Received: 26 May 2026 | Revised: 15 July 2026 | Accepted: 29 July 2026 | Published online: 13 August 2026

RESEARCH ARTICLE

Smart Wearable Technology
2026, Vol. 00(00) 1–10
DOI: 10.47852/bonviewSWT620210550

BON VIEW PUBLISHING

Real-Time Sound Mapping of Affect: A Wearable-Enabled Biometric Mapping Framework for Emotion Recognition in Music Generation

Dario Mattia[1,*] and Fabrizio Festa[1]

[1] *Department of Theory and Analysis, Composition, and Conducting, Conservatorio di Musica E.R. Duni, Italy*

Abstract: In affective computing and human–computer interaction, interpreting continuous physiological data for real-time creative applications is limited by the trade-off between bulky laboratory hardware and ergonomic commercial wearables. This paper presents a scalable node-level mapping framework refining the biometric data pipeline established by the Strings: Sounds from Human Collective Intelligence project. Grounded in Kurt Lewin's field theory and James A. Russell's circumplex model of affect, the system translates peripheral autonomic nervous system signals into continuous musical parameters. A custom wearable prototype collects electrodermal activity and heart rate variability, streaming low-latency data via UDP to Max/MSP and Ableton Live. To resolve interpersonal baselining, we introduce the Zero Point Protocol, a calibration using neutral-valence International Affective Digitized Sounds (IADS-2) stimuli to isolate participant-specific deviations from a resting floor. An affective mapping matrix then updates four-quadrant coordinates to govern tempo, harmony, and timbral modulation. For multi-user contexts, Weighted Harmonic Resolution assigns divergent signals to independent musical layers rather than destructive averaging. Bench testing confirms end-to-end operational stability with low latency; formal validation via Self-Assessment Manikin ratings and Spearman's rank correlation (ρ) against IADS-2 benchmarks is outlined as next steps. The framework provides a reproducible model for biometric sound mapping to foster nonverbal social cohesion.

Keywords: biometric sound mapping, human–computer interaction (HCI), electroacoustic music, assisted composition and performance, wearable devices

1. Introduction

Interpreting a participant's physiological signal in real time and turning it into music sits at the intersection of two problems that are usually treated separately: the technical problem of low-latency, noise-robust biosignal acquisition, and the artistic problem of designing a mapping that listeners experience as meaningful rather than arbitrary. The theoretical premises of this research are inseparable from the paradigm introduced by Picard [1], whose founding formulation established both the viability and the ethical stakes of automated emotional state recognition. This framework positions affective recognition not merely as a control interface but as a compositional substrate. Modern wearable devices make the acquisition side increasingly tractable, but the interpretive side (how to turn a noisy, individually variable physiological signal into a musically coherent and perceptually honest output) remains comparatively underdeveloped, and it is this second problem that motivates the present work. Modern wearable devices have moved beyond simple step-counting to provide a continuous stream of complex physiological data, including heart rate variability (HRV) and electrodermal activity (EDA) [2]. These signals offer a window into the autonomic nervous system (ANS), which responds directly to emotional stimuli. However, a significant gap remains between high-fidelity laboratory prototypes and the ergonomic requirements of daily use. The challenge for current human–computer interaction (HCI) research is not only in the acquisition of these data but also in their meaningful interpretation in real time while accounting for the noise inherent in mobile environments and the vast differences in participants' physiological baselines.

1.1. Problem statement: Reliability vs ergonomics

Existing research in emotion recognition often faces a trade-off. Custom-built sensor harnesses provide high-precision data but are bulky and difficult to calibrate, whereas commercial wearables offer ergonomics but lack the transparency and real-time processing required for sophisticated affective sound mapping. Furthermore, most systems rely on "absolute" biometric thresholds, failing to recognize that a state of "rest" is unique to every participant. This research addresses these challenges by building upon the established methodologies of Strings: Sounds from Human Collective Intelligence (SCHI). Originally a collective

***Corresponding author:** Dario Mattia, Department of Theory and Analysis, Composition, and Conducting, Conservatorio di Musica E.R. Duni, Italy. Email: dario.mattia@conservatoriomatera.it

biometric performance, the Strings project created a robust pipeline for real-time data filtering using Max/MSP and Ableton Live. Rather than departing from this collective focus, this research refines the constituent nodes of the framework to provide higher data resolution, ensuring a tighter coupling between collective physiology and the generation of the shared sonic environment.

1.2. Distributed affective mapping

By fostering a deeper understanding of a single node's affective mirroring, we can more accurately model how these individual nodes integrate into a harmonious collective consciousness during group activities. By repurposing the custom wearable prototype (CWP) from the Strings project, the system utilizes a Distributed Affective Mapping logic, treating each participant as a high-fidelity data source. This computational engine handles low-latency physiological data, allowing the research to focus on the psychological accuracy of the mapping within a group context. Within this loop, the soloist musician acts as a catalyst, utilizing real-time affective data to inform improvisational choices that drive the audience's psychophysiological state. Conducted within the lab team (PhD) of the Conservatory of Music E.R. Duni, Matera (Italy), this research is designed to produce both a technical framework and a tangible artistic output, culminating in a performance where technology facilitates new forms of collective expression.

1.3. Research objectives and proposed framework

The motivation for this work is twofold. Scientifically, current affective-wearable research rarely addresses the specific demands of live, multi-participant musical performance: low latency, individualized baselining, and a mapping logic that must remain perceptually coherent under real-world motion and noise. Artistically, most parameter-mapping sonification systems treat physiological data as a single control stream, which discards the individuality of each participant's contribution once several people are monitored together. The primary objective of this paper is to propose a scalable framework for node-level mapping. It is important to note that node-level mapping is not the end goal of this system but the necessary precision layer that enables accurate collective affective sound mapping. This system departs from traditional categorical models of emotion in favor of the Arousal-Valence Dimensional Model, which is more compatible with the continuous output of biometric sensors. Building on this objective, the paper makes three concrete contributions. First, it introduces the Zero Point Protocol, a standardized baseline calibration method that accounts for participant-specific physiological differences by measuring the relative change (Δ) from a controlled resting state. Second, it defines a structured Biometric-to-Musical Mapping logic for translating physiological shifts into real-time sound mapping, providing an intuitive "sonic mirror" of affective states. Third, it outlines a Transition Strategy, a roadmap for moving this logic from custom hardware to commercial wearables, so that the framework can remain both scientifically valid and ergonomically viable at ensemble scale.

