\mathcal{N}_k(\mathbf{x}_i)$ denotes the k nearest neighbors of window $\mathbf{x}_i$. The normalized Laplacian was defined as

$$L_{\text{norm}} = I - D^{-1/2} A D^{-1/2}, \quad (2)$$

with $D_{ii} = \sum_j A_{ij}$. Eigenvectors of L_{norm} were used for embedding, followed by K-means assignment.

Clustering was implemented using `scikit-learn` with `n_neighbors=50`, with kernel width σ scaled to the median distance, and the `arpak` solver (tolerance of 10^{-6} , maximum of 1000 iterations). Four values of k were tested ($\{5, 8, 10, 12\}$). The $k = 10$ solution was selected as it balanced interpretability with granularity. Validity indices including silhouette, Davies–Bouldin, and Calinski–Harabasz scores were computed.

3.4. Validation Framework

3.4.1. Comparative Clustering Analysis

Spectral clustering was chosen after comparison with K-means, Gaussian Mixture Models (GMMs), and Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN). The methods were assessed using silhouette coefficients, the Davies–Bouldin index, and post hoc anxiety enrichment. K-means produced favorable silhouette scores (0.302 at $k = 3$) but failed to isolate rare states, with the smallest cluster still containing 48,890 windows—two orders of magnitude larger than expected anxiety prevalence. HDBSCAN collapsed the manifold into a dominant baseline cluster (87% of windows) and misclassified noise, limiting its discriminative utility. GMM fragmented the manifold into many overlapping components, yielding modest enrichment ($2.3\times$) but reduced interpretability. Spectral clustering was unique in isolating rare physiological states. The smallest cluster contained 266 windows, closely aligned with the expected rarity of extreme physiological events. Rare-state preservation—a structural property of the algorithm evaluated independently of any emotion labels—was the primary basis for algorithm selection. Post hoc anxiety enrichment was subsequently used as an interpretive lens to characterize the discovered structure (31.2-fold enrichment in the smallest cluster, versus $2.3\times$ for GMM and $<1.0\times$ for K-means or HDBSCAN), but it did not drive the selection decision.

3.4.2. Statistical Validation

Cluster validity was evaluated using one-way ANOVA with Tukey’s honestly significant difference (HSD) test for feature-level discrimination, permutation testing (10,000 iterations) for enrichment significance, and bootstrap resampling (1000 iterations) for confidence intervals. Bonferroni correction ($\alpha = 0.05/54$) was applied to adjust for multiple comparisons. Clusters were examined in both the original six-dimensional feature space and two-dimensional UMAP embedding, with improved separation observed in the latter (silhouette: 0.339 vs. 0.108).

3.4.3. External Validation

External validation employed the WESAD dataset [7], which included 15 participants under TSST conditions. WESAD windows were preprocessed identically and mapped to cluster centroids in the six-dimensional feature space using Euclidean distances. This mapping enabled assessment of whether anxiety-derived phenotypes generalized to independent laboratory stress responses.

3.4.4. Post Hoc Emotion Alignment

EMA reports were temporally aligned to physiological windows with a ± 3 -minute tolerance, reflecting expected uncertainty in response timing. Alignment was performed using a window-centered merge-as-of approach, pairing each five-minute physiological window with the temporally closest EMA report from the same participant and accepting the match only if the EMA timestamp fell within ± 3 min of the window center and the window contained valid data across all modalities. Of the 890,494 physiological windows available after preprocessing, 873,573 originated from the 61 participants who contributed EMA reports and were therefore eligible for alignment. Within this subset, 1330 windows (0.15%) satisfied the strict alignment and completeness criteria. Among these, 212 contained anxi-

ety labels, establishing a baseline prevalence of 0.024%. Enrichment for each cluster C_k was computed as

$$\text{Enrichment}(C_k) = \frac{p_k}{p_{\text{baseline}}}, \quad (3)$$

where p_k is the prevalence of anxiety within C_k , and p_{baseline} is the global prevalence. This procedure enabled quantification of whether discovered physiological states corresponded to elevated anxiety incidence.

Reproducibility Notes

Fixed pseudorandom seeds were used throughout preprocessing, dimensionality reduction, and clustering to enable reproducibility. Windows requiring at least three valid minutes across all modalities were retained; incomplete windows were excluded. Bootstrap and permutation analyses used 1000 and 10,000 iterations, respectively. Cluster centroids in the standardized six-feature space are reported in Table 3. De-identified summary statistics and the full analysis pipeline (preprocessing, feature extraction, clustering, and validation) are available from the corresponding author upon reasonable request, subject to IRB approval, as noted in the Data Availability Statement.

Table 3. Six-feature physiological profiles of the ten clusters. Values are standardized deviations from the dataset mean.

Cluster	Size	Temp	PR	MET	EDA _{ϕ}	PR SD	EDA SD
0	102,887	−0.15	+1.34	+1.43	−0.11	+0.46	+0.00
1	59,907	+0.11	+1.13	+1.11	+2.92	+0.42	+0.03
2	149,925	+0.38	+0.22	−0.15	−0.32	+0.19	+0.00
3	94,369	−1.03	−0.38	−0.26	−0.28	+0.19	+0.00
4	73,751	+1.06	−0.69	−0.47	−0.36	+0.17	+0.00
5	145,622	−0.45	+0.22	+0.33	−0.29	+0.42	+0.00
6	266	−5.76	+3.47	+1.67	+1.08	+0.19	+0.01
7	97,394	+0.02	−0.59	−0.43	−0.36	+0.16	+0.00
8	73,818	−0.98	+0.25	+0.70	−0.22	+0.96	+0.00
9	92,555	+0.55	+0.68	+0.21	+0.33	+0.25	+0.01

4. Results

4.1. Discovered Physiological State Space

The unsupervised analysis of multimodal physiological recordings from 67 working professionals revealed a structured landscape of autonomic states during naturalistic workplace activities. A total of 44,443 h of data yielded 890,494 five-minute windows, each characterized by coordinated electrodermal, cardiovascular, thermal, and metabolic activity. Spectral clustering partitioned these windows into ten distinct physiological states with a markedly imbalanced distribution: Cluster 6 contained only 266 windows (0.03%), whereas Cluster 2 contained 147,421 windows (16.6%), with the remaining clusters spanning intermediate prevalences of 58,170–143,168 windows. This long-tailed distribution reflects the coexistence of common baseline states with rare physiological events.

Cluster quality was evaluated in the UMAP-embedded space using silhouette analysis. The overall coefficient was $s = 0.339$, indicating moderate separation of states along the physiological manifold. As shown in Figure 2, cohesion varied substantially across clusters: Cluster 6 achieved the highest silhouette score ($s = 0.86$) despite its rarity, reflecting both strong internal consistency and clear separation from neighboring states.

