

Personalized Emotion Recognition Using Physiological Signals (EDA & PPG): A Temperament-Informed Approach

Maryam Saidi

msaidi@aut.ac.ir

Shahrood University of Technology

Mahdi Kafaee

Shahrood University of Technology

Asmar Ghaderi

Shahrood University of Technology

Yasaman Shabani

Shahrood University of Technology

Research Article

Keywords: Temperament, Mezaj (Mizaj), Electrodermal Activity (EDA), Photoplethysmography (PPG), Traditional Persian Medicine (TPM), Emotion recognition, Affective computing

Posted Date: April 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9172405/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Individual differences significantly affect physiological responses to emotional stimuli, challenging generalized emotion recognition models. This study investigates the use of temperament as a contextual cue for personalizing emotion classification from physiological signals. Specifically, it focuses on the temperament as defined in Traditional Persian Medicine (TPM), known as Mezaaj, which is characterized along warm/cold and moist/dry dimensions. To this end, electrodermal activity (EDA) and Photoplethysmography (PPG) signals were recorded from 124 participants during the induction of four emotions: scary, joyful, relaxing, and boring. Participants also completed a temperament questionnaire. We hypothesize that temperament systematically modulates the mapping of emotional states within the arousal–valence space. Three personalization strategies were evaluated: (1) training separate classifiers for each temperament group, (2) including temperament scores as additional input features, and (3) adjusting classifier outputs based on temperament via a secondary model. Features in both time and frequency domains were extracted, and then analyzed using classical dimensionality reduction and classification approaches. Results show that incorporating temperament improves accuracy in most binary classification tasks, with the largest gains in the joy vs. relaxation condition. Four-class classification also benefits from temperament information, though improvements are smaller and more nuanced. These findings suggest that temperament-driven modulation is a complex, emotion-specific mechanism rather than a simple linear adjustment. This work highlights the potential of integrating temperament into physiological emotion recognition to enhance the development of personalized affective computing.

1. Introduction

In recent years, emotion recognition using physiological signals has emerged as a vital research area in affective computing, with applications ranging from human-computer interaction to mental health monitoring and adaptive learning systems. Physiological signals provide a more direct, unconscious measure of emotional arousal through the autonomic nervous system (ANS) (Miao et al., 2025)(Ahmad & Khan, 2022).

Despite notable advancements in signal processing and machine learning techniques, high inter-individual variability in emotional responses remains a major challenge in this field (Miao et al., 2025)(Goshvarpour et al., 2022). Many existing models implicitly assume a relatively universal mapping between physiological responses and emotional states. However, individuals often exhibit different autonomic patterns in response to similar stimuli, and conversely, similar physiological patterns may correspond to distinct subjective experiences across users (Ahmad & Khan, 2022)(Zhang et al., 2018). In addition, stable user-specific traits—such as temperament or personality—systematically influence how emotions are perceived, regulated, and physiologically expressed, thereby limiting the robustness and generalizability of generalized models(Goshvarpour et al., 2022)(Sears et al., 2006).

To address this limitation, recent studies have increasingly adopted personalized or subject-dependent approaches. These models incorporate individual-level information to better account for variability in emotional reactivity and baseline physiology, leading to improved performance (J. Li & Washington, 2024)(Zhao et al., 2018). As summarized in Table I, prior personalization strategies have considered a range of contextual and trait-based factors, including personality dimensions, mood states, demographic variables, cognitive states, and physiological baselines.

Table I. Factors commonly used in personalized emotion recognition models reported in the literature.

Category	Example Factors
Temperament (Muller, 2011)	Harm Avoidance, Novelty Seeking (e.g., TCI)
Personality Traits (Hosseini et al., 2023)(Kutt et al., 2020)(Zhao et al., 2018)	Big Five: OCEAN
Mood (Can & Ersoy, 2022)	PANAS, self-reported mood states
Demographics (Goshvarpour et al., 2022)	Age, Gender
Cognitive States (Tsouli et al., 2017)	Attention, fatigue, cognitive load
Physiological Baselines (Shu et al., 2018)	Resting HR, baseline EDA

While the benefits of personalization are widely acknowledged, comparatively few studies have systematically investigated the role of temperament in physiological emotion recognition (Shu et al., 2018). Previous research on personality-aware modeling has primarily focused on the Big Five traits(Poyraz et al., 2017). For instance, Zhao et al. (Zhao et al., 2018) demonstrated that incorporating Big Five traits improved multimodal emotion recognition performance relative to generic models. However, broad personality constructs largely capture long-term behavioral tendencies rather than directly indexing physiological reactivity (McCrae & Costa Jr, 1997).

In contrast, temperament—both in contemporary psychological frameworks and in Traditional Persian Medicine (TPM)—is conceptualized as a biologically influenced and relatively stable disposition that shapes emotional regulation, arousal thresholds, and response intensity (Tansaz et al., 2023)(Alizadeh et al., 2025)(Mojahedi et al., 2023).Unlike transient states such as mood, temperament reflects enduring patterns in how individuals perceive, experience, and respond to stimuli, making it an ideal candidate for personalizing emotion recognition systems(Canli & Amin, 2002)(Talge et al., 2008)(Sears et al., 2006).

Empirical findings support the link between temperament traits and physiological markers. Talge et al. (Talge et al., 2008) examined associations between fearful temperament—a temperament trait—and physiological responses, including cortisol, sympathetic activation, and parasympathetic withdrawal, in preschoolers performing stress-inducing tasks. They found that, although correlations were modest, higher fearful temperament was linked to increased cortisol reactivity and weakly associated with sympathetic activation, suggesting heightened physiological sensitivity in children with a fearful disposition. Similarly, Li et al. (M. Li et al., 2017) found that resting heart rate (HR) and respiratory sinus arrhythmia (RSA) profiles differentiated early temperament traits in young children, indicating stable physiological signatures associated with temperamental dimensions. These studies suggest that temperament may modulate both baseline physiological states and reactivity to emotional stimuli.

Despite such evidence, the explicit integration of temperament into computational models for physiological emotion recognition remains limited. EEG studies have shown associations between temperament-related traits (e.g., trait anxiety, negative emotionality) and neural indicators such as frontal alpha asymmetry (Vecchio & De Pascalis, 2020)(Adolph & Margraf, 2017). Likewise, dimensions such as Harm Avoidance (HA) and Novelty Seeking (NS) have been linked to differences in emotional processing biases and resting-state neural patterns (Muller, 2011). Collectively, these findings highlight temperament as a biologically grounded factor influencing emotional processing, yet its computational exploitation in peripheral signal-based emotion recognition is still scarce.

Given the diversity of temperament frameworks, it is important to clarify conceptual definitions. Table II summarizes key definitions from major psychological models alongside Avicenna’s temperament theory.

Table II. Definitions of temperament in psychological theories.

Theory/ Model	Definition of Temperament	Core Dimensions
Thomas & Chess (1977) (Thomas & Chess, 1977)	Biologically-based individual differences in emotional reactivity and self-regulation, observable from infancy.	9 dimensions: activity level, rhythmicity, approach/withdrawal, adaptability, intensity, threshold, mood, distractibility, attention span
Buss & Plomin (1984) (Buss & Plomin, 2014)	Heritable traits present from early life that remain relatively stable over time.	Emotionality, Activity, Sociability
Rothbart (2001) (Rothbart & Bates, 2006)	Individual differences in reactivity and self-regulation, influenced by biology and evident in early life.	Surgency/Extraversion, Negative Affectivity, Effortful Control
Cloninger's TCI (1993) (Cloninger et al., 1993)	Temperament consists of automatic emotional responses that are heritable and stable, separate from character traits.	Novelty Seeking, Harm Avoidance, Reward Dependence, Persistence
Avicenna's Temperament (Mezaj) Theory (~ 1025) (Avicenna, 2015)	A combination of four qualities (warm, cold, wet, dry) that determine the physical and psychological characteristics of person. It is rooted in the balance of the body's basic elements.	9 main types (e.g., warm-wet, cold-dry, balanced)

Across these frameworks, temperament is consistently described as biologically influenced and relatively stable, closely linked to neural and autonomic systems involved in emotional processing(Cloninger et al., 1993)(Rothbart & Bates, 2006).

While modern psychology relies primarily on standardized psychometric instruments, TPM conceptualizes temperament through the balance of four fundamental qualities—warm, cold, wet, and dry(Kafaee et al., 2025).

Although originating from distinct theoretical backgrounds, both perspectives classify individuals into stable types associated with predictable affective tendencies (Salmannezhad et al., 2018)(Alizadeh et al., 2025)(Ahmadi et al., 2023) (Cloninger et al., 2019). In recent decades, TPM temperament assessment has evolved through the development and validation of structured questionnaires (Salmannezhad et al., 2018)(Mojahedi et al., 2014), enabling reproducible measurement in research settings. Table III provides a comparative overview of how temperament is conceptualized and operationalized in Modern Psychology and TPM.

Table III. Comparison of Temperament in Traditional Persian Medicine and Modern Psychology.

Aspect	Modern Psychology	Traditional Persian Medicine
Theoretical basis	Neuroscience, developmental psychology, genetics	Humoral theory, philosophical foundations, genetics
Assessment	Standardized psychometric questionnaires	Standardized temperament questionnaires + Clinical observation
Focus	Emotional and behavioral regulation	Physical constitution and emotional traits (psychosomatic view)
Goal	Mental health, personality profiling	Lifestyle optimization, health maintenance, and treatment

Building upon this theoretical background, the present study introduces a temperament-aware framework for emotion recognition using peripheral physiological signals. Specifically, we focus on electrodermal activity (EDA) and photoplethysmography (PPG). EDA reflects sympathetic modulation of sweat gland activity, while PPG captures blood volume pulse dynamics and heart rate variability—both sensitive to emotional arousal(Poh et al., 2010)(Setz et al., 2009). Due to their non-invasive nature and compatibility with wearable devices, these modalities are widely adopted in real-time affective monitoring systems.

Typically, features from EDA and PPG are extracted in time, frequency, and nonlinear domains and mapped to arousal–valence dimensions using machine learning models. For example, Gohumpu et al. (Gohumpu et al., 2023) utilized EDA and PPG features from the DEAP dataset and achieved competitive performance using Support Vector Machines (SVMs). However, such models generally do not explicitly account for stable temperament-driven modulation of physiological responses. The present work addresses this gap by systematically incorporating temperament information into the emotion recognition pipeline.

This paper hypothesizes that temperament can be represented as a style-shifting transformation within the arousal–valence emotion space, denoted as $f_T(v, a)$. Building on temperament theory—characterized along dimensions such as warm–cold and dry–wet, as described later—this formulation adjusts the emotional response vector (v, a) according to an individual’s temperament T (Fig. 1). Within this framework, temperament functions as a modulatory factor that shapes the subjective experience and expression of emotional states. For instance, consider two individuals exposed to the same stimulus, such as a suspenseful video clip. Both may exhibit similar physiological responses, like comparable increases in electrodermal activity (EDA) or heart rate (PPG). However, one person may experience high anxiety while the other feels mild excitement, reflecting differences in how their temperament modulates the interpretation of the same physiological signals. This transformation may involve baseline shifts, scaling of arousal intensity, or modification of valence perception.