Relative to prior parameter-mapping sonification and biofeedback-performance systems, the proposed approach offers four practical advantages: (I) it is noninvasive and built on low-cost, off-the-shelf-compatible hardware, keeping the barrier to ensemble-scale deployment low; (II) the Zero Point Protocol removes reliance on population-level physiological thresholds, which is the main source of error in prior absolute-threshold systems; (III) the Weighted Harmonic Resolution logic scales to multiple simultaneous participants without collapsing their signals into an averaged, affectively flattened output; and (IV) the framework is explicitly dual-grounded in psychophysiology for scientific accountability and in music-psychology/sonification theory for artistic coherence, rather than optimizing for only one of the two. These advantages are conceptual and technical at this stage; the empirical demonstration of the resulting gains in mapping accuracy is addressed in Section 5.

The remainder of the paper moves from theoretical foundations (Section 2) and the Strings case study (Section 3), through the core methodological framework and its comparative positioning (Section 4), to preliminary pilot testing, the proposed validation protocol, and ethical considerations (Section 5), before concluding with future directions (Section 6).

2. Theoretical Foundations: Affect, Physiology, and Music

While the primary focus of this work is node-level mapping, the system's macroscopic feedback loop is conceptually grounded in Kurt Lewin's field theory [3]. Lewin modeled human interaction using the formula $B = f(P, E)$, where Behavior (B) is a dynamic function of the Person (P) interacting with their Environment (E). This framework directly translates that psychological abstraction into a closed-loop digital architecture by operationalizing those variables within the software engine. The Person (P) is captured via the participant's internal state through peripheral biometric streams like the relative changes in GSR and HRV. This input interacts with the Environment (E), which is represented by the shared, generative sonic landscape output by a software chain acting as the immediate external stimulus. The resulting Behavior (B) is expressed not as physical movement but as the computational output of the entire system, manifest in the real-time musical variations, harmonic resolutions, and structural tension shifts generated by the Max/MSP mapping engine. By sound mapping physiology this way, the framework converts an abstract psychological field into an interactive loop where the participant's physiological shifts alter the sonic environment, which immediately loops back to influence their subsequent psychophysiological state.

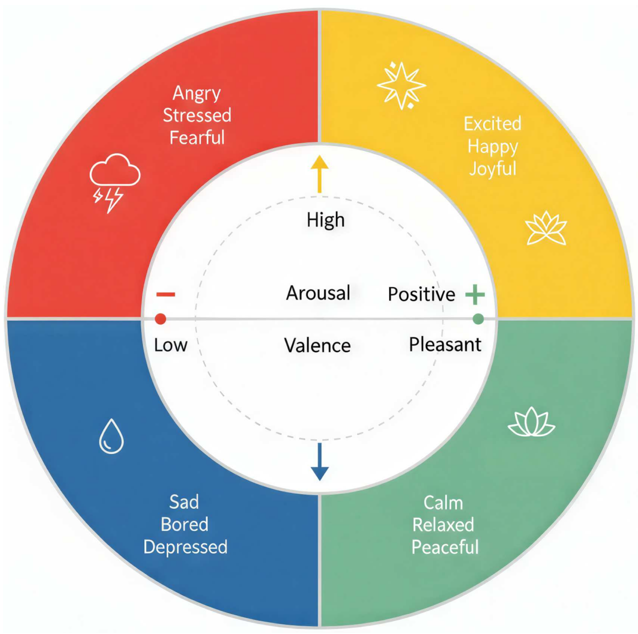
2.1. Dimensional models of affect

For real-time emotion recognition systems, the fundamental choice of an emotion model dictates the system's architecture. Two dominant theoretical approaches exist: discrete (categorical) models and dimensional models.

Discrete models, such as Paul Ekman's theory of basic emotions [4], posit that humans possess a small set of innate, universal categories: anger, disgust, fear, happiness, sadness, and surprise. While psychologically significant, this categorical view is poorly suited for real-time HCI. Biometric data—such as galvanic skin response (GSR) or HRV—are continuous, fluctuating waveforms. Translating a nonspecific voltage change into a rigid category like "Surprise" in real time is computationally intensive and often ignores the vast range of blended emotional experiences.

In contrast, dimensional models of affect conceptualize emotion as coordinates within a continuous, low-dimensional space. The primary framework adopted in this research is the Arousal-Valence model (see Figure 1) developed by Russell [5].

Figure 1
The Arousal-Valence model



This model defines emotional states along two fundamental, independent axes:

1) Arousal (Y-axis): Refers to the intensity or activation level of the emotion, ranging from calmness to excitement. This axis is strongly correlated with the sympathetic nervous system, making it directly measurable via GSR.
2) Valence (X-axis): Refers to the hedonic tone of the emotion, ranging from unpleasant/negative to pleasant/positive.

By adopting this model, the system avoids forcing analog physiological changes into rigid labels. Instead, any complex affective state is represented by a coordinate (V, A), which can be continuously updated. This allows the system to partition the affective space into four measurable quadrants for structured sound mapping as shown in Table 1.

2.2. Psychophysiology of the individual node

To bridge the Arousal-Valence model with a wearable-enabled system, we rely on psychophysiological principles demonstrating how the ANS manifests in peripheral signals. Analyzing the psychophysiology of the individual node ensures that the global collective sound mapping is derived from scientifically grounded local data. The ANS regulates involuntary bodily functions through two competing branches:

1) The Sympathetic Nervous System (SNS): The "fight-or-flight" system that prepares the body for action, increasing heart rate and perspiration. SNS activity is the primary driver of the Arousal dimension.
2) The Parasympathetic Nervous System (PNS): The "rest-and-digest" system that conserves energy and promotes relaxation.