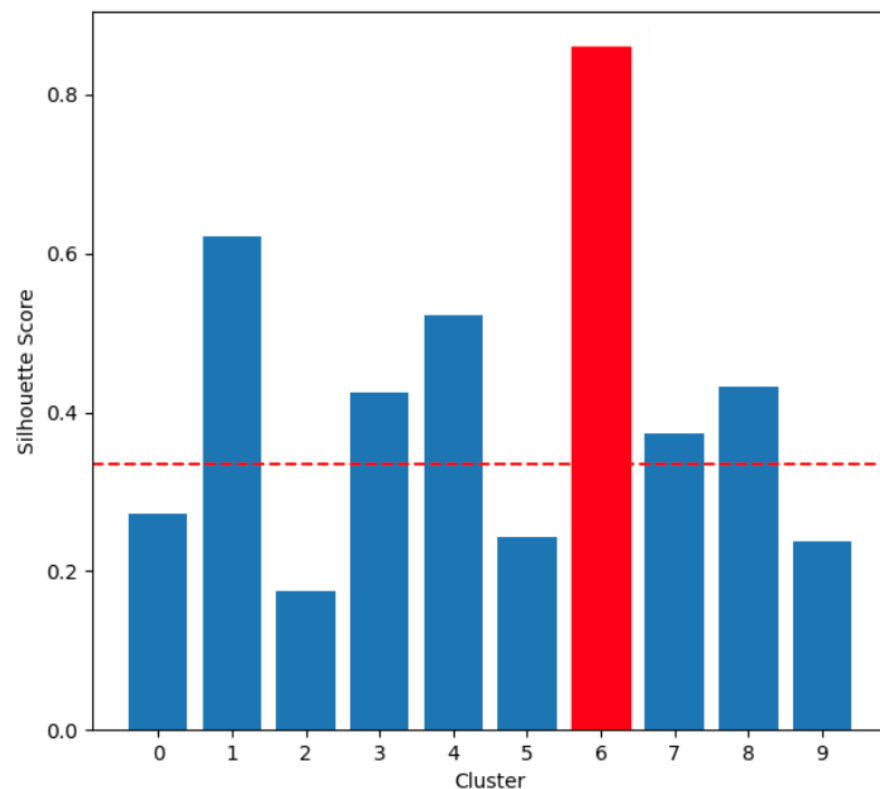


Figure 2. Silhouette scores by cluster in UMAP-embedded space. Blue bars indicate clusters other than Cluster 6, the red bar highlights Cluster 6, and the red dashed line denotes the overall mean silhouette score across clusters. Cluster 6 exhibited the strongest cohesion and separation ($s = 0.86$).

These results establish a ten-state physiological topology with highly imbalanced prevalence, providing the foundation for subsequent analyses of emotion alignment, physiological characterization, and rare-state identification.

4.2. Physiological Characterization of Clusters

The ten physiological states identified through spectral clustering exhibited distinct multi-system signatures, revealing a structured topology of autonomic nervous system activity during workplace functioning. Each cluster represented a unique configuration of temperature, cardiovascular, metabolic, and electrodermal dynamics, allowing classification into interpretable physiological state families.

Cluster centroids (Table 3) reveal systematic deviations across features, with states ranging from suppressed activation (e.g., C3, C7) to strong sympathetic engagement (e.g., C0, C1, C9).

4.2.1. Statistical Differentiation of Features

One-way ANOVA across all ten clusters confirmed significant between-group differences for each feature ($p < 0.001$). Given the very large number of windows analyzed ($n \approx 890,000$), even modest between-cluster differences yield large F-statistics; therefore we emphasize effect sizes (η^2) for interpretability. Effect sizes quantified by eta-squared ranged from $\eta^2 = 0.469$ (temperature mean) to $\eta^2 = 0.655$ (EDA phasic mean), indicating that cluster membership explained nearly half or more of the variance for every feature. Table 4 reports F-statistics and effect sizes.

Table 4. Feature discriminative power across clusters, ranked by effect size.

Feature	F-Statistic	η^2	Max–Min Clusters
EDA phasic mean	187,722	0.655	C1 (max)–C7 (min)
MET mean	142,090	0.590	C6 (max)–C4 (min)
EDA phasic std	138,780	0.584	C1 (max)–C4 (min)
PR std	122,871	0.554	C8 (max)–C7 (min)
PR mean	89,643	0.475	C6 (max)–C4 (min)
Temperature mean	87,401	0.469	C4 (max)–C6 (min)

Post hoc Tukey HSD tests confirmed that 98% of cluster pairs differed significantly ($p < 0.001$), indicating that the discovered states are physiologically distinct. Variability features (PR std, EDA phasic std) proved to be discriminative as mean-level features, underscoring the importance of dynamic fluctuations within five-minute windows.

4.2.2. Cluster Families

Based on centroid profiles (Table 3), the ten clusters were organized into four higher-order families:

- **Baseline/Resting (C2, C3, C4, and C7):** Near-baseline or suppressed activation, with temperature ranging from cold (C3) to warm (C4).
- **Moderate/Transitional (C5 and C8):** Elevated metabolism and variability despite peripheral cooling; C8 exhibited the highest PR variability.
- **High Activation (C0, C1, and C9):** Strong sympathetic engagement. C1 combined cardio-metabolic activation with high non-extreme EDA.
- **Acute Dysregulation (C6):** An isolated state with multi-system deviation; detailed analysis follows in Section 4.3.

4.3. Discovery and Validation of a Rare Anxiety Signature (Cluster 6)

Cluster 6 emerged as the rarest yet most distinctive physiological state, comprising only 266 windows. Despite its rarity, this cluster exhibited a uniquely dysregulated profile characterized by severe multi-system deviations.

4.3.1. Physiological Profile and Magnitude of Effects

Figure 3 shows how Cluster 6 diverged from all other states across temperature, PR, MET, and EDA. Effect sizes (Table 5) revealed extreme deviations: peripheral cooling, tachycardia, elevated metabolism, and heightened sudomotor activity. To contextualize, effect sizes greater than 0.8 are conventionally considered large in physiological research; Cluster 6's temperature deviation ($d = -6.77$) represents an unprecedented magnitude. This constellation of effects reflects a coherent autonomic pattern rather than random noise, combining sympathetic activation (tachycardia, elevated metabolism, and heightened EDA) with peripheral vasoconstriction (temperature suppression), consistent with acute dysregulation observed in high-anxiety or fear states.

These magnitudes far exceed conventional thresholds for large effects ($|d| > 0.8$) and surpass values typically reported in laboratory stress studies, reinforcing that Cluster 6 captures an extreme but physiologically coherent dysregulated state unlikely to arise from artifacts or random variance. Euclidean centroid distances confirmed isolation: The nearest neighbor was Cluster 8 at $d_{6,8} = 6.03$, whereas typical inter-cluster distances were < 2.0 . Silhouette analysis corroborated this distinctiveness, with Cluster 6 achieving $s = 0.86$ compared to the global average of $s = 0.339$.