To investigate this hypothesis, this study addresses the following research questions:

- **RQ1:** Does incorporating temperament information improve binary emotion classification from EDA and PPG signals?
- **RQ2:** Does temperament-aware modeling enhance multiclass emotion classification performance?
- **RQ3:** Which personalization strategy is most effective for integrating temperament into physiological emotion recognition models

To answer these questions, we implement and compare three temperament-aware modeling strategies:

1. **Temperament-Specific Models:** Developing separate classification models for each major temperament type.
2. **Temperament as Input Feature:** Including temperament scores as additional input features in a shared model to capture individual modulation of signals.
3. **Model Conditioning by Temperament:** Adjusting model parameters (e.g., attention weights or decision thresholds) dynamically based on the user’s temperament.

The remainder of this paper is structured as follows. Section 1.1 presents a comprehensive overview of the temperament framework, its psychometric grounding in TPM, and its operationalization through standardized self-report questionnaires. This section establishes the theoretical motivation for exploring temperament as a contextual factor in physiological emotion recognition. Section 2 introduces the experimental dataset, including details on participant recruitment, demographic distribution, and experimental design. The section elaborates on the recording setup and modalities, focusing on EDA and PPG signals, which were acquired under controlled emotional stimuli. Preprocessing steps such as noise reduction and normalization are described, followed by the methodology for feature extraction and classification. Section 3 presents the empirical results of the study. Accuracy of emotion recognition are reported for both binary and multiclass classification settings. Comparative analyses are conducted to quantify the contribution of temperament-aware strategies relative to baseline models. Section 4 provides an in-depth discussion of the findings. Broader implications for affective computing, personalized human–computer interaction, and adaptive mental health technologies are also discussed. Finally, Section 5 concludes the paper by summarizing the key contributions, emphasizing the novelty of integrating TPM temperament assessment with physiological signal processing.

1.1 Temperament

Personalized medicine in modern healthcare aims to tailor medical treatment to the individual characteristics of each patient, including their genetics, lifestyle, and environment, to achieve more effective and precise outcomes. One approach that aligns with this goal can be found in some traditional medicine perspectives, which focuses on the concept of temperament. The concept of Temperament offers a holistic framework for understanding an individual's unique physical and psychological traits. It explains variations in body type, metabolism, and even personality. Crucially, this system is applied proactively for prevention through tailored diets and lifestyles that maintain equilibrium. In illness, diagnosis hinges on identifying the temperamental imbalance, allowing for highly personalized treatments—from specific dietary and lifestyle to herbal remedies—designed to restore the body's natural harmony and address the root cause of the disease. Contemporary studies have shown that temperament is correlated with basal metabolic rate (BMR) (Kafaei et al., 2024) and can be described by genetics (Rezadoost et al., 2016)(Cloninger et al., 2019).

The basic elements interact with each other in different proportions in the body. As a result of this interaction, a stable and homogeneous balance is provided, known as temperament (Avicenna, 2015). This balance can be characterized by two orthogonal dimensions: the hot-cold dimension and the wet-dry dimension (Kafaei et al., 2024)(Shirbeigi et al., 2017). Individuals in a population will typically have statistical distribution in both dimensions. People who are close to the mean are considered to have a balanced temperament. Other people are considered to have an unbalanced temperament, depending on their distance from the center. Literally, a temperament in which there is an increase in warm and wet is called Sanguine, an increase in warm and dry is called Choleric, an increase in cold and dry is called Melancholic, and an increase in cold and wet is called Phlegmatic. This is while the dominance of each quality may also occur alone. The predominance of each of these qualities can lead to specific physical and psychological characteristics, each of which is described in detail in the relevant literature (Avicenna, 2015)(Shahabi et al., 2008). Temperament can be assessed using anthropometric criteria, including tactile sensitivity, muscle and fat mass, hair and skin characteristics, body shape, speed of impression, patterns of sleep and wakefulness, physical functions, the quality of excretory substances, as well as mental functions (Tansaz et al., 2023)(Salmannezhad et al., 2018). In order to quantify the two dimensions, standard questionnaire were developed (see section 2.5). By identifying temperament, the unique characteristics of individuals can be predicted.

2. Materials and Methods

2.1 Participants

A total of 124 students (64 males and 60 females), aged 18 to 25 years (mean $\pm$ SD: 21.69 $\pm$ 1.58 years), voluntarily participated in the experiment. Participants were required to follow specific preparatory instructions: they had to abstain from caffeine, tea, alcohol, tobacco, and any narcotics for at least four hours prior to the experiment. Additionally, they were asked to avoid extreme emotional states, stress, and certain medications—including sedatives, antidepressants, hypnotics, stimulants, or hallucinogens—for 24 hours before the experiment.

Individuals with pre-existing medical conditions—such as hypertension, diabetes, neurological disorders, pregnancy, menstruation (due to potential hormonal and emotional fluctuations) (Buckwalter et al., 1999) (Sundström Poromaa & Gingnell, 2014), or other relevant health issues—were not permitted to participate in the study. Throughout the experiment, participants remained unaware of the specific experimental stimuli.

2.2 Emotional Stimuli

Among several types of emotional stimulation, such as music, text, voice, and pictures we chose picture movies. In this study, four video clips were used as stimuli to induce four primary emotional states: happiness, fear, relaxation, and boredom. The selection criteria for the videos were based on the arousal level and emotional valence of each stimulus, as determined by the 'Circumplex Model of Affect' (Fig. 2)(Russell et al., 1989)(Russell, 1980). Our chosen stimuli were specifically curated to represent the four core quadrants of this model: high arousal-negative valence (scary), high arousal-

positive valence (joyful), low arousal-positive valence (relaxing), and low arousal-negative valence (boring). This approach ensures a comprehensive coverage of the affective space, allowing us to investigate the potential modulatory effect of temperament across the entire spectrum of emotional experiences, rather than being confined to a specific region.

The video contents were carefully chosen to evoke specific emotional states: a video of a child was used to induce happiness, a nature scene to evoke relaxation, the sound of dripping water to elicit boredom, and a dark room scene to provoke fear. To minimize potential order effects, the clips were presented to participants in one of four randomized sequences:

1. Scary / Joyful / Boring / Relaxing;
2. Joyful / Relaxing / Boredom / Scary
3. . Relaxing / Joyful / Scary / Boring
4. Boring / Joyful / Scary / Relaxing.

2.3 Data Acquisition Device

A dedicated multimodal biosignal acquisition system was designed and developed in our laboratory specifically for this study. The system enables simultaneous real-time recording of EDA and PPG signals and interfaces with a PC via a USB connection. All signals were sampled at 256 Hz and stored for offline processing in MATLAB.

To ensure signal fidelity, the performance of the developed system was validated against a commercial reference device (PowerLab, ADInstruments). Comparative recordings demonstrated comparable waveform morphology and amplitude consistency across both systems, confirming the reliability of the custom hardware for physiological data acquisition.

For EDA acquisition, electrodes placed on the distal phalanges of the non-dominant hand fingers (as illustrated in Fig. 3). A constant-voltage configuration was employed for skin conductance measurement. The analog front-end included appropriate amplification and anti-aliasing filtering prior to digitization.

For PPG acquisition, a reflective optical sensor was used and positioned on the index finger of the non-dominant hand (see Fig. 3). The sensor operated at a standard infrared wavelength suitable for peripheral blood volume pulse detection. Signal acquisition for both modalities was synchronized at the hardware level within the same embedded system to ensure temporal alignment.

Subsequently, digital preprocessing was applied in MATLAB, including band-limited filtering to remove motion artifacts and high-frequency noise (details provided in the Signal Processing section)

2.4 Experiment Procedure

To minimize environmental and physiological confounding factors, all recordings were carried out in a controlled laboratory setting with stable temperature, uniform lighting, and absence of external noise or distractions (Hygge & Knez, 2001). Before data collection, participants rested for 10 minutes to allow their physiological baseline (e.g., heart rate, skin conductance) to stabilize. If elevated heart rate or self-reported anxiety was observed, additional time was provided until relaxation was achieved (Measures et al., 2012).

During the experiment, each participant was seated comfortably in front of a monitor, and the procedure was explained in detail. The acquisition setup consisted of EDA and PPG sensors attached to the participant's left hand: two EDA electrodes placed on the index and ring fingers, and a PPG sensor on the middle finger. In a single case where the PPG signal amplitude was very weak in the left hand, the sensor was relocated to the right hand, which resulted in a stronger signal amplitude (Fig. 3).

Participants were instructed to remain still and refrain from talking, moving their hands, or engaging in any actions that could introduce artifacts into the recordings. Each session consisted of four video clips, each lasting 23 seconds, separated by 30-second inter-trial intervals. These intervals served as resting periods, allowing participants to return to baseline after the emotional arousal induced by the preceding stimulus (Morris et al., 2013). Throughout the entire experiment, EDA and PPG signals were recorded simultaneously.

2.5 Questionnaires

According to the objectives of the experiment, participants completed three questionnaires, described as follows:

General Health Questionnaire (GHQ)

Participants completed the 28-item General Health Questionnaire (GHQ-28) (Goldberg & Hillier, 1979), which evaluates mental health across four domains: somatic symptoms, anxiety and insomnia, social dysfunction, and depression.

Responses were rated on a 4-point Likert scale (0 = best state, 3 = worst state). In accordance with the standardized cutoff score (≥ 24 in the 0–3 scoring system; (Goldberg & Hillier, 1979)), and after adjustment for the 1–4 scoring scheme, participants whose total scores met or exceeded the equivalent cutoff of 24 were excluded from further participation. This exclusion ensured that significant psychological distress did not confound the study variables.

TPM temperament Identification Questionnaire (MIQ)

Prior to the presentation of experimental stimuli, all participants completed the standardized Salmannejad-Mojahedi Temperament Questionnaire (Salmannejad et al., 2018) to assess their baseline temperament. This self-report instrument comprises 20 items rated on a five-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). According to the scoring protocol, the total score of the first 15 items reflects the warm–cold temperament dimension (Warmth/Coldness), whereas the total score of the last 5 items represents the dry–wet temperament dimension (Dryness/Wetness). These two dimensions are regarded as orthogonal and independent.

Self-Assessment Manikin (SAM) Questionnaire

After each video stimulus, participants' emotional responses were assessed using the Self-Assessment Manikin (SAM) (Bradley & Lang, 1994), a non-verbal, pictorial tool designed to directly measure two core affective dimensions: valence (pleasantness) and arousal (activation). Developed by Bradley and Lang (1994), SAM provides a simplified alternative to verbal scales such as the Semantic Differential, reducing the required judgments from 18 items to just a few pictorial choices. Figure 4 shows this pictorial questionnaire.

Participants rated valence on a continuum ranging from a frowning, unhappy figure (low valence) to a smiling, happy figure (high valence), and arousal from a relaxed, sleepy figure (low arousal) to an excited, wide-eyed figure (high arousal). The instrument is typically presented as five pictograms per dimension, with the option to select between figures for a finer nine-point scale. This design enables rapid and intuitive assessment of emotional states.

Because of its non-verbal, culture-independent format and strong convergent validity with traditional verbal measures—particularly for valence and arousal—SAM has become one of the most widely used tools in emotion research.