The wearable framework (Figure 2) captures the activity of these branches using two primary biometric streams:

1) Electrodermal Activity (GSR/EDA): Because sweat glands are innervated almost exclusively by the SNS, skin conductance provides a near-direct measure of sympathetic activation. However, because GSR is also influenced by non-emotional factors such as ambient temperature and skin conductance baselines, immediate baseline tracking and controlled environments are essential for accurate interpretation [6, 7].
2) Heart Rate (BPM) and Heart Rate Variability (HRV): Cardiac activity is under the constant, opposing control of both the SNS and PNS. Mean heart rate generally reflects arousal, while HRV serves as a direct index of ANS balance. High HRV typically signifies dominant parasympathetic influence and emotional regulation, providing a proxy for determining valence. While valence remains notoriously difficult to isolate from peripheral biometrics alone, tracking relative departures in HRV provides a reliable index of the node's emotional regulation capacity, which serves as a functional proxy for systemic valence within our mapping engine [8, 9].

This proxy relationship should be read with caution. The affective computing literature is explicit that no single peripheral signal, and no fixed combination of them, offers a validated one-to-one mapping onto subjective valence; HRV-based estimates are correlational and probabilistic rather than diagnostic [9]. Within this framework, HRV is therefore treated strictly as one input among several driving a designed, artistic mapping rather than as a certified measurement of the participant's emotional state, and any claim of "accuracy" refers to perceptual coherence with self-report, not to physiological ground truth.

2.3. Musical correlates of emotional perception

To complete the pipeline, abstract psychological coordinates must be mapped to perceivable musical parameters (Table 2). This mapping is based on established relationships in music psychology that ensure the audio output is perceptually coherent with the participant's measured state:

By establishing this systematic relationship, technical biometric outputs are translated into specific controls within the

Table 1
Quadrants for sound mapping based on the Arousal-Valence model

Quadrant	Valence	Arousal	Example emotions	Musical parameters
I (Yellow)	Positive	High	Excitement, joy	Mid-fast tempo, consonant, high complexity
II (Red)	Negative	High	Stress, anger	Fast tempo, dissonant, high dynamics
III (Blue)	Negative	Low	Sadness, boredom	Mid-slow tempo, dissonant, low energy
IV (Green)	Positive	Low	Calmness, relaxation	Slow tempo, consonant, soft timbre

Figure 2
GSR and heart rate/HRV representation of physiological response pathways

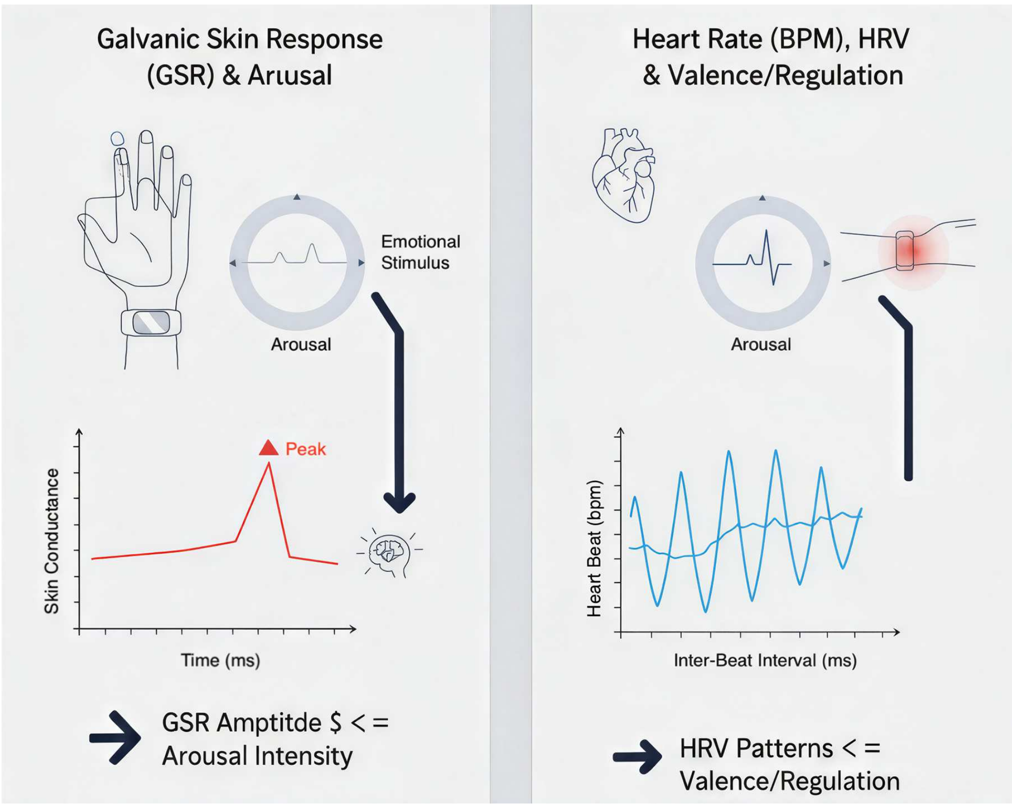


Table 2
Musical parameters and participants' state correlation

Arousal-Valence state	Dominant musical features	Examples of sound mapping
High arousal (intensity)	Fast tempo, high loudness, dense texture, sharp onsets, complex rhythms, high-frequency content	High GSR will drive an increase in the MIDI tempo and the spectral centroid of oscillators
Low arousal (calmness)	Slow tempo, low loudness, sparse texture, smooth articulation, simple rhythms, low-frequency content	Low GSR will trigger a reduction in overall sound density and the application of low-pass filters
Positive valence (pleasantness)	Major/consonant harmony, ascending melodies, bright timbre, predictable rhythm	Specific HRV patterns (suggesting good regulation) will select musical mode and consonant intervals
Negative valence (unpleasantness)	Minor/dissonant harmony, descending/disjunct melodies, rough/noisy timbre, unpredictable harmonies	Patterns suggesting stress (low HRV, high GSR) will select minor or atonal harmonies and increase timbral modulation (roughness)

Max/MSP and Ableton Live architecture. This ensures the generated music functions as a "sonic mirror," providing an intuitive representation of the node's internal state within the broader collective. This approach follows the parameter-mapping paradigm formally theorized by Hermann et al. [10]. Parameter mapping allows for designerly correspondences between data dimensions (like GSR) and acoustic parameters (like tempo), providing the expressive flexibility required for artistic output [10–13].

It should be made explicit that the specific correspondences in Table 2 (e.g., GSR driving spectral centroid, HRV selecting mode) are rule-based and expert-driven: they were defined through composer practice within the research team, informed by the parameter-mapping sonification literature [10, 14] rather than derived from a dedicated listener study. This is consistent with common practice in interactive-music mapping design, where perceptual plausibility is established through established music-psychology correspondences (e.g., tempo-arousal, mode-valence) before being tested empirically. A systematic, listener-based validation of these specific correspondences, rather than of the framework's technical accuracy alone, is identified as a necessary next step and is discussed as a limitation in Section 5.

