

The Operational Acoustic–Affective Space (OAAS): a framework for the design and evaluation of functional affective sound environments

Berardo de Jesús Rodríguez^{*}, Juliana Zapata-Cardona

Grupo de Investigación Patobiología QUIRON, Escuela de Medicina Veterinaria, Universidad de Antioquia, Calle 70 No. 52-21, Medellín 050010, Colombia

ARTICLE INFO

Keywords:

Animal vocalizations
Animal welfare
Bioacoustics
Dimensional modeling
Functional music
Principal component analysis
Sound design

ABSTRACT

Sound shapes emotional, behavioral, and physiological states in living systems. Despite growing evidence that humans and other animals respond affectively to sound, most research still relies on exposure-based paradigms and lacks translational, design-oriented methods—particularly in animal models where affect must be inferred from objective indicators. Consequently, links between acoustic structure and affective modulation remain weakly formalized, limiting reproducibility and cross-study comparability across heterogeneous sound sources. We introduce the Operational Acoustic–Affective Space (OAAS), a framework for designing and evaluating functional affective sound environments. OAAS defines a shared coordinate space in which sound environments can be positioned, compared, and intentionally modified relative to regulatory objectives. Rather than prescribing specific aesthetic forms or restricting inputs to pre-existing human music, the framework provides a coordinate geometry derived from objective acoustic descriptors and biologically grounded affective anchors, enabling interpretable analysis and controlled navigation within acoustic systems. We demonstrate OAAS using porcine-centered case studies that (1) define biologically grounded vocalization-derived reference anchors within a joint acoustic embedding space, (2) situate a veterinary functional music program for post hoc interpretation, and (3) validate operability through directed acoustic transformations that produce measurable displacement toward predefined affective reference regions. OAAS provides an extensible basis for translational affective sound design across species and contexts.

1. Introduction

The design of synthetic acoustic environments is an emerging applied field concerned with understanding and shaping how sound structures influence emotional, behavioral, and physiological states in living systems. Unlike isolated auditory stimuli, synthetic acoustic environments are conceived as structured, time-extended sound systems designed to modify the perceptual and affective context in which organisms are immersed, consistent with the soundscape paradigm and its operational definitions [1]. In recent years, synthetic sound environments have also gained relevance in computational acoustics and auditory research as controllable systems for simulating complex soundscapes and producing synthetic datasets for sound event detection, monitoring, and transfer learning [2,3].

Although existing soundscape standards and affective appraisal frameworks provide essential conceptual foundations for describing and measuring how environments are perceived, they rarely provide a unified geometric framework that supports intervention-oriented design

and reproducible navigation across heterogeneous sound environments [1,4]. In particular, there remains no operational coordinate framework that allows heterogeneous sound environments (e.g., functional music interventions, vocalizations, ecological soundscapes, and challenge reference signals) to be placed in a shared space and manipulated through explicit geometric operations—such as centroids, distances, trajectories, and displacement—under standardized acoustic representations. As a result, sound-based interventions are often treated as ad hoc exposures rather than controllable affective systems, limiting reproducibility, cross-study comparability, and the capacity to formalize “target states” and quantify directed change.

Despite extensive evidence that animals and humans exhibit emotional and physiological responses to sound, the field still lacks translational, design-oriented methodologies that connect acoustic structure with affective modulation in a reproducible way—particularly in animal models, where affective state must be inferred from objective indicators rather than verbal report [5,6]. Most studies remain exposure-based and frequently rely on pre-existing human music (often Western

^{*} Corresponding author.

E-mail addresses: berardo.rodriguez@udea.edu.co (B. de Jesús Rodríguez), juliana.zapata9@udea.edu.co (J. Zapata-Cardona).

<https://doi.org/10.1016/j.apacoust.2026.111457>

Received 15 March 2026; Received in revised form 13 June 2026; Accepted 26 June 2026

Available online 29 June 2026

0003-682X/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

repertoires), reporting heterogeneous outcomes while providing limited mechanistic insight into which acoustic properties drive regulation, engagement, or aversion [7,8]. Moreover, methodological constraints remain a persistent barrier in animal-centered music research, where controlled experimentation, acoustic standardization, and species-appropriate interpretation are often limited [9]. In addition, many approaches to “emotion in music” rely strongly on subjective self-report paradigms that are not directly portable to non-human contexts, reinforcing the need for frameworks that can operate with objective acoustic structure and organism-level inference when working across species [4,5].