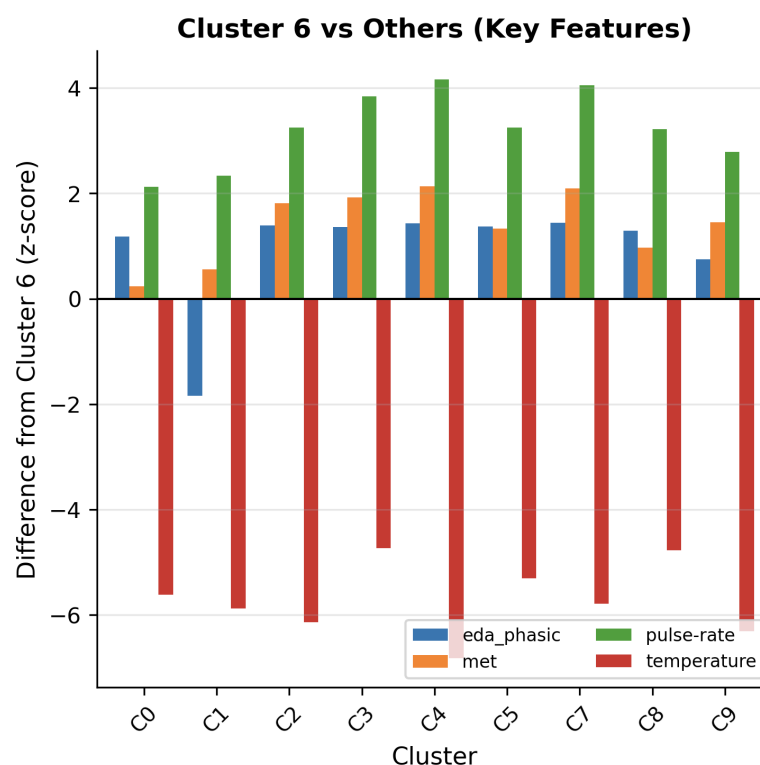


Figure 3. Cluster 6 deviations relative to other clusters across temperature, PR, MET, and EDA.

Table 5. Cluster 6 physiological signature: Cohen’s *d* effect sizes and 95% confidence intervals.

Feature	Cohen’s <i>d</i>	95% CI
Temperature mean	−6.77	[−7.08, −6.48]
PR mean	+4.41	[+4.23, +4.61]
MET mean	+2.14	[+2.03, +2.25]
EDA phasic mean	+0.89	[+0.77, +1.01]
PR SD	−0.68	[−0.76, −0.59]
EDA SD	+0.84	[+0.62, +0.99]

4.3.2. Statistical Validation and Robustness

The observed anxiety prevalence in Cluster 6 ($\psi_6 = 0.752\%$) corresponded to a ~ 31 -fold enrichment relative to the dataset baseline of 0.024% (Table 6). To evaluate significance, a permutation test with 10,000 iterations repeatedly sampled $n = 266$ windows; the observed prevalence exceeded 99.78% of permuted values, yielding $p = 0.0022$.

Table 6. Anxiety prevalence and enrichment across physiological clusters for the EMA-eligible subset of windows. Emotion coverage (%) denotes the proportion of EMA-aligned windows assigned to each cluster.

Cluster	Total Windows	Anxiety Rate (%)	Conditional Anxiety (%)	Emotion Coverage (%)	Enrichment Fold
0	101,604	0.02	5.17	0.29	0.61
1	58,170	0.04	9.81	0.37	1.49
2	147,421	0.02	7.93	0.22	0.73
3	92,752	0.03	12.81	0.22	1.16
4	72,191	0.01	9.38	0.09	0.34
5	143,168	0.03	9.41	0.28	1.09
6	266	0.75	28.57	2.63	30.98
7	95,380	0.02	9.66	0.15	0.61
8	72,929	0.03	8.27	0.38	1.30
9	89,692	0.05	15.59	0.29	1.88

Pairwise contrasts between Cluster 6 and the nine other clusters across six physiological features (54 tests) controlled familywise errors via Bonferroni correction:

$$\alpha_{\text{corrected}} = \frac{0.05}{54} = 0.000926. \quad (4)$$

Differences in temperature, PR mean, MET, and EDA measures remained highly significant under this threshold; PR variability was less uniquely characteristic.

Bootstrap resampling (1000 iterations) produced narrow 95% confidence intervals for key effects (Table 5), confirming stability despite the limited sample. Across bootstrap subsampling proportions ($p \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$), Cluster 6 was rediscovered in 100% of runs, with defining features showing exceptional stability (temperature $z \approx -5.76$, PR $z \approx +3.47$, coefficients of variation $< 1\%$). Perturbation tests adding Gaussian noise up to 20% of each feature's standard deviation preserved the extreme profile in all trials, confirming robustness to sampling variability and measurement noise (Table 7). Figure 4 illustrates consistent multi-system deviations across resamples.

Table 7. Cluster 6 stability across bootstrap subsampling (100 runs per p).

p	Mean Windows	Temp (z) Mean $\pm$ SD	PR (z) Mean $\pm$ SD
0.5	133.4 $\pm$ 7.6	−5.767 $\pm$ 0.045	3.464 $\pm$ 0.028
0.6	159.4 $\pm$ 8.1	−5.764 $\pm$ 0.040	3.469 $\pm$ 0.023
0.7	185.9 $\pm$ 5.9	−5.764 $\pm$ 0.034	3.466 $\pm$ 0.019
0.8	212.8 $\pm$ 6.6	−5.759 $\pm$ 0.022	3.466 $\pm$ 0.015
0.9	239.2 $\pm$ 5.3	−5.765 $\pm$ 0.017	3.466 $\pm$ 0.009

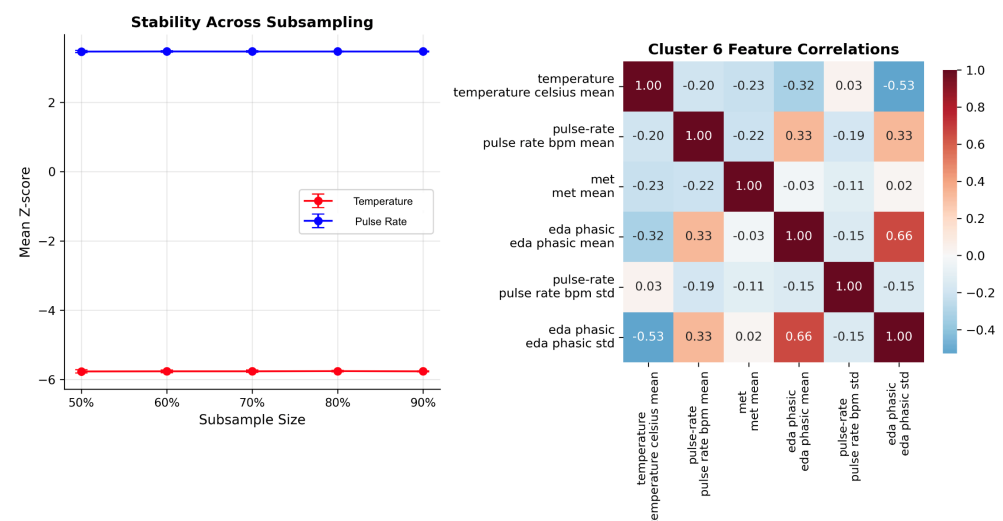


Figure 4. Cluster 6 robustness across subsampling and feature correlations. Multi-system deviations remained stable across all resampling trials.