3. Case Study: The Strings Biometric Processing Framework

The affective sound mapping system is built upon the hardware and data processing methodologies developed for Strings: SCHI. This project serves as a crucial case study, providing a tested, real-time framework for the noninvasive acquisition, filtering, and mapping of peripheral biometric data to sound. Within this research, the original Strings setup is referred to as the custom wearable prototype (CWP).

3.1. Custom wearable system architecture

The architecture of the CWP is situated within a taxonomy of digital musical instruments distinguishing sensor input, mapping, and sound synthesis subsystems. It addresses the central tension identified by Tanaka [15]: the struggle between the intimacy of physiological data and its inherent instability as a control signal. The CWP was designed to capture physiological activation from multiple participants simultaneously, creating an architecture in which robust data streams are leveraged for collective emotion recognition. The hardware layer utilizes GSR and pulse sensors connected to an ESP32 microcontroller. Serving as the embedded processing unit, the ESP32 is integrated into a wearable fabric harness to ensure participant mobility. This design choice is consistent with recent wearable biosensing approaches for embodied musical performance [16, 17] and with systematic evidence that wearable sensors are increasingly central to music-experience research [18].

The software architecture (Figure 3) follows a low-latency pipeline:

1) Firmware (Arduino IDE): Handles initial analog-to-digital conversion and basic signal conditioning.
2) Transmission (UDP): Data are streamed wirelessly to a central computer via User Datagram Protocol, prioritizing transmission speed essential for real-time musical performance (Anwar & Malik, 2022).
3) Computational Engine (Max/MSP): Receives and interpolates raw biometric data, extracting the statistical features required for mapping.
4) Synthesis (OSC and Ableton Live): Control parameters are routed via Open Sound Control to Ableton Live, which manages the digital oscillators, MIDI tracks, and effects chains [19].

3.2. Data reliability, processing, and commercial transition

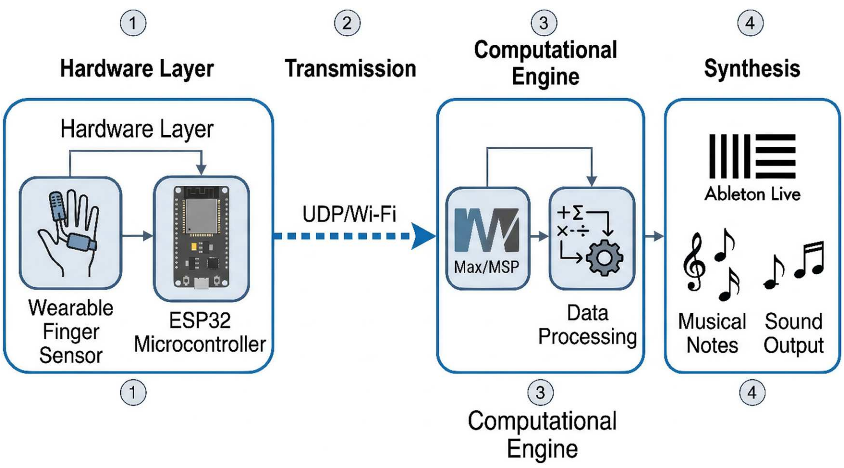
A central contribution of the Strings project is establishing a robust protocol for data reliability in live performance environments. To counter motion artifacts and environmental noise, the system employs a multi-layered filtering approach. At the hardware level, the ESP32 applies moving average filters to damp extreme electrical outliers. Within Max/MSP, a system of real-time interpolation filters ensures signal continuity, preventing the abrupt, harsh sonic clicks that result from momentary hardware dropouts.

Crucially, the system transitions from raw sensor values to statistically meaningful features extracted over rolling time windows (e.g., 5–15 s). Metrics such as the rolling mean, range, and standard deviation characterize the overall stability and intensity of the node's physiological signal, providing the refined inputs necessary for the affective mapping matrix.

While this custom prototype successfully validated the feasibility of real-time sound mapping, the next phase of this research focuses on upgrading scalability and participant ergonomics via Commercial Off-The-Shelf wearables (e.g., smartwatches with built-in EDA and heart rate hardware). This direction is consistent with the broader trend toward flexible, low-drift wearable sensing materials; for instance, Lei et al. [20] demonstrate an MXene/laser-induced-graphene strain sensor with rapid response and low signal hysteresis for continuous body-motion monitoring, illustrating the kind of motion-robust sensing substrate that could reduce the artifact burden currently handled in software by the CWP's filtering stage. This strategic transition solves several deployment challenges:

1) Ensemble Scalability: Standardized sensor packages significantly lower the setup and calibration friction for large-ensemble monitoring, making multi-participant orchestration viable [21].

Figure 3
SCHI pipeline architecture



2) Data Fidelity: Commercial devices provide internal, preprocessed metrics for HRV that are far more accurate than those produced by unshielded custom pulse sensors prone to physical drift [22].
3) Focus on Mapping Logic: By utilizing public Application Programming Interfaces (APIs) for standardized data ingestion, the research can prioritize refining the Max/MSP affective mapping logic rather than managing physical hardware fabrication and ongoing bench maintenance.

It is important to note that while the CWP was fundamental for algorithmic validation, its precision in mapping Valence via HRV was limited by motion artifacts. Consequently, the current framework treats CWP-derived HRV data as a "relative variability proxy" rather than an absolute metric. To achieve clinical-grade data fidelity, future work will involve the integration of medical-grade sensor platforms, such as those provided by Shimmer. This technical transition aims to eliminate the signal noise inherent in the prototype, ensuring the highest possible accuracy for the affective mapping matrix [23, 24].

4. Methodological Advancement: Affective Mapping

The transition from physiological measurement to emotional representation requires a structured mapping framework that accounts for the inherent variability of human biology [25]. This section details the logic used to translate biometric signals into coordinates within the Arousal-Valence model and, subsequently, into real-time musical parameters for the collective engine.