A key translational bridge for such inference is animal vocal communication: across mammals and other vertebrates, vocal signals systematically encode affect-related information through their spectro-temporal structure, supporting the notion of cross-species acoustic codes for arousal and related affective dimensions [10,11]. In farm animals, vocalizations are increasingly treated as operational welfare indicators under context-grounded emission conditions [12]. This principle is directly relevant here because the SoundWel porcine vocal library provides labeled call categories and documented emission conditions with context-associated affective valence, enabling the construction of biologically grounded acoustic reference regions for subsequent analytical calibration [13].

Recent evidence from commercial farms suggests that functional veterinary music can improve maternal behavior and reduce piglet mortality during lactation [14]. Yet such findings remain difficult to formalize and compare across contexts because interventions are rarely positioned relative to standardized baseline soundscapes or evaluated through a shared operational coordinate space. This limitation is especially relevant under real-world production conditions, where acoustic context is dynamic, uncontrollable, and often dominated by complex background noise, making it essential to define frameworks capable of quantifying baseline-to-intervention shifts and comparing different acoustic strategies under a common representation.

This paper addresses this gap by introducing the OAAS, an applied framework for the analysis, design, and evaluation of synthetic acoustic environments conceived as functional affective systems. OAAS does not prescribe a specific acoustic or musical strategy. Instead, it provides a structured coordinate space in which acoustic environments can be positioned, compared, and systematically modified according to affective and regulatory objectives, enabling explicit representation of both baseline states and intentional transformations [15–17]. In this context, “affective” is operationally grounded on biologically meaningful vocal reference signals recorded under well-defined emission conditions, such that affect-oriented displacement in OAAS reflects acoustic proximity to context-anchored positive and negative centroids rather than abstract emotion inference from geometry alone.

In this study, the term synthetic acoustic environments refers to intentionally constructed soundscapes designed to shape the affective and regulatory conditions of a living system. These environments are conceived as structured acoustic ecosystems in which spectral, temporal, dynamic, and spatial properties interact over time to generate coherent perceptual and affective contexts, consistent with the foundational soundscape tradition [18,19] and with standard soundscape conceptualizations [1]. OAAS adopts this ecological-systemic view of sound and extends it into an operational framework that enables synthetic environments to be quantitatively positioned, compared, and modified within a shared acoustic-affective coordinate space.

Within this framework, music is treated as one possible implementation of a synthetic acoustic environment rather than a privileged category. While music represents a culturally formalized mode of sound organization in humans, musicality can be understood more broadly as an adaptive capacity for structuring acoustic patterns in ways that modulate attention, arousal, and affective meaning [20,21]. Accordingly, music is conceptualized here as an intentional soundscape: a constructed acoustic micro-ecosystem whose organization can evoke

ecological resonance without requiring direct imitation of vocal expressions [18,19]. This perspective also clarifies the role of vocalizations in the present work: auditory perception is not adapted to vocal communication alone, but to the broader acoustic structure of ecosystems in which vocalizations function as biologically meaningful signals embedded within environmental soundscapes. Vocalizations are therefore used here not as musical templates, but as ecologically grounded affective anchors for defining OAAS reference regions (Table 1).

The development of OAAS builds upon an established body of empirical research using animal models, where sound-based interventions—including veterinary functional music—were designed, deployed, and validated [15–17]. In the present study, this empirical foundation is not revisited experimentally; rather, it is used to formalize an operational acoustic-affective design space and to demonstrate how biologically grounded vocal reference signals (porcine vocalizations) can be used to structure and interpret synthetic acoustic environments within a shared coordinate domain [13].

The objective of this study is to formalize OAAS as a reproducible and extensible framework that enables synthetic acoustic environments to be represented, compared, and intentionally navigated within a low-dimensional acoustic coordinate projection of the acoustic-affective space (PCA-based), supporting systematic evaluation and directed displacement toward predefined affective reference regions. The complete OAAS workflow is summarized in Fig. 1.

Schematic overview of the analysis workflow: (1) vocal-only diagnostic OAAS (“proto-atlas”) to establish biologically grounded anchor geometry; (2) OAAS joint embedding to position functional music, challenge stimuli, and noise reference signals relative to the vocal anchors; (3) construction of POS/NEG vocal centroids; (4) computation of centroid distances (d_{POS} , d_{NEG}), bias (pos_bias), and OAAS-derived acoustic-affective index (AAI_{OAAS}); and (5) operational ranking/selection for sound design and intervention planning (Table 4; Supplementary Table S7).

2. Materials and methods

This section describes the methodological basis of the proposed framework. Rather than presenting a single experimental protocol, it outlines the conceptual design principles, analytical tools, and evaluation strategies that constitute the OAAS, drawing on a body of empirical work conducted across multiple studies and applied contexts.