Within the six-dimensional physiological manifold, Cluster 6 maintained the highest internal validity (silhouette = 0.86) and was geometrically isolated from all other centroids (minimum distance $d_{6,8} = 6.03$). These indices confirm that its distinctiveness is structural rather than an artifact of clustering initialization or projection space.

Interpretation and Generalizability: Although the convergence of permutation significance ($p = 0.0022$), bootstrap confidence intervals, subsampling stability, and perturbation robustness supports Cluster 6 as a genuine physiological signature, the small absolute count implies that prevalence estimates should be interpreted with appropriate uncertainty. The rarity of this state is consistent with its interpretation as an extreme anxiety phenotype

in a working professional population; however, replication in independent occupational cohorts will be necessary to confirm generalizability.

Summary: The convergence of extreme effect sizes, centroid isolation, high silhouette, permutation inference ($p = 0.0022$), multiplicity-controlled contrasts ($\alpha_{\text{corrected}}$), bootstrap precision, and robustness to perturbation confirms Cluster 6 as a genuine, rare anxiety-related physiological signature in naturalistic contexts.

4.4. Multi-Phenotype Anxiety Model

Anxiety manifestation was analyzed across all ten physiological states, yielding a multi-phenotype model that integrates prevalence, physiological intensity, and emotional context.

4.4.1. Distribution of Anxiety Across Clusters

The overall baseline prevalence of anxiety across all windows was 0.024%. Clusters displayed strong heterogeneity. Cluster 6, despite representing only 266 windows, exhibited the highest prevalence ($\psi_6 = 0.752\%$), corresponding to a 31-fold enrichment relative to the baseline ($p = 0.0022$). Cluster 9 also exceeded baseline prevalence (1.9-fold enrichment), driven by its larger size (89,692 windows). The full distribution is shown in Figure 5.

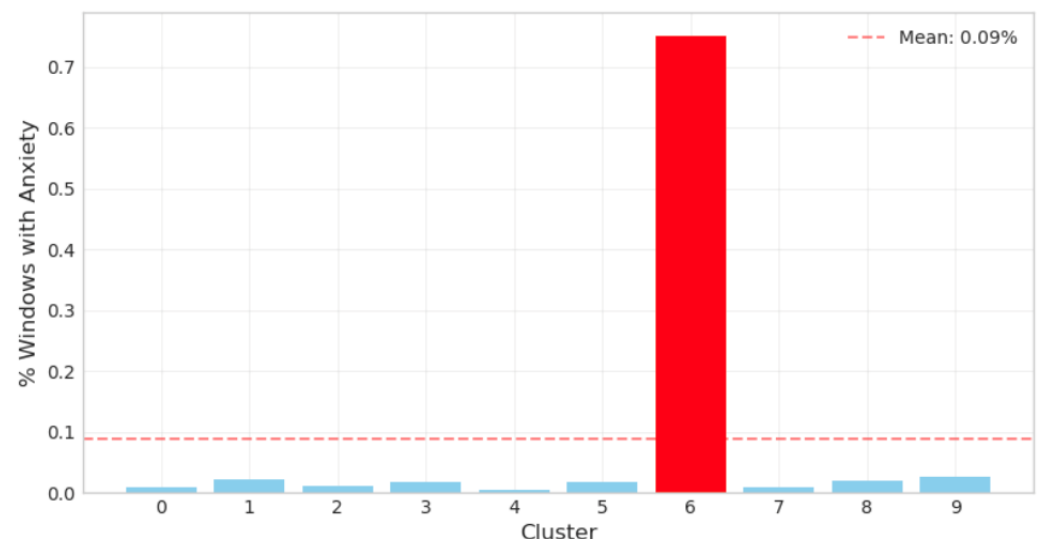


Figure 5. Cluster-level anxiety prevalence across the ten discovered physiological states. Light-blue bars indicate clusters other than Cluster 6, the red bar highlights Cluster 6, and the red dashed line denotes the mean cluster-level anxiety prevalence across clusters. Cluster 6 showed the strongest enrichment.

4.4.2. Phenotype Classification of Anxiety Intensity

A heuristic physiological intensity score, weighted by effect sizes from Cluster 6, ranked clusters by their anxiety-linked activation. The score for cluster k is defined as

$$I_k^{\text{raw}} = w_T \bar{z}_{k,\text{Temp}} + w_{HR} \bar{z}_{k,\text{PR}} + w_{\text{MET}} \bar{z}_{k,\text{MET}} + w_{\text{EDA}} \bar{z}_{k,\text{EDA}}, \quad (5)$$

$$I_k = 100 \cdot \frac{I_k^{\text{raw}} - \min_j I_j^{\text{raw}}}{\max_j I_j^{\text{raw}} - \min_j I_j^{\text{raw}}}, \quad (6)$$

where $\bar{z}_{k,f}$ denotes the mean z-score of feature f within cluster k , and the weights, w_f , are set proportional to the absolute Cohen's d effect sizes of Cluster 6 (Table 5). Cluster 6 was chosen as the reference anchor because it exhibited the largest multi-system effect sizes, the strongest geometric isolation, and the highest silhouette score of any discovered state.

Because Cluster 6's own weights parameterize the score, its classification as the *Extreme* phenotype is a construction of the scoring anchor rather than an independent result; however, the ordering among the remaining nine clusters is not fixed by this construction, and the resulting high > moderate > mild gradient is independently validated by WESAD STAI-Y6 scores (Section 4.6). Four phenotypes emerged:

- **Extreme:** Cluster 6, defined by severe peripheral cooling with concurrent cardiovascular, metabolic, and electrodermal activation.
- **High:** Clusters 1 and 8, characterized by elevated PR and metabolism, with heightened EDA in C1.
- **Moderate:** Clusters 0, 3, and 5, showing intermediate sympathetic engagement.
- **Mild:** Clusters 2, 4, 7, and 9, generally near baseline physiology with limited deviations.

This stratification partially aligns with the state families described in Section 4.2.2, but it also reveals anxiety-specific differences.

4.4.3. Conditional Rates and Emotion Co-Occurrence

Conditional anxiety rates quantify the likelihood of anxiety relative to other emotions within each cluster. Cluster 6 again stood out, with $\phi_6 = 28.6\%$, exceeding the 5–15% range observed elsewhere. To contextualize these findings, clusters were ranked by conditional anxiety rates and their dominant co-occurring emotions summarized (Table 8). Fear dominated across clusters (85–100%), with anger and sadness emerging in warm/active states (e.g., C9) and surprise occasionally appearing in cold–rest states (e.g., C3). Occasional “happy” reports during anxiety (e.g., C0, C2, C7, C8) likely reflect mixed affect or contextual influences.

Table 8. Clusters ranked by conditional anxiety rate with top co-occurring emotions.

Cluster	% Data	Cond. Anxiety (%)	Top Anxiety-Linked Emotions
6	0.03	28.57	Fear (100%)
9	10.27	15.59	Fear (73%), Anger (10%)
3	10.62	12.81	Fear (88.5%), Surprise (7.7%)
1	6.66	9.81	Fear (95.2%), Sadness (4.8%)
7	10.92	9.66	Fear (92.9%), Happy (7.1%)
5	16.39	9.41	Fear (86.8%), Anger (10.5%)
4	8.26	9.38	Fear (100%)
8	8.35	8.27	Fear (91.3%), Happy (8.7%)
2	16.88	7.93	Fear (84.6%), Happy (11.5%)
0	11.63	5.17	Fear (86.7%), Happy (13.3%)

Translational Note. These physiological patterns could inform workplace wellness monitoring systems and affect-aware feedback tools; they are not intended for clinical diagnosis.