4.1. Node-specific baseline calibration

A significant challenge in wearable-based emotion recognition is the "baselining" problem: a high heart rate for one participant may represent a resting state for another. To ensure accuracy across the ensemble, the framework utilizes a Standardized Baseline Calibration Protocol (Figure 4). Rather than assuming universal physiological constants, the methodology requires a pre-performance session where participants are placed in identical environmental conditions, involving neutral lighting and a controlled seated position. The Zero Point Protocol extends the methodology of Yuksel et al. [14], who first demonstrated the importance of adaptive calibration to a single user's individual cognitive and physiological baseline rather than population norms [26]. To further standardize this "Zero Point of Restfulness," the protocol includes the playback of neutral-valence auditory stimuli from the International Affective Digitized Sounds (IADS-2) database. This ensures that the initial baseline is a controlled response to a neutral acoustic environment rather than mere silence, increasing the consistency of the relative change (Δ) measurements across different nodes. Additionally, the controlled calibration environment serves to stabilize ambient temperature, mitigating it as a confound for GSR readings. During this period, Max/MSP captures the baseline values for both GSR and heart rate. This relative thresholding allows the system to distinguish between a node's natural physiological floor and the fluctuations caused by active emotional contribution.

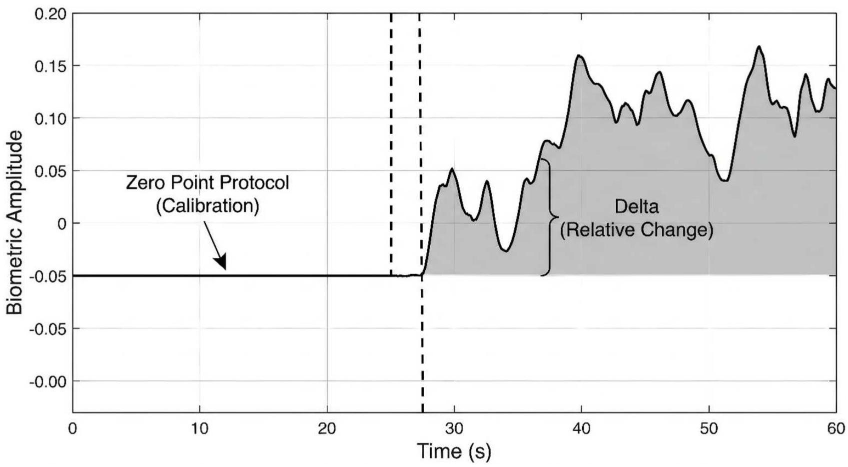
4.2. The biometric-to-affective state mapping matrix

The central logic of the system is the affective mapping matrix, which functions as the bridge between physiology and psychology. This matrix uses deviations from the established Zero Point to determine a participant's position within the four quadrants of the Arousal-Valence space. The logic is divided into two primary streams:

1) Arousal Mapping (Y-axis): Derived from GSR amplitude. When skin conductance rises significantly above the node's Zero Point, it indicates sympathetic activation, driving the coordinate into the High Arousal quadrants.
2) Valence Mapping (X-axis): Determined through the relationship between heart rate (BPM) and HRV proxies. While high HRV generally suggests parasympathetic regulation, this mapping is framed as probabilistic rather than deterministic; high HRV increases the likelihood of positive valence within the system logic but does not serve as a biological confirmation.

To mitigate sensor noise and motion artifacts inherent in custom sensors, the calculation utilizes the relative deviation from the Zero Point rather than instantaneous raw values. By processing these inputs simultaneously, Max/MSP identifies a dominant quadrant at a high refresh rate, allowing for a continuous, "smooth" emotional trajectory that prevents jarring musical jumps.

Figure 4
Relative changes from the Zero Point diagram



4.3. Affective sound mapping design

Once the affective state is identified, the coordinates are transmitted via OSC to Ableton Live. The sound mapping is governed by Musical Mapping Rules that ensure the audio output is perceptually aligned with the Arousal-Valence model. In Quadrant I (High Arousal, Positive Valence), characterized by excitement, the system increases the MIDI tempo and shifts the musical mode to a major scale while increasing the spectral brightness of the synthesizers. In Quadrant II (High Arousal, Negative Valence), representing stress or anger, the tempo remains high, but the harmony shifts to dissonant intervals and timbral "roughness" is introduced through frequency modulation.

For lower-activation states, Quadrant III (Low Arousal, Negative Valence) utilizes slow, descending melodic patterns and darker, low-pass filtered timbres to represent sadness or boredom. Finally, Quadrant IV (Low Arousal, Positive Valence) focuses on consonant, ambient textures with minimal rhythmic movement and soft attacks to map through sound states of calmness and relaxation. This structured design ensures that the music acts as a real-time "mirror" of the participant's affective state, providing an intuitive and immersive sonic representation of the node's internal physiology.

4.4. Collective state resolution

Resolving several simultaneous affective states into one musical output is not only a signal-processing problem: it is also a question of what it means for a group to share an experience without collapsing individual participants into an undifferentiated average. The theoretical foundation for this resolution is provided by Lévy's [27] seminal formulation of collective intelligence. For Lévy, collective intelligence is neither the sum of individual intelligences nor their average: it is the dynamic, networked configuration of distinct subjective contributions, each preserving its specificity while participating in a larger, emergent whole. The proposed architecture instantiates this structure at the level of physiological data, networking distinct individual signals into a collaborative whole rather than flattening them into uniformity. Where Donnarumma's [28] Xth Sense pushed the body-as-instrument logic to its limit for a single performer, this framework scales that same expressive logic to a collective, addressing inter-individual synchronization through the Weighted Harmonic Resolution logic described below [27–29].