2.1. OAAS as a multidimensional acoustic embedding space

Building on the conceptual considerations introduced in Section 1, OAAS is implemented as an applied framework that operationalizes multidimensional acoustic feature configurations into a geometric coordinate space. This representation allows synthetic acoustic environments to be positioned, compared, and intentionally transformed with respect to biologically grounded reference anchors. Within OAAS, each stimulus—whether a vocalization, a musical intervention, or a control signal—is encoded as a vector of standardized acoustic descriptors and mapped as a point in a shared embedding, enabling distance- and trajectory-based analyses within a common coordinate domain.

In this study, the term *operational* indicates that OAAS is specified through measurable and actionable variables. Feature vectors are standardized and projected into a reduced coordinate system (PC1–PC3) that supports quantitative geometric operations. Because affect-related interpretation is inherently species- and context-dependent, OAAS-POS and OAAS-NEG are defined here as context-anchored reference centroids derived from labeled subsets of porcine vocalizations in the SoundWel corpus [13]. Although the present work instantiates OAAS using vocalizations as biologically contextualized anchors, the framework itself is not restricted to this domain: alternative anchor sets may be defined depending on species, application context, or regulatory objective (e.g., ecological soundscapes, behavioral reference signals, or

Table 1

Core operational acoustic descriptors used for OAAS PCA embedding.

Domain	Descriptor	Variable	Operational definition
Spectral / informational	Spectral entropy mean	se_mean	Mean frame-wise spectral entropy from short-time Fourier transform (STFT) magnitude
Spectral / informational	Spectral entropy SD	se_std	Standard deviation of frame-wise spectral entropy
Spectral / informational	Spectral entropy 95th percentile	se_p95	95th percentile of frame-wise spectral entropy
Spectral texture	Spectral flatness mean	flatness_mean	Mean frame-wise spectral flatness
Spectral texture	Spectral flatness SD	flatness_std	Standard deviation of frame-wise spectral flatness
Spectral texture	Spectral flatness 95th percentile	flatness_p95	95th percentile of frame-wise spectral flatness
Temporal / dynamic	RMS temporal entropy	te_rms	Entropy of RMS envelope distribution
Temporal / informational	Multiscale envelope entropy	mse_env	Multiscale entropy descriptor computed from the RMS envelope
Harmonic organization	Harmonic ratio	harm_ratio	Ratio of harmonic to total energy derived from harmonic–percussive source separation (HPSS)

Note: Acoustic descriptors shown in this table represent the reduced operational feature set used for PCA-based OAAS embedding in the present implementation. These variables were selected from a broader exploratory descriptor pool based on robustness, interpretability, stability across heterogeneous sound environments, and compatibility with joint embedding of vocalizations, musical stimuli, and reference signals. OAAS itself is not restricted to this specific feature configuration and can incorporate additional descriptors depending on species, application context, and design objectives.

species-specific target repertoires) (Table 2).

Importantly, OAAS should not be interpreted as a geometric “emotion detector.” Affective interpretation does not emerge from geometry in isolation, but from proximity to contextualized reference anchors recorded under known ethological emission conditions. Accordingly, OAAS quantifies structured acoustic relationships among stimuli relative to these operational references rather than claiming direct measurement of internal emotional state.

Within this framework, OAAS supports three complementary functions: (1) representation of baseline acoustic states in a shared coordinate domain; (2) analysis of similarity structure through distances, centroid relations, and displacement trajectories; and (3) design of controlled transformations by moving stimuli toward or away from reference regions defined by biologically grounded anchors. Two-dimensional OAAS plots are presented only for visualization, whereas quantitative operations are computed in the full three-dimensional coordinate space (PC1–PC3). The construction of the acoustic feature geometry and projection procedure is detailed in Section 2.2.

2.2. OAAS design parameters

Within the OAAS framework, synthetic acoustic environments are defined and manipulated through a set of design parameters describing how sound is structured, delivered, and experienced over time. Parameters do not prescribe specific sound sources or aesthetic forms; instead, they characterize affect-relevant acoustic properties that can be systematically analyzed and controlled [15].

For operational purposes, OAAS parameters are organized into four interrelated domains: spectral, temporal, dynamic, and informational.

Spectral domain. Spectral parameters describe frequency-related properties such as bandwidth, spectral centroid, spectral slope, harmonicity, and energy distribution across bands. Spectral configuration determines perceptual salience and biological accessibility, and relates to timbral qualities and similarity to natural vocal or ecological sound patterns [15].

Temporal domain. Temporal parameters characterize event timing, rhythmic organization, tempo and pulse regularity, temporal density, and modulation rates. Temporal structure influences predictability, entrainment, attention, and arousal regulation across exposure regimes [5,15,22].