2.6 Analysis Overview

As mentioned previously, this paper investigates three strategies to integrate temperament into emotion recognition models using EDA and PPG signals, as illustrated in Fig. 5

1. Temperament as an input feature (Temperament Fusion-a in Fig. 5)— involves concatenating temperament scores with physiological features to create a unified model. Although various machine learning classifiers can be employed in this

approach (e.g., Logistic Regression, Random Forest, XGBoost, or k-Nearest Neighbors), the primary objective of this study is to compare three temperament-integration strategies under a controlled modeling framework. To ensure a fair and unbiased comparison across approaches, we intentionally adopted a single classifier architecture for all experiments. Support Vector Machines (SVMs) were selected due to their strong generalization capability in small-to-moderate sample sizes, robustness to high-dimensional feature spaces, and well-established performance in physiological emotion recognition tasks. By keeping the classifier fixed across all three strategies, performance differences can be attributed to the temperament-integration mechanism rather than variations in model architecture.

2. Temperament-weighted adaptation (Temperament-Weighted-b in Fig. 5)– Fine-tuning a base model by adjusting feature weights or loss functions based on temperament group. In this framework, a primary SVM classifier was first trained exclusively on physiological features extracted from EDA and PPG signals. Instead of directly using the predicted class labels, the posterior class probabilities produced by the primary SVM (obtained via probability calibration) were retained. To incorporate temperament information, these probability outputs were concatenated with the normalized temperament scores of each participant to form an extended feature vector. A secondary SVM classifier was then trained on this combined representation. In this configuration, temperament does not alter the physiological signal features themselves; rather, it modulates the final decision stage by conditioning the classifier on both physiological-derived emotion probabilities and stable temperament traits

3. Temperament-specific models (Temperament-Specific -c in Fig. 5)– Training separate classifiers for each temperament group, such as warm vs. cold. Specifically, two grouping strategies were examined. In the first analysis, participants were divided into warm and cold groups based on a threshold value of 48 on the Coldness–Warmness scale, and emotion classification performance was evaluated separately for each group. In the second analysis, participants were categorized into dry and wet groups using a threshold value of 15 on the Dryness–Wetness scale, and the same classification procedure was repeated to assess the impact of this temperament dimension.

Preprocessing

Out of 124 participants, one was excluded based on GHQ criteria. Data from 14 participants were discarded due to incomplete acquisition of both modalities throughout the experimental session. The final dataset included recordings with complete synchronized PPG and EDA signals.

PPG signals were sampled at 256 Hz. A 6th-order Infinite Impulse Response (IIR) band-pass digital filter (bandwidth: ~0.3–22 Hz) was applied in zero-phase forward–backward mode to preserve pulsatile components while attenuating baseline drift, respiratory interference, and high-frequency noise. Systolic peaks were detected using an adaptive amplitude–distance thresholding algorithm. Inter-beat intervals (IBIs) were derived using an automated peak detection algorithm based on adaptive amplitude thresholding combined with physiological constraints on minimum and maximum pulse intervals (corresponding to plausible heart rate limits). Detected intervals falling outside the predefined range (40–180 bpm equivalent) were automatically flagged as artifacts and corrected using cubic spline interpolation.

For the EDA signal, short-term fluctuations were smoothed using a moving average filter with a 100-sample window, which attenuated high-frequency noise without substantially distorting the slower phasic and tonic components. To separate tonic and phasic components, convex optimization-based decomposition was applied. This method models the EDA signal as the sum of: a slowly varying tonic component (Skin Conductance Level; SCL), sparse phasic responses (Skin Conductance Responses; SCRs). Phasic driver activity was extracted, and event-related SCR features were computed, such as SCR amplitude, SCR frequency, Area under the SCR curve and Rise time

To preserve inter-individual variability while preventing scale dominance, z-score normalization (mean = 0, standard deviation = 1), was applied within-subject, not across participants.

Feature extraction

The features listed in Table IV were extracted from PPG and EDA signals based on previous studies in stress and emotion recognition. They are categorized into three main groups:

1. Statistical features: including measures such as mean, standard deviation, skewness, and kurtosis, which capture the overall distribution and variability of the signals.
2. Time domain and morphologic features: reflecting the shape and morphology of the signal waveform, such as peak-to-peak intervals, signal slopes, and pulse width.
3. Frequency-domain features: derived from spectral analysis, including power spectral density and band-specific energy components, which characterize the signal's oscillatory behavior.

These features have been widely used in affective computing research and provide complementary information for reliable stress and emotion recognition(Khaleghi et al., 2024).

Table IV. Extracted Features from PPG and EDA Signals. The corresponding feature number is shown in parentheses beside each feature name

EDA signal	Statistical Features	Time domain Features	Frequency Features
	Difference of 65th and 15th percentile (1), Mean of EDA (3), Std. of EDA (4), Mean of Normalized EDA (5), Std of Normalized EDA (6), Minimum of EDA (13), Maximum of EDA (14), Mean of EDA derivate (15), Mean negative EDA derivative (16), Number of negetive-derivative samples (17), Log of EDA mean (18), RMS of EDA (19), Skewness of EDA (20), Kurtosis of EDA (21), Mean of relative difference between EDA and its derivate (22),	SCRs count (7), Amplitude of max SCRs (2), Mean SCRs amplitude (8), Mean SCRs time (9), Mean SCRs AUC (10), Mean SCRs slop (11), Linear regression slope of EDA (12)	Log spectral power 0–2.4 Hz, 10 sub-bands (23–33)
PPG signal	Inter-beat Interval Mean(34), Std. of IBIs (35), Mean Heart Rate (36), Std. of Heart Rate (37), Std. of Successive Differences (38), RMS of Successive Differences (39), Difference between the Third and First Quantiles of Successive Differences (40), Median of Absolute Values of Successive Differences (41), Sum of Absolute Values of Successive Differences greater than 50 (42), NN50/Number of Successive Differences (43) 85th & 90th percentile of heart rate derivative (64, 65), 65th–75th percentile of heart rate values (66, 67), Total variation (68), Mean P–P interval(69), Min P–P interval (70), P–P range (71)	Triangular Indexes of HRV histogram (44, 45), Correlation Dimension (61), SD1/SD2 poicare (62), Sample entropy(63),	Energy in the following frequency bands: 0.1-0.2Hz (46), 0.2-0.3Hz (47), 0.3-0.4Hz (48), 0.04 – 0.01 Hz (49), 0.15 – 0.04 Hz (50), 0.01–0.08 Hz (51), 0.08–0.15 Hz (52), 0.4 – 0.15 Hz (53), 0.15-0.5Hz (54), 0-0.08 Hz (55), LF/HF (56), LF_0.01to_0.08/ (57) HF_0.15to_0.5 (58), (LF + MF)/HF (59), HF /(HF + LF) (60).

Classification

To normalize the features, z-normalization was applied. Given the large initial feature set, a two-step feature reduction procedure was employed. First, from the original 71 features, those with a p-value < 0.2 in the training data were retained. For binary classification tasks, the p-values were computed using the independent two-sample t-test (equivalent to one-way ANOVA for two groups), whereas for multi-class tasks, one-way ANOVA was applied. Second, Principal Component Analysis (PCA) was performed, retaining 50% of the eigenvectors corresponding to the largest eigenvalues. Importantly, the PCA transformation was computed exclusively on the training set to prevent data leakage.

The resulting reduced feature set was classified using a SVM. Model performance was evaluated using a Leave-One-Subject-Out (LOSO) cross-validation strategy. In each fold, feature reduction and classifier training were performed on all subjects except the one held out for testing, ensuring a fully subject-independent evaluation.

3. Results

3.1 Temperament Rating

Participants' temperaments are determined using questionnaire scores, which represent a two-dimensional model. The first dimension corresponds to warmth-coldness, ranging from 29 to 64, while the second dimension pertains to dryness-wetness, ranging from 7 to 23. Figure 6 illustrates the distribution of scores for individuals across these two temperament dimensions.

3.2 Emotional rating

Figure 7 presents the participants' ratings of the emotional stimuli. The distributions of arousal and valence for each stimulus are shown in different colors, with star markers indicating the $\frac{mean}{std}$ across participants. As illustrated, the subjective ratings for each stimulus closely align with the intended emotional targets, confirming that the stimuli effectively elicited the desired emotional states. The only notable deviation is observed for happiness, where lower arousal scores were reported compared to the expected levels. This discrepancy may reflect the tendency of individuals to perceive the arousal associated with positive emotions less intensely than that associated with negative emotions.

Previous research has shown that individual differences in temperament and personality traits can significantly influence emotional processing and recognition. For instance, individuals with high positive affectivity or extraversion tend to be more responsive to positive emotional cues, while those characterized by high neuroticism or behavioral inhibition are more sensitive to negative stimuli and tend to perceive emotional content more negatively [19]. Such temperamental biases can modulate both the subjective appraisal of emotional stimuli and the physiological responses they evoke [20].

In light of these findings, this section of the present study investigates the relationship between participants' subjective emotion ratings and their temperament dimensions. Table V, presents Pearson correlation coefficients and corresponding p-values for the four stimuli. As observed, a significant correlation exists only between valence ratings and individuals' warm-cold temperament dimension for scary stimulation. Specifically, individuals with warmer temperaments tended to assign higher valence scores to scary stimuli (P-Coef = 0.192), suggesting that underlying temperamental traits may modulate emotional appraisal. For the other stimuli, however, no temperament dimension showed a systematic relationship with participants' subjective emotional ratings.

Table V. Relation between subjective emotional rating and temperament scores

Emotional stimulation	Coldness-Warmness				Wetness/dryness			
	Arousal		Valance		arousal		Valance	
	P-Coef	P-value	P-Coef	P-value	P-Coef	P-value	P-Coef	P-value
Scary	-0.020	0.836	0.192	0.045*	-0.007	0.945	0.135	0.162
Joyful	0.124	0.199	0.143	0.137	-0.031	0.745	-0.039	0.681
Borings	0.048	0.617	0.146	0.130	0.096	0.316	0.003	0.975
Relaxing	0.105	0.274	0.095	0.325	0.057	0.553	-0.022	0.814

3.3 Physiological Signal Features across Temperament

This study suggests that temperament traits can modulate physiological responses to emotional stimuli, reflecting stable individual differences in autonomic and affective regulation. To explore this, we examined whether specific features extracted from PPG and EDA signals differed significantly across temperament groups. For each extracted feature (Table 4), statistical comparisons were performed between participants categorized into contrasting temperament groups based on the TPM scale.

For the warm–cold dimension, participants were dichotomized using a threshold value of 48 on the coldness–warmness score. The thresholds were primarily established according to the cut-off values reported for the TPM questionnaire by Salmannejad et al. (Salmannejad et al., 2018). Importantly, these cut-off points were also aligned with the median distribution of scores in our sample. Individuals with scores > 48 were classified as warm ($n = 52$), and those with scores < 48 were classified as cold ($n = 58$).

Similarly, for the wet–dry dimension, a threshold value of 15 was used on the wetness–dryness score. Participants with scores > 15 were classified as wet ($n = 53$), and those with scores < 15 were classified as dry ($n = 57$).