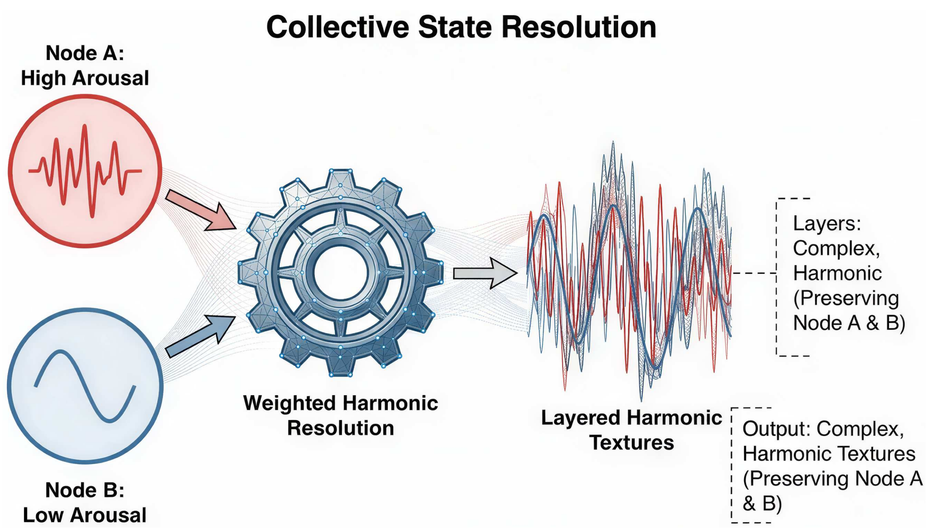
In fact, when multiple nodes are processed simultaneously, the system must resolve divergent affective states into a unified musical output through a Weighted Harmonic Resolution logic (Figure 5). Within the proposed framework, sudden arousal peaks from a highly reactive node can dynamically open global filter coefficients or increase structural tension across the entire mix, while lower-activation nodes simultaneously govern background ambient layers, spatialization density, or sparse rhythmic sub-structures. By preserving this physiological polyphony, the generative audio maps through sound the authentic friction and complexity of the collective field, treating the ensemble as a dynamic web of shared expression rather than a compromised mathematical average [30, 31].

Indeed, this resolution framework represents a proposed design principle to be empirically evaluated in future multi-participant performance studies. Rather than simply averaging data coordinates—which risks neutralizing the emotional impact into a flattened, unexpressive middle ground—the system is designed to assign divergent data streams to independent compositional layers.

4.5. Comparative positioning within existing frameworks

The framework proposed here can be positioned relative to three established lines of work. Yuksel et al.'s [14] BRAAHMS system pioneered individualized calibration of physiological control signals but was designed for a single performer and does not address the multi-participant resolution problem. Donnarumma's [28] Xth Sense similarly operates on a single body, treating muscle sound as an expressive signal with high intimacy but no mechanism for combining multiple simultaneous physiological streams. Li et al. [31], conversely, address multi-participant physiological synchrony directly, using machine learning to predict team performance from interactional-level synchrony and coherence features, which is closer in spirit to the convergence-based logic this framework explicitly avoids. The node-level mapping framework differs from all three in combining (a) individualized,

Figure 5
Weighted Harmonic Resolution representation



protocol-driven baselining, (b) an explicit four-quadrant, rule-based sound-mapping layer grounded in music psychology, and (c) a multi-participant resolution logic (Weighted Harmonic Resolution) designed to preserve rather than converge individual signals. This positioning is qualitative: a direct quantitative benchmark against these systems would require access to their implementations or comparable datasets and is identified as a valuable direction for future comparative work once the pilot validation described in Section 5 has produced usable data.

5. Validation and Discussion

To ensure the mapping logic is accurate, we propose a validation process focused on a central question: Does the sound mapping generated by the system accurately reflect the participants' internal states? This is achieved by comparing the system's real-time output against standardized emotional benchmarks, bridging the gap between objective physiological data and subjective experience.

5.1. Preliminary pilot testing

Before proposing the full validation protocol below, the authors carried out informal bench testing of the complete acquisition-to-sound pipeline: CWP sensing, UDP transmission, Max/MSP feature extraction and mapping, and OSC-driven Ableton Live synthesis. These tests confirmed that the pipeline runs end to end with a latency low enough not to be perceptible as lag during live sound mapping and that the Zero Point calibration step reliably produces a stable baseline reading under controlled conditions. These observations are technical and anecdotal: no Self-Assessment Manikin (SAM) ratings or other structured perceptual measures were collected, and no claims about mapping accuracy or emotional coherence can be drawn from them. Their purpose is limited to demonstrating engineering feasibility ahead of the structured validation study described in Section 5.2, which remains, at the time of writing, still to be conducted.

5.2. Validation protocol and stimuli

The evaluation study is conducted within the same controlled acoustic environment utilized for system calibration. Once the group of participants is fully baselined via the Zero Point Protocol, they are simultaneously exposed to a series of specific audio samples chosen from the emotionally rated range of the IADS-2 database cross-validated across cultural samples [32]. These acoustic stimuli are carefully selected for their established, cross-validated emotional impact, ranging from high-valence pleasant environments to high-arousal jarring soundscapes.

As the participant listens to each stimulus, the framework processes their physiological response in real time and generates a corresponding musical sound mapping. To verify the psychological validity of this generated music, participants provide immediate, post-stimulus feedback using the SAM. This visual, nonverbal pictorial scale allows participants to indicate their levels of arousal and pleasure through intuitive icons rather than complex descriptive text, keeping the feedback loop clean and consistent with the continuous nature of the biometric architecture [33]. This validation strategy assumes that peripheral signals are not merely correlated with emotion but are constitutively part of it, a position consistent with recent evidence linking peripheral cardiac regulation to emotional valence during social interaction [34].

The primary goal of the subsequent data analysis is to establish system coherence—the degree to which the system's musical "interpretation" aligns with the participant's reported inner feeling. Rather than relying on raw observation alone, the study checks for statistical correlations (such as Spearman's Rho) to determine if higher physiological arousal consistently triggers high-arousal musical responses in the engine.

Spearman's Rho (ρ) is specifically utilized because SAM ratings provide ordinal data and physiological signals cannot be assumed to follow a normal parametric distribution, making a nonparametric correlation the most mathematically rigorous choice. This step confirms whether the mapping engine is accurately "reading" the node across various affective states. Furthermore, the "emotional transparency" of the system is evaluated through qualitative post-performance interviews, where participants describe whether the music functioned as a natural sonic mirror of their internal state. This validation seeks to demonstrate that "affective synchrony" can be both systematically induced and measured.

5.3. Discussion: Technical reliability and wearable evolution

High accuracy across diverse participants would validate the hypothesis that relative thresholding is superior to absolute biometric monitoring for affective computing; this remains a hypothesis pending the validation study in Section 5.2, though it is consistent with the preliminary pilot observations reported in Section 5.1. However, technical challenges such as motion artifacts—which can trigger "false positives" in arousal mapping—remain a concern for unrefined hardware.