Dynamic domain. Dynamic parameters describe amplitude-related behavior over time, including sound level, dynamic range, onset and decay profiles, and amplitude modulation. Dynamic control shapes energetic profiles and supports avoidance of overstimulation in long-term exposures [5,6].

Informational domain. Informational parameters capture higher-level organization such as complexity, variability, redundancy, novelty, and structural coherence. This domain is relevant for

distinguishing neutral, enriching, and disruptive environments and for supporting stability or exploratory engagement [22,23].

Integration of domains. OAAS treats these domains as integrated, not independent. Each synthetic acoustic environment corresponds to a configuration in this multidimensional parameter space, which can be analyzed directly or projected into low-dimensional OAAS coordinates. This integration supports: (1) comparative analysis; (2) identification of parameter ranges linked to affective outcomes; and (3) design decision-making via controlled transitions in OAAS.

For the operational PCA implementation used in the present study, a reduced set of core acoustic descriptors was extracted to construct the OAAS embedding. These descriptors were selected for robustness, interpretability, and stability across heterogeneous acoustic environments.

2.3. Acoustic and musical design pipeline

Within OAAS, the production of synthetic acoustic environments is conceived as a design process guided by affective objectives rather than by predefined aesthetic categories. OAAS links diagnostic analysis, design decision-making, and post-production evaluation within a shared acoustic–affective space.

Contextual diagnosis. A baseline acoustic context is characterized (vocalizations, environmental sound, or existing interventions). OAAS parameters identify dominant spectral, temporal, dynamic, informational, and spatial features. Output: a reference region in OAAS.

Target definition. A target region is defined based on objectives such as stress reduction, affiliative engagement, or arousal modulation. Operationally, target regions can be specified as proximity objectives relative to context-anchored vocalization reference centroids (e.g., OAAS-POS vs. OAAS-NEG), rather than as direct emotional labels. Target definition constrains parameter selection.

Iterative design decisions. Sound production proceeds through controlled modifications in OAAS parameter space. In musical design: instrumentation, timbre, harmonic density, rhythmic structure, dynamics, and pacing. In non-musical/hybrid design: source selection, layering, noise shaping, spatial diffusion.

Post-production OAAS verification. Stimuli are analyzed and projected into OAAS to verify alignment with the intended target. Discrepancies inform iterative refinement.

Spatial configuration is treated as an integral design component: diffusion, movement, and positioning influence perceptual integration, affecting OAAS positioning depending on acoustic reproduction conditions, including speaker configuration, enclosure acoustics, spatial layout, and calibrated sound pressure levels during playback.