Between-group comparisons were conducted using an independent-samples t-test. This analysis aimed to identify which physiological markers are sensitive to underlying temperamental differences and may thus contribute to personalized models of emotion recognition. The statistical significance of group differences (p-values) is visualized in the Heatmap (Fig. 8). The total of 71 features were included in this analysis, comprising 33 features derived from EDA signals and the remaining features obtained from PPG signals. The evaluation was performed separately for each of the four emotional stimuli, enabling the identification of features that exhibited significant differences between the two temperament groups. These findings suggest that it is not possible to attribute stronger temperamental group differences to a single physiological signal, as both EDA and PPG exhibited significant differences in some of their features. However, for scary-evoking stimuli, the number of significantly different features between warm and cold groups was comparatively smaller.

A similar heatmap was generated for the dryness/wetness dimension of temperament (Fig. 8), which confirms the previous findings. Notably, this analysis shows an increased number of significant EDA features in response to the joyful stimulus.

3.4 Binary Classification Results

As part of the evaluation process, we examined the effectiveness of the initial strategy, which involved incorporating temperament scores as additional input features. Binary classification was conducted for each pair of emotional classes under two conditions: with and without the inclusion of temperament scores as input features. Figure 9 illustrates the performance differences between these two conditions. The results show that the inclusion of temperament scores consistently improves classification accuracy across all binary classifiers. The improvement is minimal for the Scary vs. Relaxing classifier, while the most significant increase in accuracy is observed in the Joyful vs. Relaxing condition.

Subsequently, we assessed the second approach, which involved adjusting the output of the emotion classifier—trained solely on physiological signals—based on individual temperament scores.

In the third approach, participants were grouped based on their temperament type, and separate classifiers were trained for each group. For the warm–cold dimension, two independent SVM classifiers were constructed: one for the warm group and one for the cold group. The overall performance was computed as the average classification accuracy across the two groups.

In addition to the warm–cold axis, the moist–dry dimension was evaluated using the same procedure. Separate classifiers were trained for moist and dry temperament groups within all three personalization strategies. However, results revealed

no difference in classification accuracy when incorporating the moist–dry dimension compared to the baseline non-personalized model. Therefore, only the results corresponding to the warm–cold dimension are reported in Fig. 9.

Figure 9 illustrates the accuracy of all three temperament-integration approaches. For binary emotion classification, integrating temperament through the first (temperament-specific models) and second (temperament-weighted) approaches generally resulted in improved accuracy, although no improvement was observed for the Scary vs. Relaxing classifier. In contrast, the third approach—based on training separate models for each temperament type—led to reduced accuracy in two classifiers (Scary vs. Boring and Joyful vs. Boring).

3.5 Multiple Classification Results

In this section, a four-class SVM classifier was employed to analyze the data. The results were 0.375, 0.388, 0.396, and 0.375 for the four conditions—without temperament, Approach I, Approach II, and Approach III, respectively. Although these values are above the chance (0.25), the overall classification accuracy remains relatively low. The observed improvement from incorporating temperament was minimal in this case. One possible reason for this outcome could be the relatively lower performance of classical classifiers, including SVM, in multi-class tasks compared to binary classification. This difference further suggests that the temperament–emotion modulation function, illustrated in Fig. 1, is not a simple linear relationship. Instead, the way temperament influences emotional appraisal may vary across different emotions, such that a particular temperamental trait may modulate the intensity or direction of response for one emotional stimulus but not for others.

Identifying this complex modulation function could play a crucial role in the development of personalized emotion recognition models, as understanding how individual characteristics, such as temperament, affect emotional processing allows for designing systems that can more accurately predict users' emotional responses.

The primary aim of this study was to demonstrate the impact of temperament on emotion recognition, which was observed even using classical classifiers. However, uncovering the precise modulation function requires the application of nonlinear and more advanced methods, which is deferred to future research.

4. Discussion

The present study investigated the role of individual temperament as a personalization factor in enhancing emotion recognition from physiological signals, specifically EDA and PPG. By evaluating three distinct approaches—(1) incorporating temperament as an input feature, (2) adjusting model weights based on temperament categories—, and (3) building separate models for each temperament type, we aimed to assess the utility of temperament-aware modeling.

Based on the SAM questionnaire scores, in which participants indicate their subjective evaluation of emotional stimuli, we observed that individuals with warmer temperaments assigned higher valence ratings specifically to fear-inducing stimuli. Previous research in temperament and emotion consistently shows that individual differences in temperament are linked to how people appraise and respond to emotionally salient events. Temperament has been conceptualized as biologically based differences in emotional reactivity and self-regulation that influence subjective and behavioral responses to affective stimuli (Rothbart & Bates, 2006). Evidence from developmental and personality research suggests that temperamental characteristics—particularly those related to fear and affective reactivity—are associated with differential appraisal of threat and emotion regulation strategies, such that individuals with higher fearfulness or sensitivity often show enhanced attention and subjective responses to aversive or threatening stimuli. For example, studies on temperament and appraisal indicate that threat appraisal mediates the effects of temperamental fear on emotional responses in stressful contexts, highlighting that temperament can shape how individuals evaluate and cope with emotionally charged situations (Talge et al., 2008) (Childs et al., 2014). Our finding that warmer-tempered individuals assigned higher valence scores specifically to fear-eliciting stimuli may reflect similar temperament-linked appraisal

processes, where certain dispositional traits bias emotional evaluation toward higher subjective intensity for negative or threatening content. At the same time, the absence of systematic temperament effects for other stimulus types aligns with previous research suggesting that temperamental modulation of emotional appraisal is often stimulus-specific and not uniformly observed across all emotional categories(Monsonet et al., 2026).

The findings provide valuable insights into the impact of incorporating temperament profiles into physiological-signal-based emotion recognition models. Our results indicate that adding temperament information substantially improved classification accuracy across several tested models, underscoring the importance of personalized, user-aware approaches in affective computing.

The observed improvements are consistent with prior research suggesting that inter-individual variability in physiological responses to emotion is strongly shaped by stable traits (Zhao et al., 2018) (Kutt et al., 2020). These gains highlight the potential of temperament-aware frameworks to overcome the limitations of "one-size-fits-all" models, which often fail to capture the nuanced ways individuals experience and express emotions physiologically. Such adaptability is particularly crucial in real-world applications and mobile contexts, where variability in emotional triggers and environmental conditions is inherently high.

Our findings indicate that incorporating temperament information as an additional input feature improved the performance of all binary classifiers. This strategy enables the model to internally capture the interactions between temperament and physiological patterns. Such an approach is particularly advantageous when dealing with limited sample sizes, where building and validating separate models for each temperament group may not be feasible. In these cases, a unified model enriched with temperament features can still leverage individual differences, ensuring that the personalization benefits are retained without the need for large datasets per subgroup.

The second approach, which involves adjusting model weights based on temperament, also improved model performance, as same as first approach.

The third method, where separate models were trained for each temperament type, yielded mixed results. Its effect on final accuracy varied across emotion classes. This variation may be due to the reduction of training data within the same temperament group.

The results from the four-class classification provide further evidence of the challenges involved in modeling complex emotional states. Their results remain relatively low compared to the improvements observed in binary classifications. The improvements did not reach statistical significance ($p > 0.05$), suggesting that the temperament–emotion modulation effect is not consistent across all emotion categories. Instead, certain temperament traits may amplify or dampen responses for specific emotional categories but not others, making the relationship more heterogeneous and complex. These findings underscore the necessity of employing more advanced, nonlinear, or deep learning methods in future work to fully capture the intricate role of temperament in multi-class emotion recognition

Previous studies have explored personalization in emotion recognition using demographic factors (e.g., age, gender), contextual cues (e.g., time of day), or cognitive traits (e.g., attention level). However, few have systematically examined temperament or mood-related traits as criteria for personalization. Our findings are consistent with initial research(Kutt et al., 2020)(Zhao et al., 2018)(Hosseini et al., 2023), who proposed that stable personality traits could be used for adaptive emotion recognition systems. However, our work is the first to apply temperament-based modeling with physiological data and compare multiple integration strategies.

However, several limitations must be acknowledged. First, only four emotional states are considered potentially limiting the application of temperament-based transformer functions to the arousal-valence space. Second, we only focused on EDA and PPG; including additional modalities like facial expression, EEG, or respiration may further enhance the performance of temperament-aware systems. Third, three strategies proposed primarily highlight the importance and

potential of incorporating temperament for enhancing emotion recognition, but they do not provide a precise model describing how temperament modulates the emotional feature vectors. Developing a detailed modulation model that captures the interaction between temperament and emotional signals is therefore considered an important direction for our future work. Such a model could provide deeper insights into the mechanisms underlying temperament-based personalization and enable more accurate and adaptive emotion recognition systems

As part of our future work, we plan to explore the reverse direction investigated in some previous studies—namely, predicting an individual's temperament from their emotional or physiological responses. Specifically, we aim to examine whether temperament can be inferred directly from physiological signals such as EDA and PPG. This approach is motivated by recent research suggesting that physiological features, including heart rate variability and electrodermal activity, can serve as informative indicators of stable personality traits such as extraversion or conscientiousness (Biswas et al., 2023)(Tseng et al., 2023).

Inferring temperament from physiological signals could provide a non-invasive, objective, and real-time method for individual profiling. Such a capability would enhance the personalization of affective computing systems, reduce reliance on subjective self-reports, and potentially enable more adaptive and responsive emotion recognition models tailored to each user's temperament.

5. Conclusion

This study provides empirical evidence for the role of temperament in enhancing emotion recognition from physiological signals, using data collected from 124 participants. EDA and PPG signals were recorded during the presentation of four distinct emotional stimuli, while participants' temperament profiles were assessed via the standardized TPM temperament questionnaire. Three personalization strategies were evaluated: training separate classifiers for each temperament group, incorporating temperament scores as additional input features, and adjusting model outputs according to temperament using a secondary SVM.

Results indicate that incorporating temperament information consistently improves binary classification accuracy, particularly in distinguishing joy versus relaxation, confirming the modulatory effect of individual differences on physiological-emotional responses.

Overall, these findings demonstrate that temperament-aware models leverage both EDA and PPG features to capture inter-individual variability, offering a data-driven, personalized approach to affective computing. Future work should explore multimodal physiological inputs, continuous temperament representations, and longitudinal data to further optimize model performance and generalizability.

Declarations

Funding: No funding was received to assist with the preparation of this manuscript.

Competing interests: The authors have no competing interests to declare that are relevant to the content of this article.

Availability of data and material: The data prepared in this study are available on request from the corresponding author.

Code availability: The code used in this study are available on request from the corresponding author.

Ethical approve: All procedures were performed in compliance with the World Medical Association Declaration of Helsinki and have been approved by the institutional ethical review committee (2024-07-02, IR.SHAHROODUT.REC.1403.016).

Consent to participate: The privacy rights of human subjects have been observed and that informed consent was obtained for experimentation with human subjects.

Declaration of generative AI and AI-assisted technologies in the writing process: Generative AI tools were used to assist with language editing and improving clarity of the manuscript text. The authors critically reviewed, edited, and approved all AI-assisted suggestions, and take full responsibility for the content and scientific integrity of the work. No AI tools were used for data analysis, statistical inference, or drawing scientific conclusions.

Authors' contribution: All authors contributed to the preparation and writing of the manuscript. They have read and approved the final version of the manuscript.