This highlights the necessity of the strategic shift toward commercial wearable devices. Standardized smartwatches and medical-grade sensors include advanced, proprietary algorithms for motion compensation that can refine the GSR and heart rate signals before they reach the Max/MSP mapping layer. Transitioning to these devices ensures that the robust mapping logic developed here can be deployed in more dynamic, real-world collective environments without sacrificing data integrity.

5.4. Ethical considerations

Because the system continuously monitors physiological signals and infers affective states from them, its deployment raises ethical questions that go beyond standard data-protection practice [35]. First, informed consent must cover not only the collection of raw GSR/HRV data but also its real-time transformation into a public, audible output; participants should understand, before calibration, that their physiological state will be perceptible to an audience through the music and should retain the ability to withdraw from monitoring at any point in a performance without technical or social penalty. Second, privacy and data ownership require that physiological recordings, including Zero Point baselines, are stored no longer than necessary for the stated research or artistic purpose, are not repurposed for unrelated analysis without renewed consent, and remain under the participant's control regarding retention and deletion. Third, because the system deliberately shapes a shared sonic environment in response to individual and collective arousal, there is a risk of emotional manipulation: the same feedback loop that lets the music reflect a participant's state could, in principle, be used to drive it in a particular direction rather than merely mirror it. The framework as designed is intentionally mirroring rather than directive, and this distinction should be maintained and made explicit to participants in any future deployment or publication of results. Finally, because affective inference from peripheral signals is probabilistic rather than certain (Section 2.2), any public or performative display of a participant's inferred emotional state carries a risk of misrepresentation that should be disclosed to both participants and audiences as an inherent limitation of the system, not a guaranteed reading of an individual's inner state.

6. Conclusion and Future Directions

This paper has detailed the evolution of a robust biometric processing framework, transitioning from a collective experimental prototype into a refined system for high-resolution node mapping. By leveraging the technical foundations of the Strings (SCHI) project, we have demonstrated that physiological data can be more than just raw numbers—it can function as a dynamic, real-time musical language that bridges the gap between individual biology and collective expression.

6.1. Summary of contributions

The primary contribution of this research is the development of a mapping framework that prioritizes node-specific calibration. By establishing a "Zero Point of Restfulness" through a standardized baseline protocol, the system successfully addresses the common pitfalls of biometric monitoring, such as interpersonal physiological variability. Furthermore, the creation of a structured affective mapping matrix provides a scalable method for translating internal ANS states into intuitive, perceptually coherent soundscapes. At its current stage, this contribution is a validated engineering pipeline (Section 5.1) combined with a conceptually and theoretically grounded mapping design (Sections 2–4); the claim that the resulting music constitutes an accurate "sonic mirror" of participants' emotional states, rather than a plausible design intended to function as one, remains to be empirically confirmed through the validation study proposed in Section 5.2.

6.2. The path forward: Multidisciplinary synthesis and societal impact

The strategic transition from the CWP toward medical-grade, off-the-shelf wearables marks a significant turning point for the scientific validity of this research. This evolution moves beyond the limitations of isolated prototyping toward a multidisciplinary framework where music, science, and engineering converge to study human collectiveness.

Future work will focus on the integration of medical industry-grade devices previously reserved for laboratory settings, allowing for a more rigorous understanding of how physiological nodes respond to esthetic stimuli and grounding musical performance in verified biological truth. Furthermore, collaboration with computer engineering is essential to enhance the real-time interaction between biological data and generative systems, developing more sophisticated algorithms for multimodal output—not only in sound mapping but also in immersive visual environments that respond to the collective physiological presence of the group.

Beyond the conservatory, this framework offers a new lens through which to view human interaction in society. By mapping through sound the unspoken emotional states of groups, we can explore new forms of nonverbal social cohesion, communal empathy, and collective regulation. This approach reimagines technology not as an isolating tool but as a mediator that facilitates a deeper, more resonant dialog between individuals within the social fabric. Ultimately, this work proves that by combining robust signal processing with a deep, cross-disciplinary understanding of human psychophysiology, we can create technology that not only monitors the human body but also translates the physiological polyphony of the collective into a meaningful, shared cultural experience.

Ethical Statement

The authors declare that all potential respondents were fully informed about the survey and participation was voluntary.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data are available from the corresponding author upon reasonable request.

Author Contribution Statement

Dario Mattia: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Fabrizio Festa:** Conceptualization, Resources, Writing – review & editing, Supervision, Project administration.

References

[1] Picard, R. W. (1997). *Affective computing*. USA: MIT Press. https://doi.org/10.7551/mitpress/1140.001.0001

[2] Gedam, S., & Paul, S. (2021). A review on mental stress detection using wearable sensors and machine learning techniques. *IEEE Access*, *9*, 84045–84066. https://doi.org/10.1109/ACCESS.2021.3085502

[3] Lewin, K. (1936). *Principles of topological psychology*. USA: McGraw-Hill.

[4] Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3-4), 169–200. https://doi.org/10.1080/02699939208411068

[5] Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://psycnet.apa.org/doi/10.1037/h0077714

[6] Boucsein, W. (2012). *Electrodermal activity*. Germany: Springer Science & Business Media. https://doi.org/10.1007/978-1-4614-1126-0

[7] Greco, A., Valenza, G., & Scilingo, E. P. (2016). *Advances in electrodermal activity processing with applications for mental health: From heuristic methods to convex optimization*. Germany: Springer. https://doi.org/10.1007/978-3-319-46705-4

[8] Critchley, H. D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *Journal of Comparative Neurology*, *493*(1), 154–166. https://doi.org/10.1002/cne.20749

[9] Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health*, *5*, 258. https://doi.org/10.3389/fpubh.2017.00258