4.5. Temporal Dynamics of Anxiety States

The temporal dynamics of physiological states were examined to determine whether anxiety occurred during sustained or transient episodes. *Episodes* were defined as contiguous runs of identical cluster assignments; any interruption longer than six minutes initiated a new episode. Episode counts and median durations are reported in Table 9.

To quantify the relationship between anxiety and persistence, the following ratio was computed for each cluster:

$$R_k = \frac{\overline{D}_k^{\text{anx}}}{\overline{D}_k^{\text{all}}}, \quad (7)$$

where $\overline{D}_k^{\text{anx}}$ denotes the mean duration of episodes containing at least one anxiety annotation, and $\overline{D}_k^{\text{all}}$ denotes the mean duration of all episodes in cluster k .

Table 9. Episode statistics by cluster (median duration only).

Cluster	Episodes (N)	Median Duration (min)
0	22,687	15.0
1	5852	25.0
2	30,502	15.0
3	16,017	15.0
4	12,823	25.0
5	45,394	10.0
6	12	22.5
7	17,431	20.0
8	21,202	15.0
9	18,735	15.0

Across most clusters, episodes containing anxiety were found to be longer than typical episodes (e.g., C8: $R = 2.43$; C9: $R = 3.24$). This indicates that anxiety was generally embedded within sustained physiological states rather than arising during brief fluctuations.

An exception was observed in Cluster 6. Although C6 episodes had a median duration comparable to other clusters (22.5 min; Table 9), anxiety-containing C6 episodes were markedly shorter than the cluster norm ($R = 0.14$). Only 12 C6 episodes were detected in total (derived from 266 windows across all participants), underscoring the rarity of this dysregulated state. These findings suggest that, although C6 persisted physiologically once entered, the co-occurrence of anxiety within it was transient. The results therefore highlight an asymmetry in temporal expression: anxiety tended to persist during high-activation states (C8 and C9) but was brief within the prolonged dysregulated state of C6.

4.6. External Validation with WESAD

To evaluate the generalizability of the discovered physiological states, external validation was performed using the WESAD dataset. WESAD also incorporates task-level state anxiety assessments (STAI-Y6) collected before and after each condition. STAI-Y6, a validated six-item short form of the State-Trait Anxiety Inventory, captures transient anxiety (“how one feels right now”) with high reliability and sensitivity to task-induced changes. This makes WESAD suitable for assessing whether the anxiety-derived physiological phenotypes identified in professionals generalize to controlled laboratory contexts, thereby strengthening the ecological validity of the findings.

4.6.1. Mapping Methodology

Each WESAD window was preprocessed using the same pipeline as the professional dataset (low-pass EDA, tonic–phasic decomposition, five-minute windows with 80% overlap). The same six clustering features were derived. Windows with less than 80% valid label coverage were excluded, yielding 533 analyzable windows. Each window was assigned to the nearest professional-cluster centroid using Euclidean distances in the six-dimensional feature space:

$$\hat{c}(x) = \arg \min_{k \in \{0, \dots, 9\}} \|x - \mu_k\|_2, \quad (8)$$

where μ_k denotes the centroid of cluster k .

4.6.2. Anxiety Distribution Across Clusters

A total of 533 WESAD windows were assigned to the professional-cohort clusters. The mean STAI-Y6 varied systematically across both the anxiety phenotypes and the broader physiological families.

Phenotype-Level Validation: Across WESAD, mean STAI-Y6 values followed the expected gradient: high > moderate > mild. This pattern mirrors the structure observed in

the professional cohort, confirming that the unsupervised clustering captured physiological patterns corresponding to anxiety intensity levels. No WESAD windows were assigned to Cluster 6, suggesting that this *Extreme* phenotype represents a distinct pattern present only in the workplace dataset and not elicited under controlled laboratory conditions. Figure 6 illustrates this graded relationship between phenotype and the mean STAI-Y6.

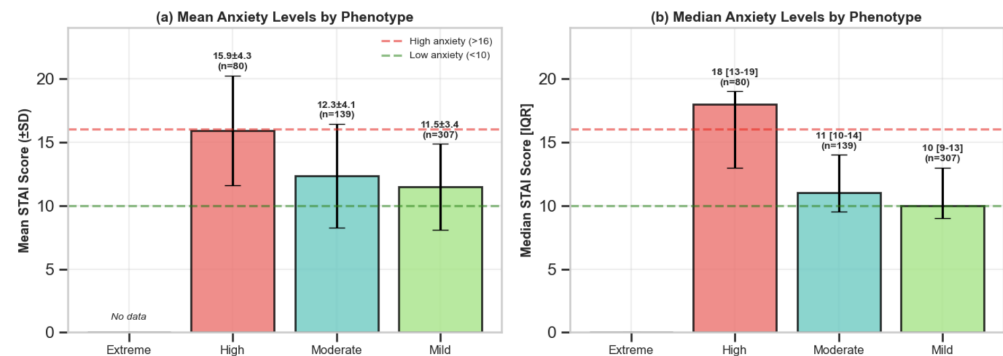


Figure 6. Mean STAI-Y6 across WESAD phenotype assignments. Bracketed values in the figure indicate STAI-Y6 value ranges, not reference citations. Extreme unoccupied.

Family-Level Context: At the supporting family level, the same hierarchy was evident: High Activation > Moderate/Transitional > Baseline/Resting in the mean STAI-Y6 scores. Table 10 summarizes the distribution of windows and mean anxiety values by cluster and family. Anxiety was highest in the *High-Activation* clusters (C1: 18.32 ± 2.54 ; C0: 16.35 ± 4.83 ; C9: 15.56 ± 3.94), intermediate in *Moderate/Transitional* clusters (C5: 12.15 ± 3.10 ; C8: 11.19 ± 3.06), and lowest in *Baseline/Resting* clusters (mean ≈ 10 – 11). No WESAD windows were assigned to Cluster 6.

Table 10. WESAD assignment and STAI-Y6 by professional-cohort clusters, grouped by physiological state families.