References

1. Adolph, D., & Margraf, J. (2017). The differential relationship between trait anxiety, depression, and resting frontal α -asymmetry. *Journal of Neural Transmission*, 124(3), 379–386.
2. Ahmad, Z., & Khan, N. (2022). A survey on physiological signal-based emotion recognition. *Bioengineering*, 9(11), 688.
3. Ahmadi, M., Shirafkan, H., & Mozaffarpur, S. A. (2023). Assessment of the diagnostic methods of mizaj in persian medicine: a systematic review. *Diagnostics*, 13(5), 818.
4. Alizadeh, M., Gorji, N., Ghods, R., Shirzad, M., Vaghasloo, M. A., Mokaberinejad, R., Parvizi, M. M., Parsa, E., Mortaji, Z., & Heydarirad, G. (2025). Major and Minor Criteria of Heart Mizaj (Temperament) Diagnosis from the Perspective of Persian Medicine: An Expert Panel Study. *Shiraz E-Medical Journal*, 26(26).
5. Avicenna, H. (2015). *Canon of medicine complete five volume set*. Chicago, IL: Kazi Publications.
6. Biswas, A., Li, J., Shubho, F. H., Gedeon, T., Rahman, S., & Hossain, M. Z. (2023). Personality Perception Using Scenario Based Stimulation and Physiological Signals. *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1884–1891.
7. Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49–59.
8. Buckwalter, J. G., Stanczyk, F. Z., McCleary, C. A., Bluestein, B. W., Buckwalter, D. K., Rankin, K. P., Chang, L., & Goodwin, T. M. (1999). Pregnancy, the postpartum, and steroid hormones: effects on cognition and mood. *Psychoneuroendocrinology*, 24(1), 69–84.
9. Buss, A. H., & Plomin, R. (2014). *Temperament (PLE: Emotion): Early developing personality traits*. Psychology press.
10. Can, Y. S., & Ersoy, C. (2022). Smart affect monitoring with wearables in the wild: An unobtrusive mood-aware emotion recognition system. *IEEE Transactions on Affective Computing*, 14(4), 2851–2863.
11. Canli, T., & Amin, Z. (2002). Neuroimaging of emotion and personality: Scientific evidence and ethical considerations. *Brain and Cognition*, 50(3), 414–431.
12. Childs, E., White, T. L., & de Wit, H. (2014). Personality traits modulate emotional and physiological responses to stress. *Behavioural Pharmacology*, 25(5 and 6).
https://journals.lww.com/behaviouralpharm/fulltext/2014/09000/personality_traits_modulate_emotional_and.14.aspx
13. Cloninger, C. R., Cloninger, K. M., Zwir, I., & Keltikangas-Järvinen, L. (2019). The complex genetics and biology of human temperament: a review of traditional concepts in relation to new molecular findings. *Translational Psychiatry*, 9(1), 290.
14. Cloninger, C. R., Svrakic, D. M., & Przybeck, T. R. (1993). A psychobiological model of temperament and character. *Archives of General Psychiatry*, 50(12), 975–990.
15. Gohumpu, J., Xue, M., & Bao, Y. (2023). Emotion recognition with multi-modal peripheral physiological signals. *Frontiers in Computer Science*, 5, 1264713.

16. Goldberg, D. P., & Hillier, V. F. (1979). A scaled version of the General Health Questionnaire. *Psychological Medicine*, 9(1), 139–145.
17. Goshvarpour, A., Goshvarpour, A., & Abbasi, A. (2022). A Predictive Model for Emotion Recognition Based on Individual Characteristics and Autonomic Changes. *Basic and Clinical Neuroscience*, 13(3), 285.
18. Hosseini, M. S. K., Firoozabadi, S. M., Badie, K., & Azadfallah, P. (2023). Personality-based emotion recognition using EEG signals with a CNN-LSTM network. *Brain Sciences*, 13(6), 947.
19. Hygge, S., & Knez, I. (2001). Effects of noise, heat and indoor lighting on cognitive performance and self-reported affect. *Journal of Environmental Psychology*, 21(3), 291–299.
20. Kafaee, M., Daviran, E., Heydari, S., & Gharibzadeh, S. (2024). A new model for estimating basal metabolic rate (BMR) using temperament in traditional Persian medicine (TPM). *Advances in Traditional Medicine*, 24(2), 541–551.
21. Kafaee, M., Moazen, F., & Saidi, M. (2025). The relationship between pupillary light response (PLR) and temperament in traditional Persian medicine (TPM). *Advances in Traditional Medicine*, 1–11.
22. Khaleghi, A., Shahi, K., Saidi, M., Babaee, N., Kaveh, R., & Mohammadian, A. (2024). Linear and nonlinear analysis of multimodal physiological data for affective arousal recognition. *Cognitive Neurodynamics*, 18(5), 2277–2288.
23. Kutt, K., Drążyk, D., Bobek, S., & Nalepa, G. J. (2020). Personality-based affective adaptation methods for intelligent systems. *Sensors*, 21(1), 163.
24. Li, J., & Washington, P. (2024). A comparison of personalized and generalized approaches to emotion recognition using consumer wearable devices: Machine learning study. *JMIR AI*, 3(1), e52171.
25. Li, M., Deater-Deckard, K., Calkins, S. D., & Bell, M. A. (2017). Getting to the heart of personality in early childhood: Cardiac electrophysiology and stability of temperament. *Journal of Research in Personality*, 67, 151–156.
26. McCrae, R. R., & Costa Jr, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, 52(5), 509.
27. Measures, S. for P. R. A. H. C. on E., Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Fillion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, 49(8), 1017–1034.
28. Miao, Y., Wang, H., Wang, W., & Song, L. (2025). DuMoNet: Enhancing Emotion Recognition with a Dual-Modality Structure from EEG and ECG. *2025 25th International Conference on Digital Signal Processing (DSP)*, 1–5.
29. Mojahedi, M., Mozaffarpur, S. A., Zaidi, S. M. A., & Shirafkan, H. (2023). The weight of Mizaj (temperament) indices in Persian Medicine: A Delphi study. *Caspian Journal of Internal Medicine*, 14(3), 513.
30. Mojahedi, M., Naseri, M., Majdzadeh, R., Keshavarz, M., Ebadini, M., Nazem, E., & Isfeedvajani, M. S. (2014). Reliability and validity assessment of Mizaj questionnaire: a novel self-report scale in Iranian traditional medicine. *Iranian Red Crescent Medical Journal*, 16(3), e15924.
31. Monsonet, M., Fagian-Núñez, K., Kwapil, T. R., & Barrantes-Vidal, N. (2026). Affective temperaments, self-schemas, and their interplay with emotional distress. *Frontiers in Psychiatry, Volume 16-2025*.
<https://doi.org/10.3389/fpsyt.2025.1732425>
32. Morriss, J., Taylor, A. N. W., Roesch, E. B., & van Reekum, C. M. (2013). Still feeling it: the time course of emotional recovery from an attentional perspective. *Frontiers in Human Neuroscience*, 7, 201.
33. Muller, J. (2011). *The influence of Harm Avoidance and Novelty Seeking temperament traits on emotional processing*. University of Pretoria (South Africa).
34. Poh, M.-Z., Swenson, N. C., & Picard, R. W. (2010). A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *IEEE Transactions on Biomedical Engineering*, 57(5), 1243–1252.
35. Poyraz, B. Ç., Sakallı Kani, A., Aksoy Poyraz, C., Öcek Baş, T., & Arıkan, M. K. (2017). Cognitive psychophysiological substrates of affective temperaments: a P300 study. *Clinical EEG and Neuroscience*, 48(2), 96–102.

36. Rezadoost, H., Karimi, M., & Jafari, M. (2016). Proteomics of hot-wet and cold-dry temperaments proposed in Iranian traditional medicine: a Network-based Study. *Scientific Reports*, 6(1), 30133.
37. Rothbart, M. K., & Bates, J. E. (2006). Temperament. *Handbook of Child Psychology*, 3, 99–166.
38. Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161.
39. Russell, J. A., Lewicka, M., & Niit, T. (1989). A cross-cultural study of a circumplex model of affect. *Journal of Personality and Social Psychology*, 57(5), 848.
40. Salmannezhad, H., Mojahedi, M., Ebadi, A., Mozaffarpur, S. A., Alipoor, A., Saghebi, R., & Montazeri, A. (2018). Design and validation of Mizaj identification questionnaire in Persian medicine. *Iranian Red Crescent Medical Journal*, 20(11), 9.
41. Sears, K. E., Behringer, R. R., Rasweiler, J. J. 4th, & Niswander, L. A. (2006). Development of bat flight: morphologic and molecular evolution of bat wing digits. *Proceedings of the National Academy of Sciences of the United States of America*, 103(17), 6581–6586. <https://doi.org/10.1073/pnas.0509716103>
42. Setz, C., Arnrich, B., Schumm, J., La Marca, R., Tröster, G., & Ehler, U. (2009). Discriminating stress from cognitive load using a wearable EDA device. *IEEE Transactions on Information Technology in Biomedicine*, 14(2), 410–417.
43. Shahabi, S., Hassan, Z. M., Mahdavi, M., Dezfouli, M., Rahvar, M. T., Naseri, M., Jazani, N. H., & Khalkhali, H. R. (2008). Hot and Cold natures and some parameters of neuroendocrine and immune systems in traditional Iranian medicine: a preliminary study. *The Journal of Alternative and Complementary Medicine*, 14(2), 147–156.
44. Shirbeigi, L., Zarei, A., Naghizadeh, A., & Vaghasloo, M. A. (2017). The concept of temperaments in traditional Persian medicine. *Traditional and Integrative Medicine*, 143–156.
45. Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., Xu, X., & Yang, X. (2018). A review of emotion recognition using physiological signals. *Sensors*, 18(7), 2074.
46. Sundström Poromaa, I., & Gingnell, M. (2014). Menstrual cycle influence on cognitive function and emotion processing –from a reproductive perspective. *Frontiers in Neuroscience*, 8, 380.
47. Talge, N. M., Donzella, B., & Gunnar, M. R. (2008). Fearful temperament and stress reactivity among preschool-aged children. *Infant and Child Development: An International Journal of Research and Practice*, 17(4), 427–445.
48. Tansaz, M., Akhtari, M., Naseri, M., Majdzadeh, R., Isfeedvajani, M. S., Shams-Ardakani, M. R., Nejabatbakhsh, F., Ebadi, A., & Mojahedi, M. (2023). Relationship between anthropometric indices and Mizaj (temperament) in Persian medicine. *Caspian Journal of Internal Medicine*, 14(2), 205.
49. Thomas, A., & Chess, S. (1977). *Temperament and development*. Brunner/Mazel.
50. Tseng, C.-H., Lin, H.-C. K., Huang, A. C.-W., & Lin, J.-R. (2023). Personalized programming education: Using machine learning to boost learning performance based on students' personality traits. *Cogent Education*, 10(2), 2245637.
51. Tsouli, A., Pateraki, L., Spentza, I., & Nega, C. (2017). The effect of presentation time and working memory load on emotion recognition. *J Psychol Cognition*. 2017; 2 (1): 61-66. *J Psychol Cognition 2017 Volume 2 Issue 1*, 62.
52. Vecchio, A., & De Pascalis, V. (2020). EEG resting asymmetries and frequency oscillations in approach/avoidance personality traits: a systematic review. *Symmetry*, 12(10), 1712.
53. Zhang, X., Xu, C., Xue, W., Hu, J., He, Y., & Gao, M. (2018). Emotion recognition based on multichannel physiological signals with comprehensive nonlinear processing. *Sensors*, 18(11), 3886.
54. Zhao, S., Ding, G., Han, J., & Gao, Y. (2018). Personality-Aware Personalized Emotion Recognition from Physiological Signals. *IJCAI*, 1660–1667.