[10] Hermann, T., Hunt, A., & Neuhoff, J. G. (2011). *The sonification handbook*. Germany: Logos Verlag.
[11] Minciacchi, D., Bravi, R., & Rosenboom, D. (2023). Sonification, aesthetic representation of physical quantities. *Frontiers in Neuroscience*, *17*, 1162383. https://doi.org/10.3389/fnins.2023.1162383
[12] Peng, T., Choi, H., & Berger, J. (2024). The sound of data: Designing a framework for parameter mapping sonification. *Leonardo*, *57*(5), 566–572. https://doi.org/10.1162/leon_a_02568
[13] Zanella, A., Harrison, C. M., Lenzi, S., Cooke, J., Damsma, P., & Fleming, S. W. (2022). Sonification and sound design for astronomy research, education and public engagement. *Nature Astronomy*, *6*(11), 1241–1248. https://doi.org/10.1038/s41550-022-01721-z
[14] Yuksel, B. F., Afergan, D., Peck, E., Griffin, G., Harrison, L., Chen, N., . . . , & Jacob, R. (2015). BRAAHMS: A novel adaptive musical interface based on users' cognitive state. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 136–139.
[15] Tanaka, A. (2000). Musical performance practice on sensor-based instruments. In M. M. Wanderley & M. Battier (Eds.), *Trends in Gestural Control of Music* (pp. 389–405). IRCAM – Centre Pompidou.
[16] Reed, C. N., Skach, S., Strohmeier, P., & McPherson, A. P. (2022). Singing knit: Soft knit biosensing for augmenting vocal performances. In *Proceedings of the Augmented Humans International Conference 2022*, 170–183. https://doi.org/10.1145/3519391.3519412
[17] Reed, C. N., & McPherson, A. P. (2023). The body as sound: Unpacking vocal embodiment through auditory biofeedback. In *Proceedings of the Seventeenth International Conference on Tangible, Embedded, and Embodied Interaction*, 1–15. https://doi.org/10.1145/3569009.3572738
[18] Volta, E., & Di Stefano, N. (2024). Using wearable sensors to study musical experience: A systematic review. *Sensors*, *24*(17), 5783. https://doi.org/10.3390/s24175783
[19] Turchet, L., & Tomasetti, M. (2023). Immersive networked music performance systems: Identifying latency factors. In *2023 Immersive and 3D Audio: From Architecture to Automotive* (pp. 1–6). https://doi.org/10.1109/I3DA57090.2023.10289169
[20] Lei, X., Fan, H., Zhao, Y., Zhong, M., Wu, Z., Li, L., . . . , & Ren, G. (2025). MXene-enhanced laser-induced graphene flexible sensor with rapid response for monitoring pilots' body motion. *Micromachines*, *16*(5), 513. https://doi.org/10.3390/mi16050513
[21] Hickey, B. A., Chalmers, T., Newton, P., Lin, C. T., Sibbritt, D., McLachlan, C. S., . . . , & Lal, S. (2021). Smart devices and wearable technologies to detect and monitor mental health conditions and stress: A systematic review. *Sensors*, *21*(10), 3461. https://doi.org/10.3390/s21103461
[22] Sarhaddi, F., Kazemi, K., Azimi, I., Cao, R., Niela-Vilén, H., Axelin, A., . . . , & Rahmani, A. M. (2022). A comprehensive accuracy assessment of Samsung smartwatch heart rate and heart rate variability. *PloS One*, *17*(12), e0268361. https://doi.org/10.1371/journal.pone.0268361
[23] Hossain, M. B., Posada-Quintero, H. F., Kong, Y., McNaboe, R., & Chon, K. H. (2022). Automatic motion artifact detection in electrodermal activity data using machine learning. *Biomedical Signal Processing and Control*, *74*, 103483. https://doi.org/10.1016/j.bspc.2022.103483
[24] Llanes-Jurado, J., Carrasco-Ribelles, L. A., Alcañiz, M., Soria-Olivas, E., & Marín-Morales, J. (2023). Automatic artifact recognition and correction for electrodermal activity based on LSTM-CNN models. *Expert Systems with Applications*, *230*, 120581. https://doi.org/10.1016/j.eswa.2023.120581
[25] Clemente, P., Festa, F., & Giove, D (Eds.). (2026). *Mappe intelligenti e algoritmi cartografi*. Italy: Quodlibet.
[26] 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. https://doi.org/10.2196/52171
[27] Lévy, P. (1997). *Collective intelligence: Mankind's emerging world in cyberspace*. USA: Perseus Books.
[28] Donnarumma, M. (2012). Xth sense: A study of muscle sound as expressive means for musical performance. In *Proceedings of the International Computer Music Conference*.
[29] Pizarro, J., Zumeta, L., Bouchat, P., Włodarczyk, A., Rimé, B., Basabe, N., . . . , & Páez, D. (2022). Emotional processes, collective behavior, and social movements: A meta-analytic review of collective effervescence outcomes during collective gatherings and demonstrations. *Frontiers in Psychology*, *13*, 974683. https://doi.org/10.3389/fpsyg.2022.974683
[30] Algumaei, M., Hettiarachchi, I., Veerabhadrappa, R., & Bhatti, A. (2023). Physiological synchrony predict task performance and negative emotional state during a three-member collaborative task. *Sensors*, *23*(4), 2268. https://doi.org/10.3390/s23042268
[31] Li, M., Rojas, E., Bracken, B., Cooke, N., Desrochers, P., Gorman, J., . . . , & Zhou, S. (2025). Physiological synchrony as a predictor of team performance: A machine learning approach. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *69*(1), 333–337. https://doi.org/10.1177/10711813251357937
[32] Naal-Ruiz, N. E., Alonso-Valerdi, L. M., Ibarra-Zarate, D. I., Serrano-Cena, A., & Navas-Reascos, G. (2022). Mexican validation of the International Affective Digitized Sounds second edition (IADS-2) and additional sounds. *Scientific Reports*, *12*(1), 21824. https://doi.org/10.1038/s41598-022-26320-w
[33] 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. https://doi.org/10.1016/0005-7916(94)90063-9
[34] Deits-Lebehn, C., Smith, T. W., Williams, P. G., & Uchino, B. N. (2023). Heart rate variability during social interaction: Effects of valence and emotion regulation. *International Journal of Psychophysiology*, *190*, 20–29. https://doi.org/10.1016/j.ijpsycho.2023.06.004
[35] Mohammad, S. M. (2022). Ethics sheet for automatic emotion recognition and sentiment analysis. *Computational Linguistics*, *48*(2), 239–278. https://doi.org/10.1162/coli_a_00433

How to Cite: Mattia, D., & Festa, F. (2026). Real-Time Sound Mapping of Affect: A Wearable-Enabled Biometric Mapping Framework for Emotion Recognition in Music Generation. *Smart Wearable Technology*. https://doi.org/10.47852/bonviewSWT620210550