Family	Cluster	Windows (N)	STAI-Y6 Mean $\pm$ SD
High Activation	C1	53	18.32 ± 2.54
	C0	26	16.35 ± 4.83
	C9	59	15.56 ± 3.94
Moderate/Transitional	C5	39	12.15 ± 3.10
	C8	27	11.19 ± 3.06
Baseline/Resting	C2	69	11.09 ± 2.84
	C3	74	11.03 ± 3.34
	C4	116	10.23 ± 1.92
	C7	63	10.29 ± 2.47
Acute Dysregulation	C6	0	—
Total		533	

4.6.3. Absence of Cluster 6 in Laboratory Data

No WESAD windows were assigned to Cluster 6, the acute dysregulation state. This observation should be interpreted alongside its base rate: Given Cluster 6's prevalence of $266/890,494 = 0.030\%$ in the professional cohort, the expected number of assignments in a WESAD sample of 533 windows is $533 \times 0.000299 \approx 0.16$, corresponding to a probability of observing zero assignments of approximately 0.85 even under the null hypothesis that WESAD's underlying rate matches the professional cohort. The absence therefore does not, on its own, constitute strong statistical evidence of laboratory insensitivity. That said, the qualitative Cluster 6 signature—peripheral cooling of $-5.76z$ concurrent with tachycardia

of $+3.47z$ (Table 3), a combination inconsistent with exercise-driven activation and atypical of TSST-elicited acute stress—suggests that Cluster 6 may represent an ecological anxiety phenotype less readily reproduced by scripted stressors, though larger external validation cohorts would be required to establish this with statistical confidence.

4.6.4. Implications for Ecological Validity

The external validation yields three key implications. First, when interpreted alongside the base-rate constraint discussed in Section 4.6.3, the absence of Cluster 6 assignments in WESAD is consistent with the *Extreme* phenotype—defined by severe peripheral cooling and tachycardia—being less readily elicited under controlled laboratory conditions. This observation is consistent with the possibility that standard stress paradigms may not routinely elicit high-intensity autonomic dysregulation of this magnitude, though additional cohorts would be needed to confirm this. Second, the ordered gradient of STAI-Y6 across remaining phenotypes (High > Moderate > Mild) confirms that the unsupervised clustering captured meaningful physiological differences aligned with anxiety intensity. Third, the consistent hierarchy among state families—with High-Activation states showing higher STAI-Y6 than Baseline/Resting—demonstrates that the autonomic structure discovered in professionals generalizes across datasets, linking laboratory stress to the lower end of the real-world anxiety spectrum.

5. Discussion

5.1. Theoretical Implications

The findings demonstrate that anxiety can manifest as distinctive physiological patterns identifiable without extensive labeling. The identification of Cluster 6—comprising only 266 windows yet showing 31-fold anxiety enrichment ($p = 0.0022$)—indicates that extreme anxiety episodes create rare but distinctive autonomic signatures detectable through unsupervised analysis. Specifically, the concurrent presentation of severe peripheral cooling ($d = -6.77$), tachycardia ($d = +4.41$), elevated metabolism ($d = +2.14$), and heightened EDA ($d = +0.89$) represents a coordinated multi-system response not observed in any other state. This signature's isolation in feature space (minimum distance > 6.0 to any other cluster) and exceptional silhouette score (0.86 vs. global 0.339) confirm that it reflects a qualitatively distinct physiological phenomenon rather than an extension of normal arousal.

Environmental factors such as ambient office temperature or air conditioning may contribute to peripheral cooling; however, Cluster 6 is unlikely to reflect baseline temperature differences, as all features were z-scored within participants and cooling co-occurs with heightened cardiovascular, metabolic, and electrodermal activation.

The multi-phenotype model, with four distinct anxiety expressions ranging from extreme dysregulation to mild perturbation, aligns with dimensional theories of affect while providing empirical evidence for heterogeneous anxiety manifestation. The observation that fear dominated as the co-occurring emotion across all phenotypes (85–100% of anxiety windows), yet conditional anxiety prevalence varied dramatically (5.17–28.57%), suggests that physiological context modulates emotional expression beyond simple arousal levels.

The complete absence of Cluster 6 in WESAD's 533 windows—despite that dataset containing substantial stress responses under TSST—is consistent with a distinction between laboratory-induced stress and ecological anxiety, though the base-rate constraint discussed in Section 4.6.3 tempers strong inference from a single external cohort. Whereas WESAD stress concentrated in High-Activation clusters (C0, C1, and C9) with up to 96.3% stress prevalence, no windows exhibited the extreme multi-system dysregulation defining Cluster 6. This gap between controlled and naturalistic data suggests that certain anxiety

signatures may emerge more readily under real-world conditions than under scripted laboratory protocols.

5.2. Methodological Contributions

The unsupervised framework further demonstrated that spectral clustering was uniquely capable of isolating rare but meaningful physiological patterns. Unlike K-means, HDBSCAN, or GMM, which either absorbed minority states into dominant clusters or achieved only marginal enrichment, spectral clustering preserved a distinct 266-window cluster with 31-fold anxiety enrichment. This capacity to recover rare states highlights the importance of manifold-aware approaches for affective computing.

Comparisons with prior work highlight fundamental methodological differences. Supervised approaches dominate the literature, with reported accuracies of 75–96% on controlled datasets such as WESAD [34,35] but at the cost of extensive labeling. Alkurdi et al. [36] reported F1 scores up to 0.99 on WESAD but only 0.18–0.66 when applied to real-world data, illustrating the generalization challenge. Similarly, Sahu et al. [37] achieved AUC values of 0.60–0.73 but required complete annotation of all data. In contrast, meaningful anxiety signatures were discovered here from 890,494 windows with only 212 containing anxiety labels, demonstrating that rare anxiety states can be uncovered without reliance on large labeled corpora.

The systematic feature ablation revealed optimal performance with six features (silhouette = 0.339) compared to four features (0.215) or twelve features (0.284). The selected features—means and standard deviations for temperature, PR, MET, and phasic EDA—are all obtainable from commercial wearables, increasing translational potential.

5.3. Practical Applications

The discovered signatures enable objective anxiety monitoring implementable on existing wearable platforms. Cluster 6's extreme physiological profile provides clear detection thresholds: Peripheral temperature drops exceeding five standard deviations concurrent with PR elevations above three standard deviations could trigger alerts. The multi-phenotype stratification supports tiered intervention strategies, with the Extreme phenotype (Cluster 6) indicating potential need for workplace wellness support, while Mild phenotypes (5–10% conditional rates) might prompt preventive recommendations.

The temporal stability demonstrated through bootstrap validation—with Cluster 6 consistently recovered across 50–90% subsamples with <1% feature variation—suggests these signatures are robust enough for real-world deployment.

5.4. Study Design Rationale

The focus on experienced STEM professionals (minimum of four years) was deliberate, providing ecological validity through authentic workplace stressors such as project deadlines, technical problem-solving, and performance metrics. This design ensured exposure to genuine occupational anxiety distinct from the student samples commonly used in prior research, thereby strengthening the real-world relevance of the findings for affective computing.

5.5. Limitations and Future Work

Several limitations warrant consideration. The participant cohort was predominantly male (82%), which may limit generalizability given known sex differences in autonomic physiology.

The post hoc validation relied on temporal alignment between physiological windows and EMA self-reports, with only 0.15% of windows meeting strict alignment and quality criteria (± 3 -minute tolerance, >80% signal validity). This sparsity introduces potential

biases: EMA entries may over-represent emotionally salient moments when participants were motivated to log their state; reporting may concentrate during breaks rather than peak-stress periods; and sparse temporal coverage limits comprehensive characterization of anxiety episodes. While the 31-fold enrichment in Cluster 6 ($p = 0.0022$) and the consistent physiological gradients across clusters indicate that the labeled subset captured meaningful anxiety–physiology associations, broader validation coverage would strengthen population-level inferences. Future work should employ denser labeling protocols—such as prompted EMA at fixed intervals or experience sampling triggered by physiological patterns—to improve validation coverage and reduce reporting bias. A formal sensitivity analysis of enrichment estimates across alternative tolerances (e.g., ± 1 min, ± 5 min) was not performed and is flagged as a specific direction for future work.