Figures

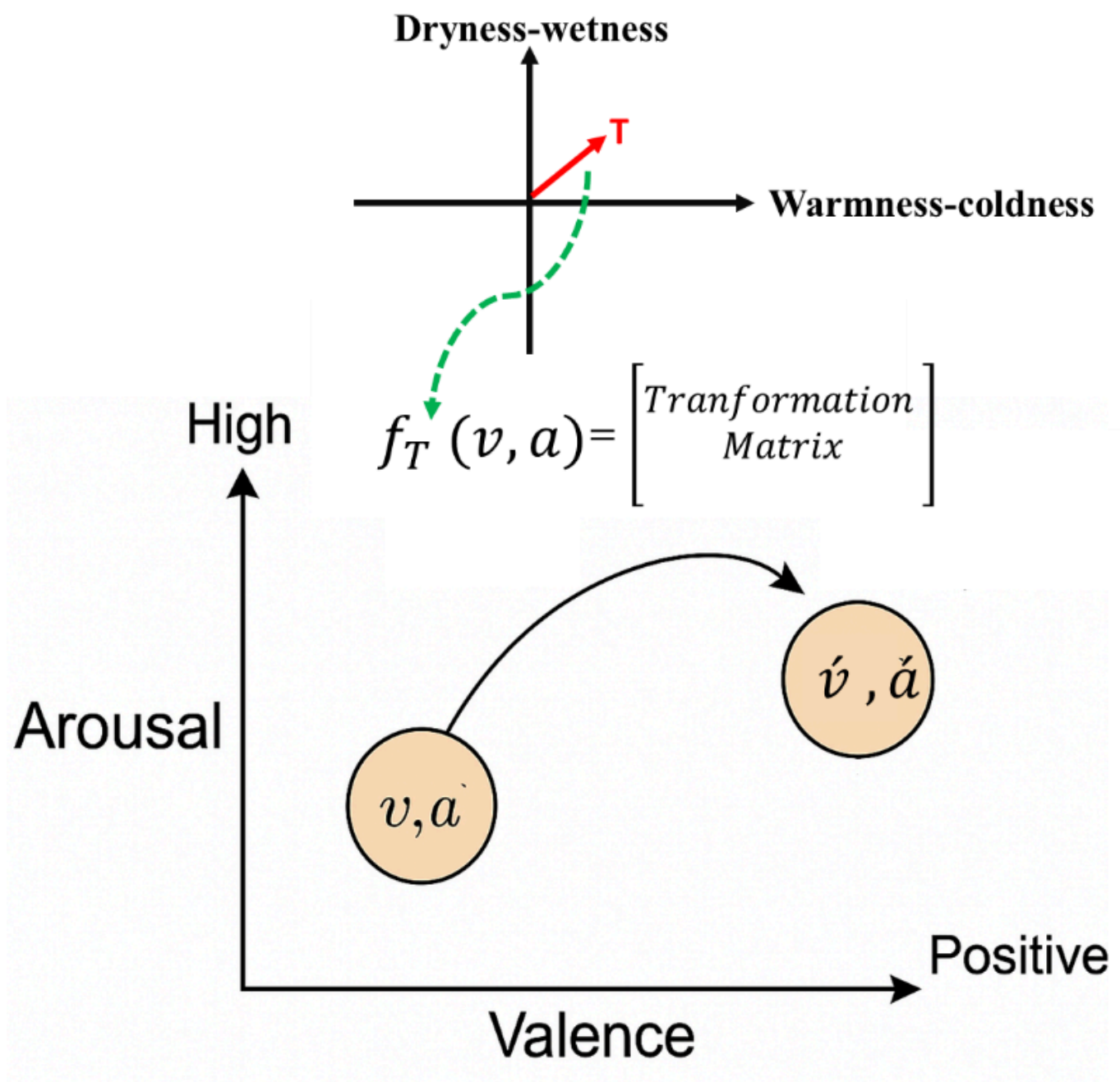


Figure 1

Temperament can be modeled as a transformer function that shifts an individual's position within the arousal–valence space.

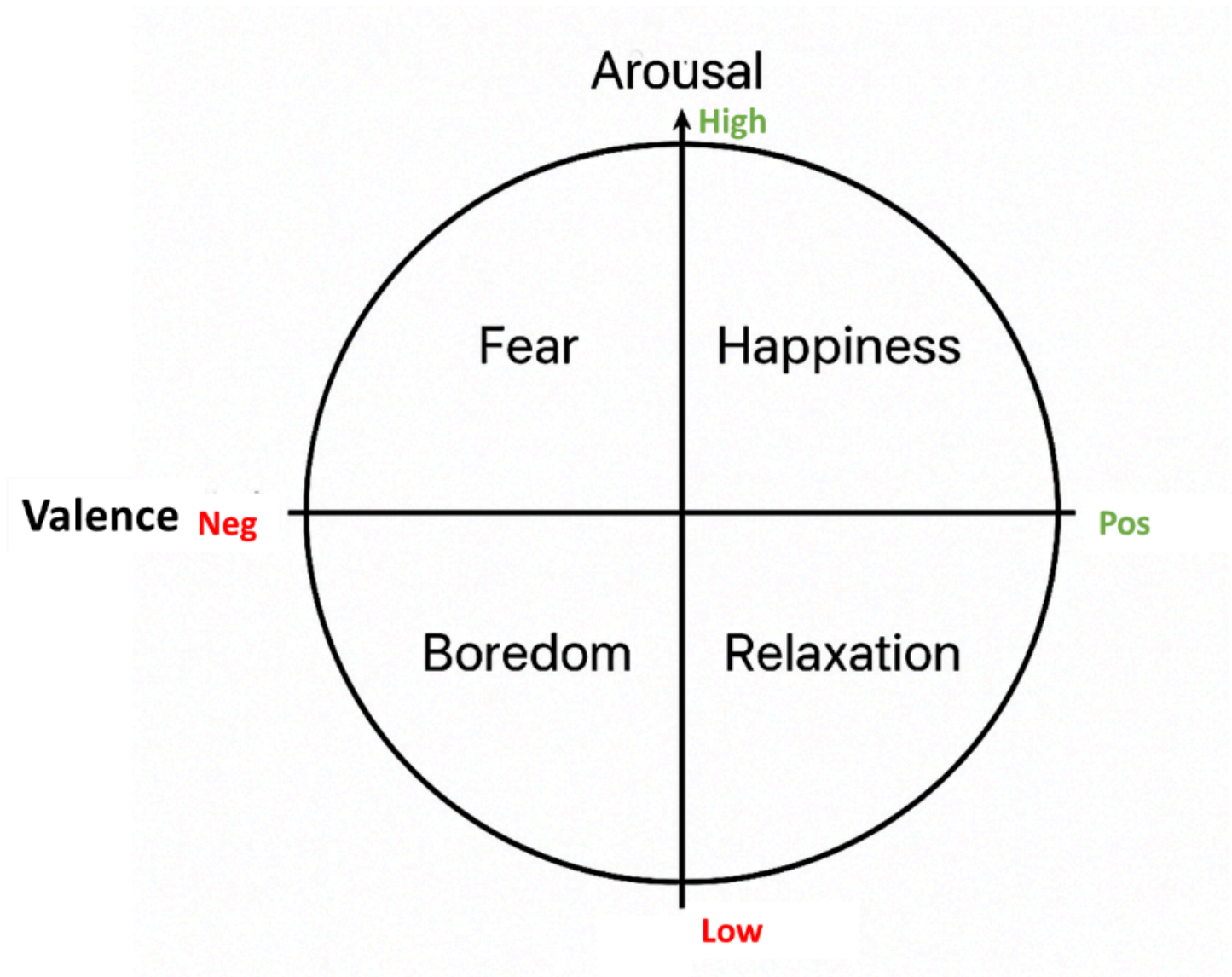


Figure 2

Circumplex model of affect, adapted from(Russell, 1980).

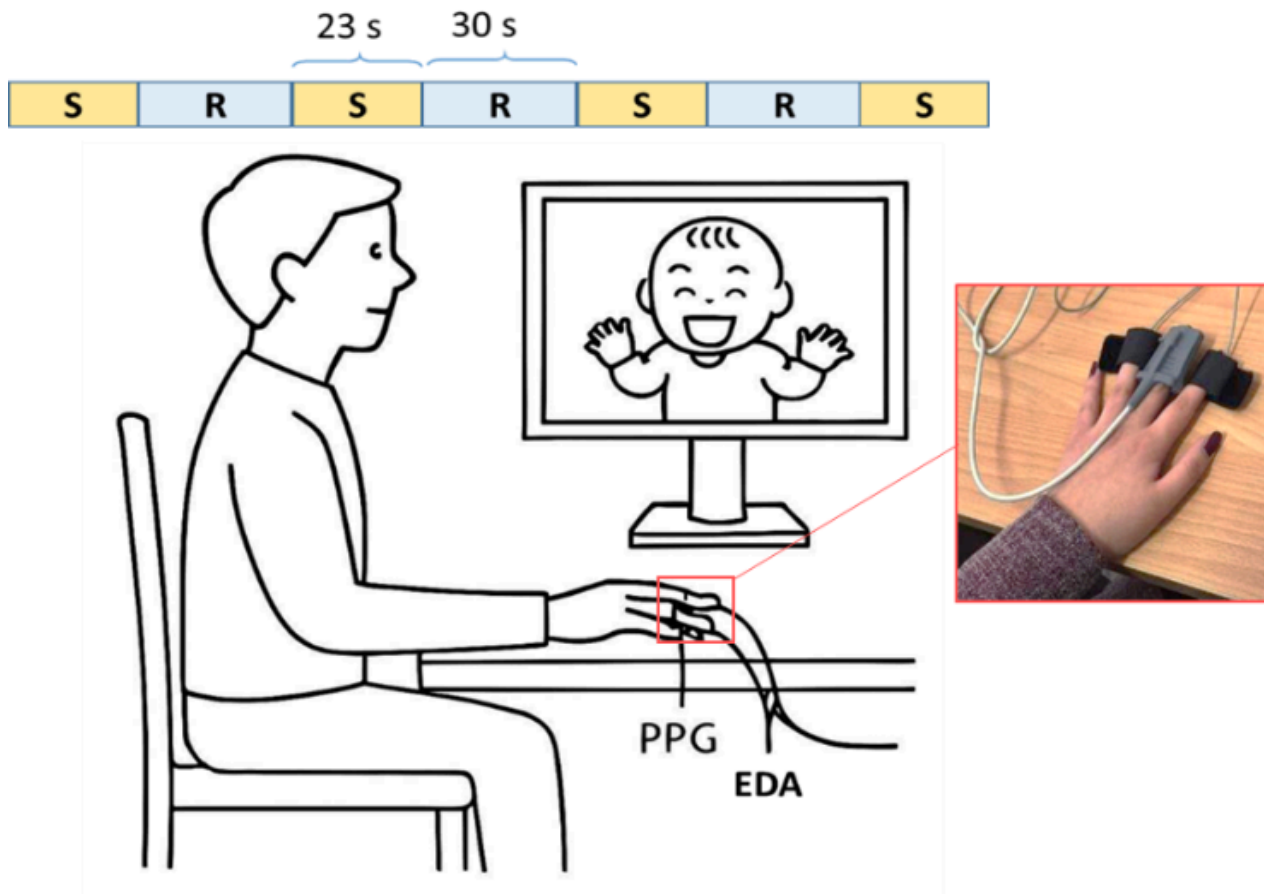


Figure 3

Experimental setup and stimulation protocol. Participants were seated in front of a monitor while PPG and EDA signals were recorded from the hand. Each emotional stimulus (S) was presented for 23 s, followed by a 30 s resting interval (R).