Additionally, individual differences (e.g., trait anxiety, coping styles) were not analyzed, which may influence physiological signatures. Relatedly, anxiety was analyzed as a unitary affective construct rather than differentiated into clinical subtypes or trait-level profiles, reflecting common constraints in ecological wearable research where diagnostic labels and dense subtype-specific self-reports are rarely available [16]. Future work integrating wearable sensing with standardized clinical assessments and richer ecological momentary protocols could examine how the autonomic phenotypes identified here map onto specific anxiety subtypes and individual vulnerability profiles.

While the 31-fold enrichment in Cluster 6 is statistically significant ($p = 0.0022$), the absolute prevalence (0.75%) means this extreme state remains rare. The multi-phenotype model partially addresses the base rate challenge by distinguishing extreme from mild manifestations, but deployment as a screening tool would still require careful calibration to balance sensitivity and specificity. This rarity limits fine-grained analyses of recurrence patterns and demographic variability, and replication in independent occupational cohorts will be necessary to assess population-level generalizability of this extreme phenotype.

The absence of intervention data represents an additional gap. Moreover, although ecologically grounded, situational context (e.g., task demands, social interactions, time of day, environmental conditions) was not systematically modeled, limiting attribution of when and why physiological–anxiety states arise.

Future research should extend validation to other professional domains (e.g., health-care, finance, education), conduct within-person longitudinal tracking to assess cluster stability and transitions, and integrate contextual data (calendar events, project deadlines, and communication patterns) to reduce false positives. Intervention protocols triggered by detection of Cluster 6 or sustained High-Activation states should be developed and empirically tested.

6. Conclusions

The first large-scale unsupervised discovery of anxiety signatures from working professionals in naturalistic settings was presented in this study. A total of 44,443 h of continuous physiological data from 67 participants, collected over nine months, were processed. Spectral clustering of 890,494 five-minute windows revealed ten distinct autonomic states. Among these, a rare dysregulated state (Cluster 6, 0.03% prevalence) demonstrated 31-fold anxiety enrichment ($p = 0.0022$) and was characterized by extreme multi-system deviations: peripheral cooling ($d = -6.77$), tachycardia ($d = +4.41$), metabolic elevation ($d = +2.14$), and heightened electrodermal activity ($d = +0.89$).

Beyond this extreme signature, anxiety was found to manifest heterogeneously across four phenotypes, with conditional anxiety rates ranging from 5.17% to 28.57% when emotions were present. External validation using WESAD confirmed that the discovered states generalized to independent stress data, while the absence of Cluster 6 in labora-

tory conditions highlighted the specificity of extreme signatures to ecological contexts. The ability of the framework to identify meaningful patterns using only 0.15% labeled windows demonstrated that continuous emotion monitoring need not rely on prohibitive annotation efforts.

By establishing that anxiety can emerge as a discoverable physiological pattern through unsupervised analysis, both theoretical understanding of anxiety's autonomic signatures and practical pathways toward scalable workplace mental health monitoring were advanced. The discovered signatures provide an objective foundation for detection systems that complement traditional assessment while maintaining ecological validity in professional environments.

Author Contributions: Conceptualization, N.A., F.S. and B.Z.; Methodology, N.A.; Software, N.A.; Validation, N.A.; Formal analysis, N.A.; Investigation, N.A.; Resources, N.A. and B.Z.; Data curation, N.A.; Writing—original draft, N.A.; Writing—review & editing, N.A.; Visualization, N.A.; Supervision, N.A., F.S., B.Z. and T.H.; Project administration, N.A. and B.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Texas A&M University (protocol code #2023-0831D; approved on 22 September 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: The data supporting the findings of this study are not publicly available due to privacy restrictions on sensitive physiological recordings and participant confidentiality. De-identified summary statistics and analysis code may be made available from the corresponding author upon reasonable request, subject to IRB approval.

Acknowledgments: Special thanks go to the Texas A&M Master of Engineering Technical Management (METM) program and the cohorts who generously contributed data to this study. Their participation and trust provided the foundation for this research. The authors also gratefully acknowledge the computational resources provided by Texas A&M High Performance Research Computing (HPRC), without which the large-scale analyses described in this work would not have been feasible. During the preparation of this manuscript, the authors used generative AI tools for the purposes of LaTeX formatting assistance. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Javaid, S.F.; Hashim, I.J.; Hashim, M.J.; Stip, E.; Samad, M.A.; Ahababi, A.A. Epidemiology of anxiety disorders: Global burden and sociodemographic associations. *Middle East Curr. Psychiatry* **2023**, *30*, 12. [\[CrossRef\]](#)
2. World Health Organization. Mental Health at Work. 2024. Available online: <https://www.who.int/news-room/fact-sheets/detail/mental-health-at-work> (accessed on 15 December 2025).
3. Boucsein, W. *Electrodermal Activity*, 2nd ed.; Springer Science + Business Media: New York, NY, USA, 2012; pp. xviii, 618. [\[CrossRef\]](#)
4. Benedek, M.; Kaernbach, C. A continuous measure of phasic electrodermal activity. *J. Neurosci. Methods* **2010**, *190*, 80–91. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Ancillon, L.; Elgendi, M.; Menon, C. Machine Learning for Anxiety Detection Using Biosignals: A Review. *Diagnostics* **2022**, *12*, 1794. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Shikha, S.; Sethia, D.; Indu, S. A Systematic Review on Physiology-based Anxiety Detection using Machine Learning. *Biomed. Phys. Eng. Express* **2025**, *11*, 042002. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Schmidt, P.; Reiss, A.; Duerichen, R.; Marberger, C.; Van Laerhoven, K. WESAD (Wearable Stress and Affect Detection); ICMI: Boulder, CO, USA, 2018. [\[CrossRef\]](#)