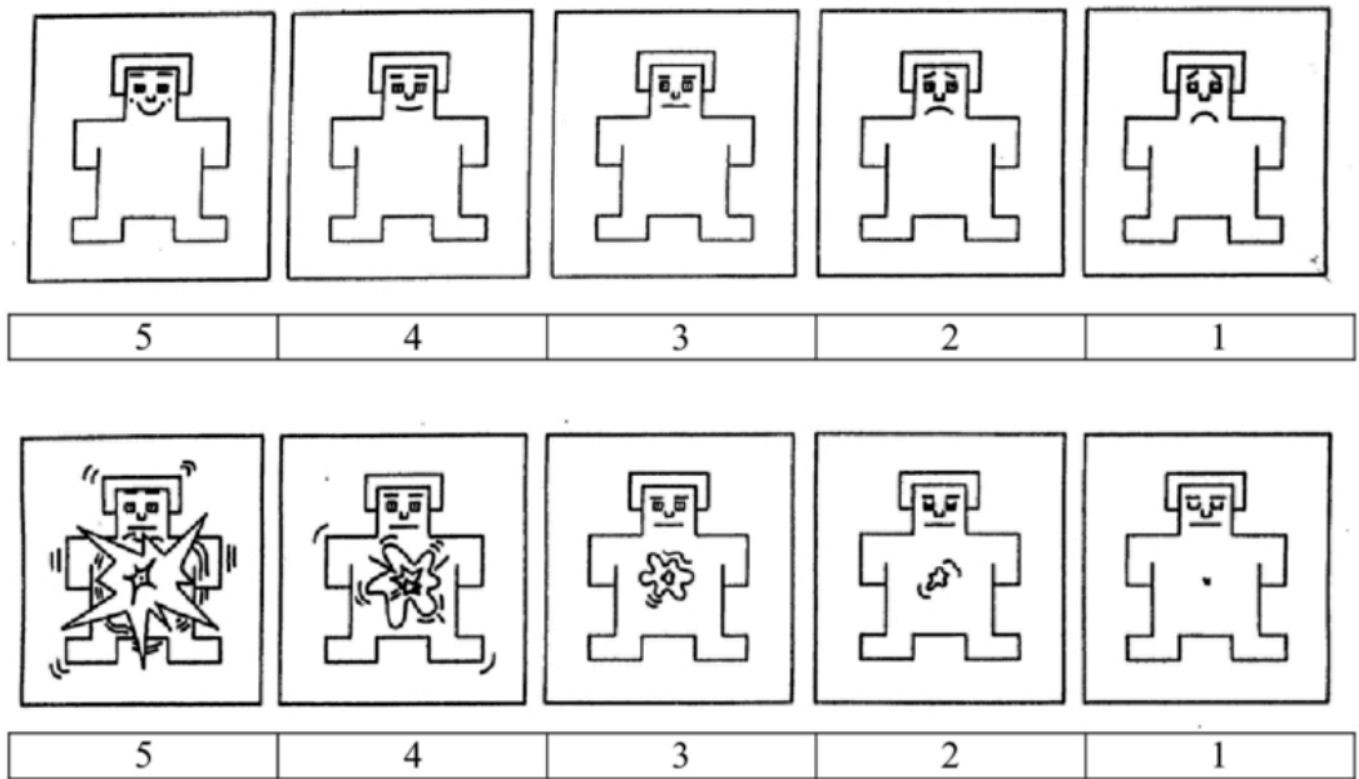


Figure 4

The Self-Assessment Manikin (SAM), a non-verbal pictorial assessment tool used to measure participants' subjective ratings of valence and arousal in response to the presented stimuli (Bradley & Lang, 1994).

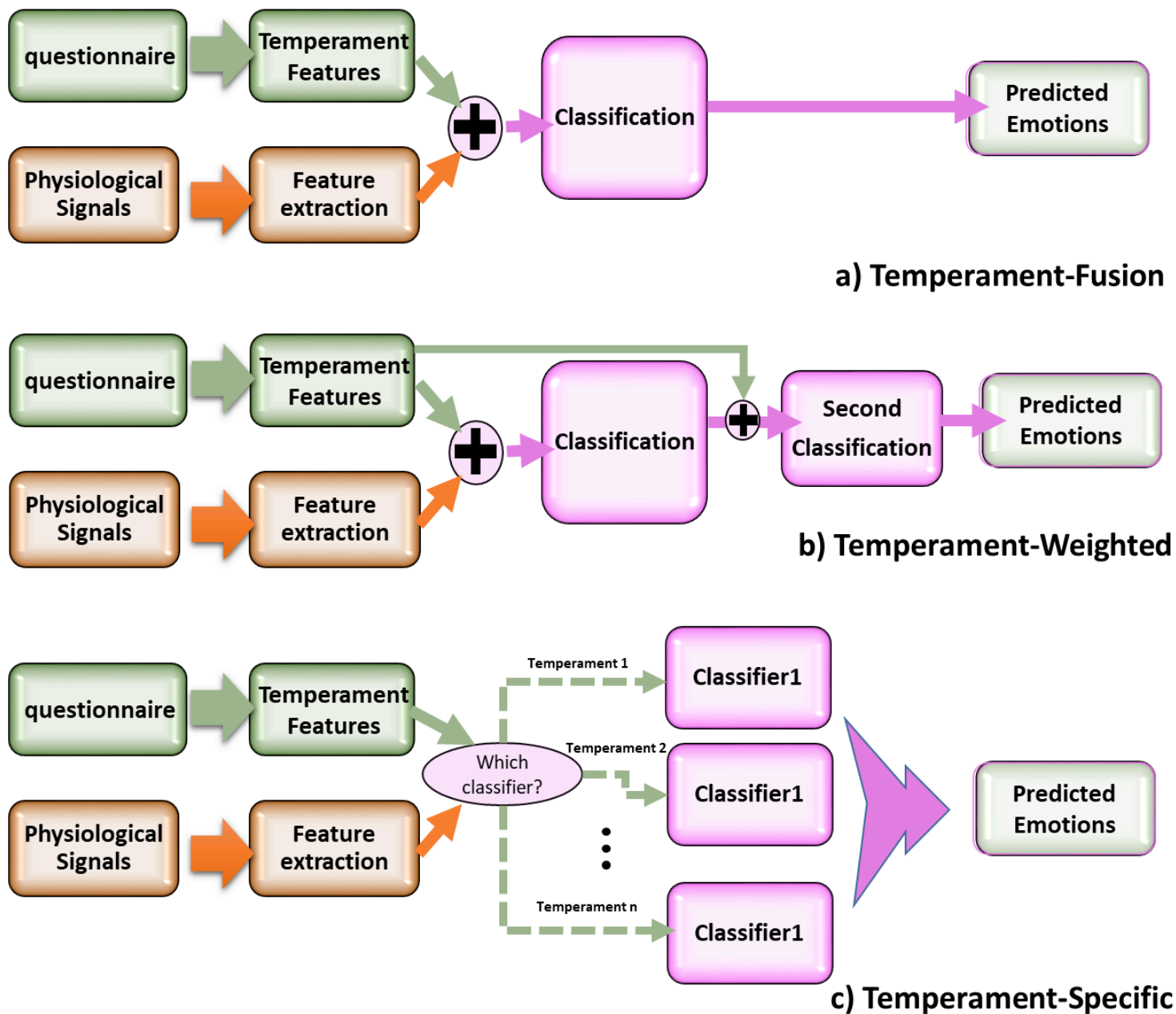


Figure 5

Schematic illustration of the three proposed strategies for integrating temperament into physiological emotion recognition models.

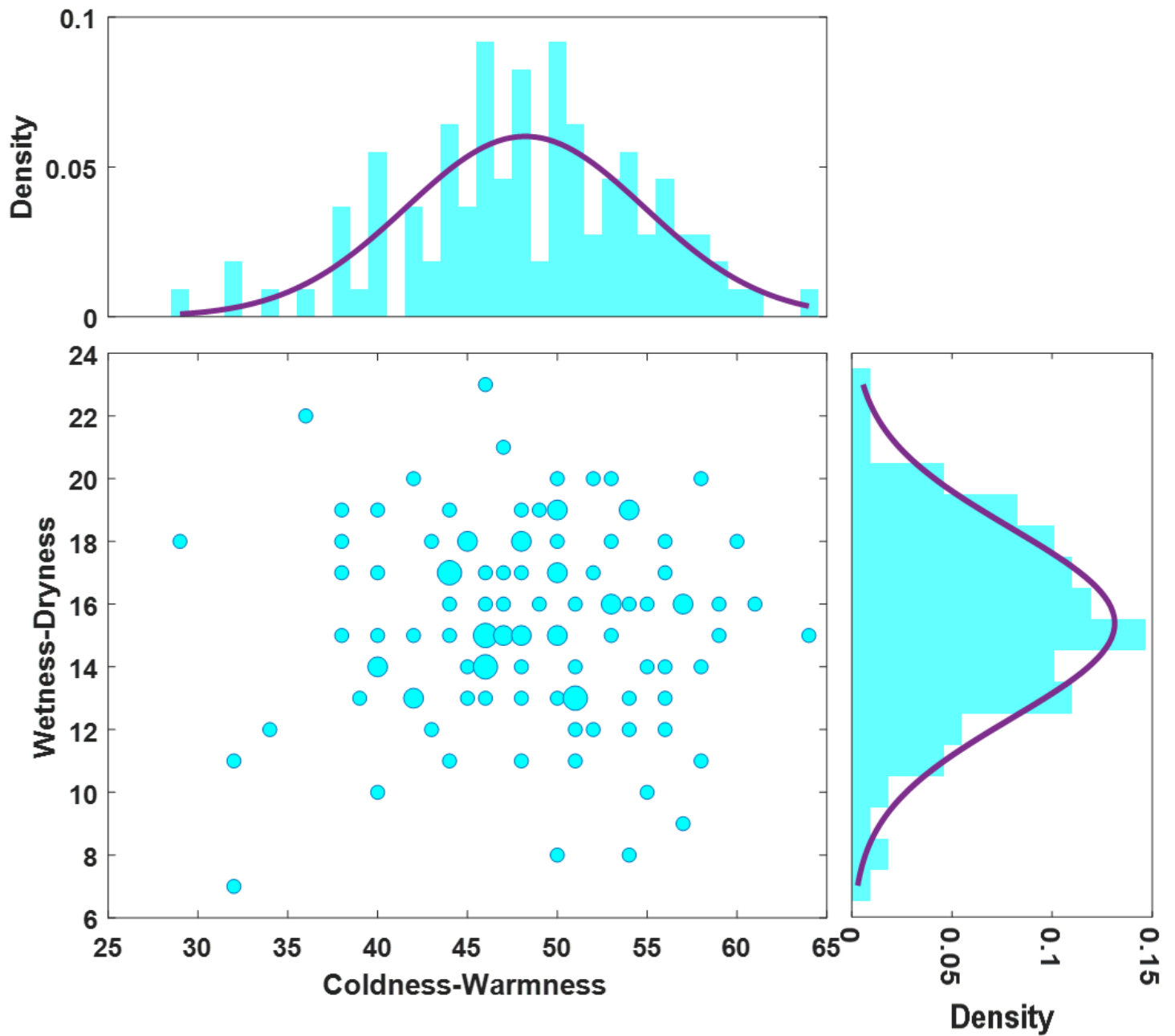


Figure 6

Distribution of participants' temperament scores along the Cold–Warm and Dry–Moist dimensions. The scatter plot shows individual data points, while the marginal histograms represent the normalized density of scores for each dimension. Kernel density estimation curves are overlaid on the histograms to highlight the overall distribution patterns.

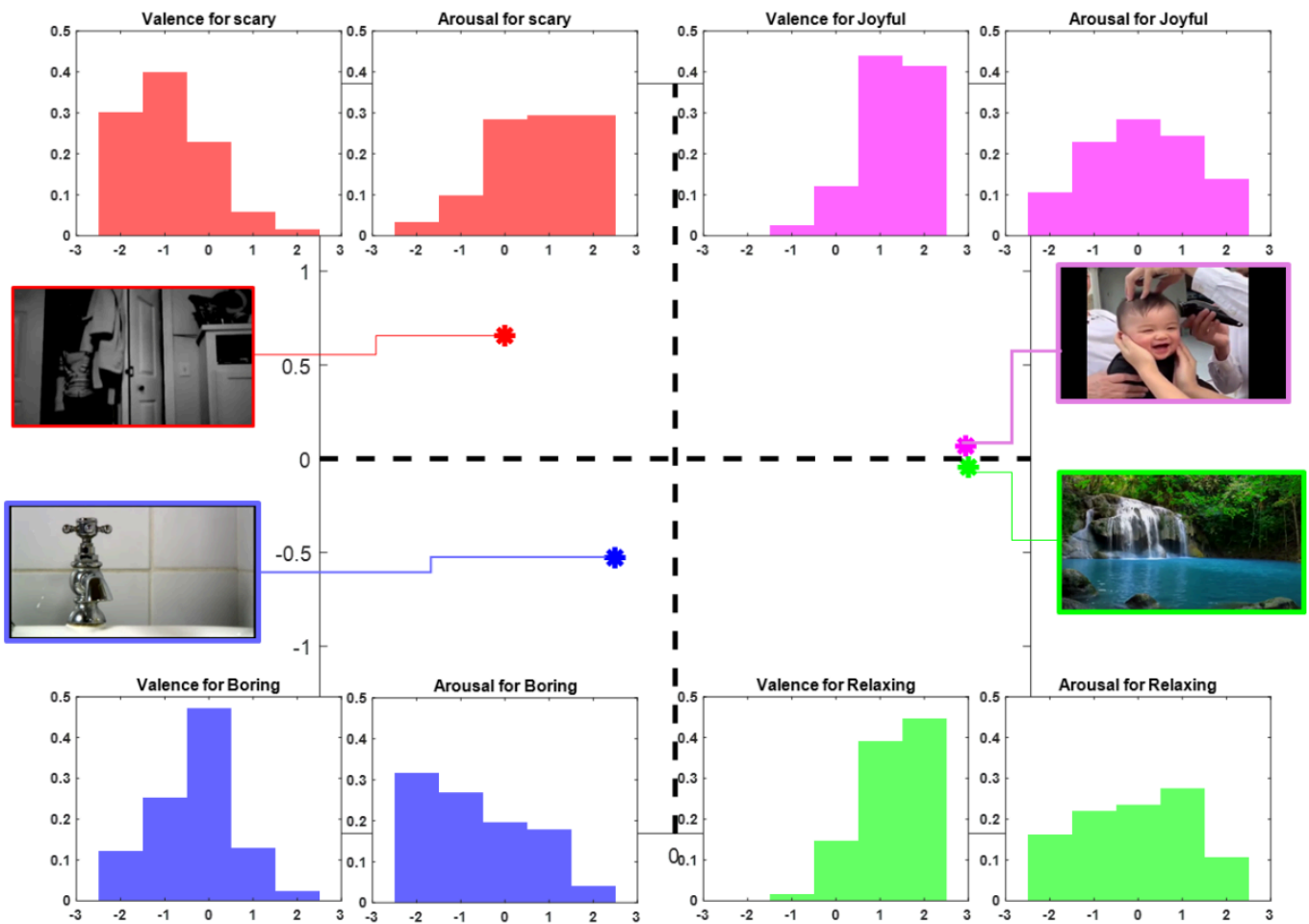


Fig 7. Distribution of stimuli in the arousal–valence space based on SAM ratings. Each colored point represents a specific stimulus ($\frac{mean}{std}$ of scores), with corresponding example frames shown in the image insets. Histograms along the axes depict the distributions of arousal and valence ratings for each stimulus category.

Figure 7

See image above for figure legend.

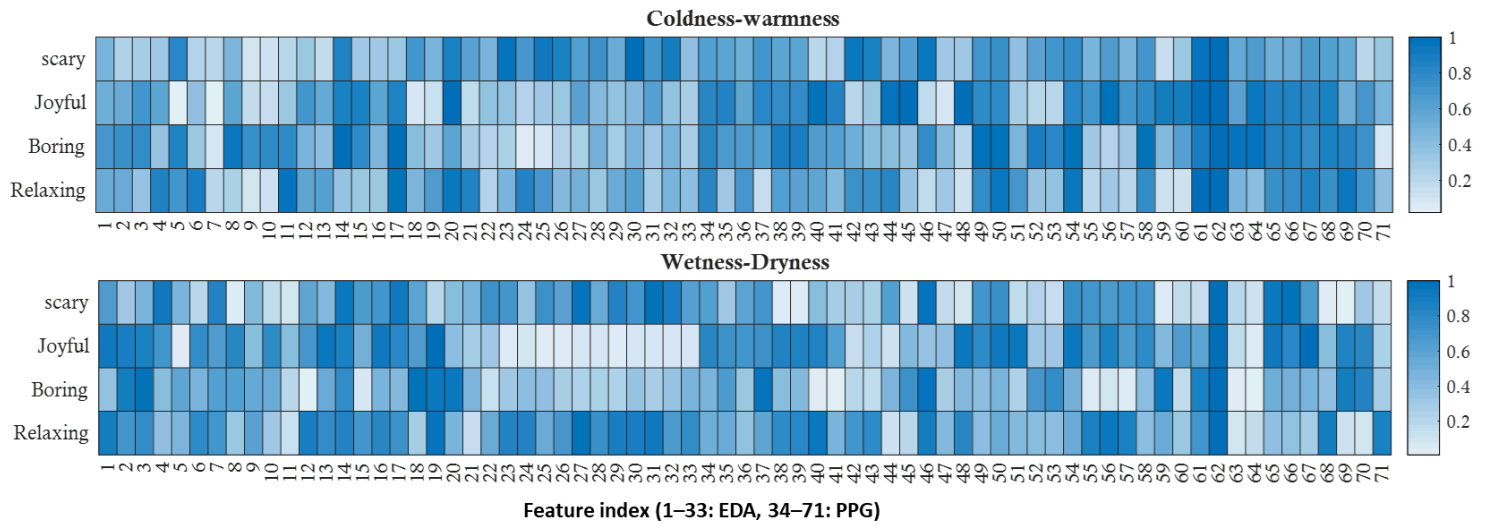


Figure 8

Heatmap of p-values for physiological signal features across coldness/warmness and wetness-dryness dimension of temperament. The analysis includes 71 features (33 from EDA, 38 from PPG) for four emotional stimuli.

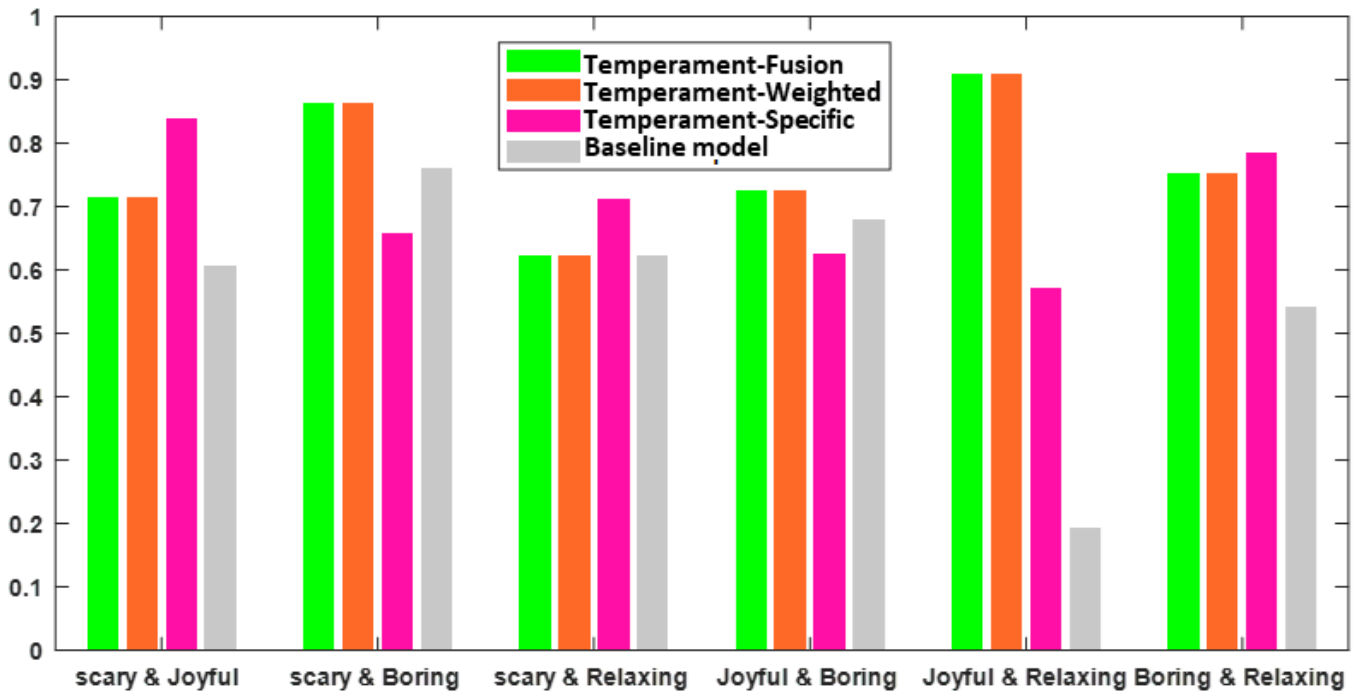


Figure 9

Classification performance of all binary emotion pairs using the baseline model (without temperament) and three temperament-aware approaches (First, Second, and Third).