8. Baghdadi, A.; Aribi, Y.; Fourati, R.; Halouani, N.; Siarry, P.; Alimi, A.M. *DASPS Database*; IEEE: Piscataway, NJ, USA, 2021. [\[CrossRef\]](#)
9. Elgendi, M.; Galli, V.; Ahmadizadeh, C.; Menon, C. Dataset of Psychological Scales and Physiological Signals Collected for Anxiety Assessment Using a Portable Device. *Data* **2022**, *7*, 132. [\[CrossRef\]](#)
10. Kirschbaum, C.; Pirke, K.M.; Hellhammer, D.H. The 'Trier Social Stress Test'—A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology* **1993**, *28*, 76–81. [\[PubMed\]](#)
11. Leichsenring, F.; Leweke, F. Social Anxiety Disorder. *N. Engl. J. Med.* **2017**, *376*, 2255–2264. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Yamamori, Y.; Robinson, O.J. Computational perspectives on human fear and anxiety. *Neurosci. Biobehav. Rev.* **2023**, *144*, 104959. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Ahmadi, N.; Hammond, T. Recognizing and Responding to Human Emotions: A Survey of Artificial Emotional Intelligence for Cooperative Social Human-Machine Interactions. In Proceedings of the 2023 IEEE Frontiers in Education Conference (FIE), College Station, TX, USA, 18–21 October 2023; pp. 1–5. [\[CrossRef\]](#)
14. Saviola, F.; Pappaianni, E.; Monti, A.; Grecucci, A.; Jovicich, J.; De Pisapia, N. Trait and state anxiety are mapped differently in the human brain. *Sci. Rep.* **2020**, *10*, 11112. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Sylvers, P.; Lilienfeld, S.O.; LaPrairie, J.L. Differences between trait fear and trait anxiety: Implications for psychopathology. *Clin. Psychol. Rev.* **2011**, *31*, 122–137. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Ahmadi, N.; Abbasian, P.; Hou, Q.; Powell, L.; Hammond, T. Physiological Sensing and Machine Learning Approaches Toward Anxiety Detection: A Systematic Review. *IEEE Sens. Rev.* **2026**, *3*, 89–116. [\[CrossRef\]](#)
17. Canals, J.; Voltas, N.; Hernández-Martínez, C.; Cosi, S.; Arija, V. Prevalence of DSM-5 anxiety disorders, comorbidity, and persistence of symptoms in Spanish early adolescents. *Eur. Child Adolesc. Psychiatry* **2019**, *28*, 131–143. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Sau, A.; Bhakta, I. Screening of anxiety and depression among the seafarers using machine learning technology. *Inform. Med. Unlocked* **2019**, *16*, 100149. [\[CrossRef\]](#)
19. Khalid, M.; Willis, E. A Brief Survey of Machine Learning Methods for Emotion Prediction using Physiological Data. *arXiv* **2022**, arXiv:2201.06610. [\[CrossRef\]](#)
20. Tripathi, A.; Choudhury, T. Permuted layer-based CNN for Emotion Detection with Multi-Modality Physiological Signals. In Proceedings of the 2023 IEEE International Conference on Contemporary Computing and Communications (InC4), Bangalore, India, 21–22 April 2023; IEEE: New York, NY, USA, 2023; Volume 1, pp. 1–5. [\[CrossRef\]](#)
21. Egger, M.; Ley, M.; Hanke, S. Emotion Recognition from Physiological Signal Analysis: A Review. *Electron. Notes Theor. Comput. Sci.* **2019**, *343*, 35–55. [\[CrossRef\]](#)
22. Mishra, R.K.; Ansari, A.R.; Mishra, V.; Jothi, A.A. Machine Learning for Prediction of Emotion Towards Digital Assets. In Proceedings of the International Conference on Advances in Signal Processing and Communication Engineering, ICASPACE 2023, Hyderabad, India, 28–29 April 2023; Lecture Notes in Electrical Engineering; Springer: Singapore, 2024; Volume 1157, pp. 489–499. [\[CrossRef\]](#)
23. Setz, C.; Arnrich, B.; Schumm, J.; La Marca, R.; Tröster, G.; Ehlert, U. Discriminating stress from cognitive load using a wearable EDA device. *IEEE Trans. Inf. Technol. Biomed. Publ. IEEE Eng. Med. Biol. Soc.* **2010**, *14*, 410–417. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Petrescu, L.; Petrescu, C.; Mitruț, O.; Moise, G.; Moldoveanu, A.; Moldoveanu, F.; Leordeanu, M. Integrating Biosignals Measurement in Virtual Reality Environments for Anxiety Detection. *Sensors* **2020**, *20*, 7088. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Vaz, M.; Summavielle, T.; Sebastião, R.; Ribeiro, R.P. Multimodal Classification of Anxiety Based on Physiological Signals. *Appl. Sci.* **2023**, *13*, 6368. [\[CrossRef\]](#)
26. Alkurdi, A.; Clore, J.; Sowers, R.; Hsiao-Wecksler, E.T.; Hernandez, M.E. Resilience of Machine Learning Models in Anxiety Detection: Assessing the Impact of Gaussian Noise on Wearable Sensors. *Appl. Sci.* **2025**, *15*, 88. [\[CrossRef\]](#)
27. Henry, J.; Lloyd, H.; Turner, M.; Kendrick, C. On the Robustness of Machine Learning Models for Stress and Anxiety Recognition from Heart Activity Signals. *IEEE Sens. J.* **2023**, *23*, 14428–14436. [\[CrossRef\]](#)
28. Panda, R.; Kumar, R.; Biradar, O. Anxiety Detection on ECG Signal using Fuzzy Deep Learning. *Procedia Comput. Sci.* **2025**, *258*, 1823–1832. [\[CrossRef\]](#)
29. Saha, U.; Sammya, S.D.; Saha, P.; Fattah, S.A.; Shahnaz, C. An Attention Guided Lightweight Network-Based Scheme for Anxiety Detection Using Multimodal Analysis of Single-Channel Wearable ECG and RSP Sensor Signals. *IEEE Sens. Lett.* **2025**, *9*, 1–4. [\[CrossRef\]](#)
30. Hoemann, K.; Khan, Z.; Feldman, M.J.; Nielson, C.; Devlin, M.; Dy, J.; Barrett, L.F.; Wormwood, J.B.; Quigley, K.S. Context-aware experience sampling reveals the scale of variation in affective experience. *Sci. Rep.* **2020**, *10*, 12459. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Alberdi, A.; Aztiria, A.; Basarab, A.; Cook, D.J. Using smart offices to predict occupational stress. *Int. J. Ind. Ergon.* **2018**, *67*, 13–26. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Plutchik, R. The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *Am. Sci.* **2001**, *89*, 344–350. [\[CrossRef\]](#)

33. Shaffer, F.; Ginsberg, J.P. An Overview of Heart Rate Variability Metrics and Norms. *Front. Public Health* **2017**, *5*, 258. [[CrossRef](#)] [[PubMed](#)]
34. Jain, A.; Kumar, R. Machine Learning based Anxiety Detection using Physiological Signals and Context Features. In Proceedings of the 2024 2nd International Conference on Advancement in Computation & Computer Technologies (InCACCT), Gharuan, India, 2–3 May 2024; pp. 116–121. [[CrossRef](#)]
35. Sebastião, R. Classification of Anxiety Based on EDA and HR. In *Proceedings of the IoT Technologies for HealthCare*; Goleva, R., Garcia, N.R.d.C., Pires, I.M., Eds.; Springer: Cham, Switzerland, 2021; pp. 112–123. [[CrossRef](#)]
36. Alkurdi, A.; He, M.; Cerna, J.; Clore, J.; Sowers, R.; Hsiao-Weckler, E.T.; Hernandez, M.E. Extending Anxiety Detection from Multimodal Wearables in Controlled Conditions to Real-World Environments. *Sensors* **2025**, *25*, 1241. [[CrossRef](#)] [[PubMed](#)]
37. Sahu, N.K.; Gupta, S.; Lone, H.R. Are Anxiety Detection Models Generalizable? A Cross-Activity and Cross-Population Study Using Wearables. *arXiv* **2025**, arXiv:2504.03695. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.