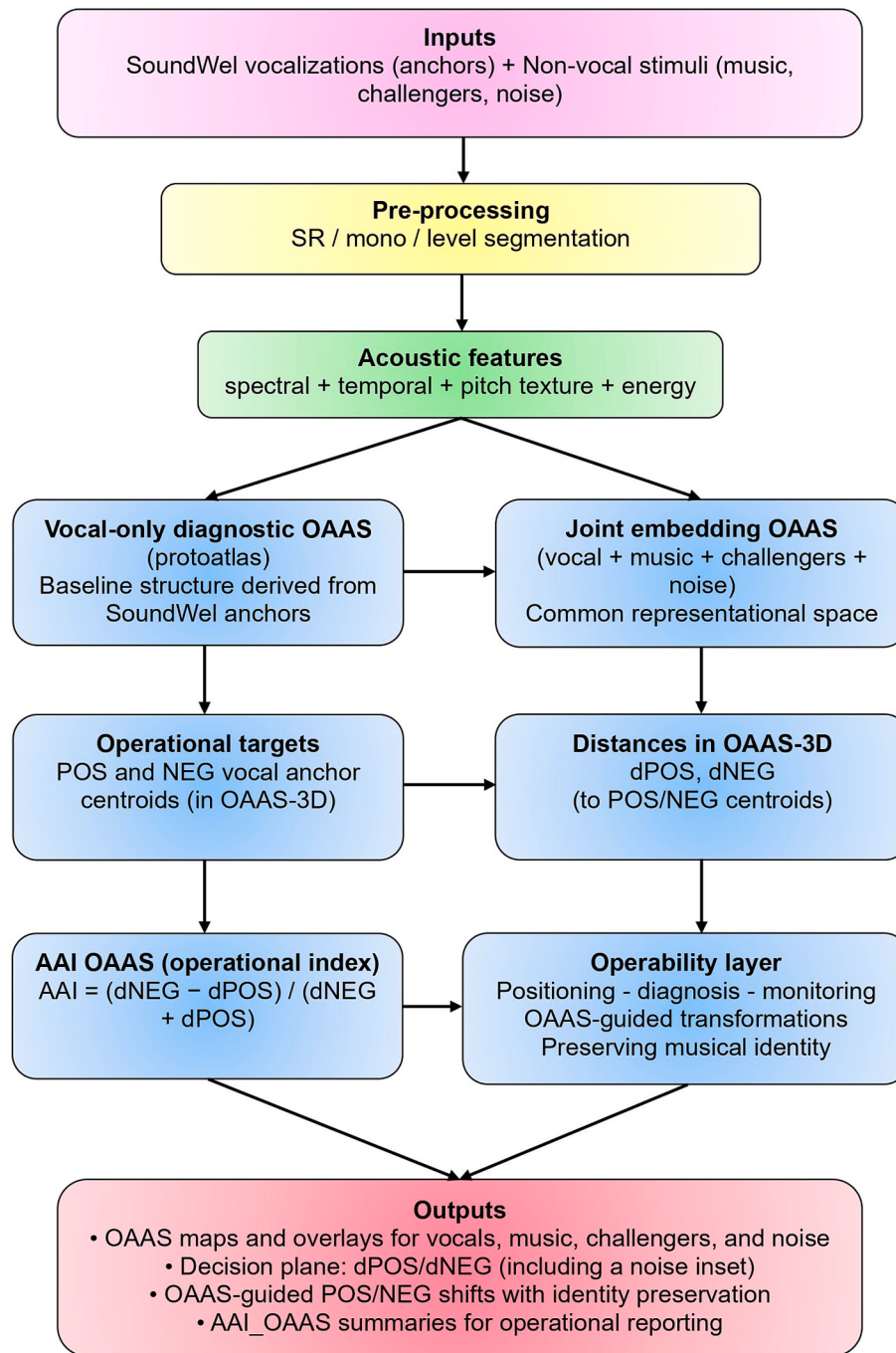


Fig. 1. OAAS framework workflow (from biological anchors to operational selection/design).

Detailed production protocols and implementation examples for the veterinary functional music program used as an intervention reference are reported in Zapata-Cardona et al. [23], including the associated [Supplementary Information](#) with audio excerpts and methodological details [17].

2.4. Vocal-only diagnostic OAAS (ethological embedding)

We implemented a vocal-only diagnostic OAAS (hereafter vOAAS) as an ethologically grounded reference configuration. In the framework, this vocal-only embedding serves as the biological anchor geometry used to interpret and operationalize the OAAS joint embedding: it delineates contextual vocal regions, supports the construction of POS/NEG vocal centroids, and provides a stable baseline against which functional

music, challenge stimuli, and noise reference signals can be positioned and quantified.

2.4.1. Vocal-only dataset

The diagnostic embedding was fitted exclusively on the SoundWel contextual vocalization ensembles [13,24]. Because very short calls can generate unstable microstructure in valence–arousal projections, vocalizations were assembled into ~ 180 s context-specific ensembles to ensure temporal comparability with the musical stimuli. The resulting diagnostic dataset comprised $N = 36$ ensembles representing 18 contextual categories (two ensembles per context).

2.4.2. Standardization and PCA fit

All acoustic features were z-standardized using a StandardScaler

Table 2
OAAS-based design pipeline for synthetic acoustic environments.

Stage	Goal	OAAS operation	Output
Diagnosis	Characterize baseline	Positioning	OAAS baseline
Objective	Define target	Targeting	OAAS target
Strategy	Select approach	Path selection	Design plan
Production	Build candidate	Navigation	Candidate environment
Verification	Evaluate fit	Projection + comparison	Evaluated configuration
Refinement	Reduce mismatch	Iterative adjustment	Finalized environment
Deployment	Assess impact	Outcome validation	Validated intervention

Note: “OAAS operation” refers to standardized geometric use (positioning, targeting, navigation, and comparison) under a common acoustic descriptor space.

fitted on the vocal-only dataset. PCA was then computed on the standardized vocal-only feature matrix, yielding the diagnostic coordinate system (vOAAS1–vOAAS3). PCA was selected for its linear interpretability, stability under re-embedding, and suitability for explicit geometric operations (centroids, distances, and constrained displacement) within a reproducible coordinate system. Importantly, this PCA is label-blind and reflects only covariance structure of the acoustic descriptors, not contextual valence assignments.

2.4.3. Context centroids and operational anchors

For each contextual category c , a centroid μ_c was computed in the full three-dimensional vOAAS space:

$$\mu_c = \frac{1}{n_c} \sum_{i \in c} z_i$$

Dispersion was summarized by the median radial distance around each centroid. Positive and negative operational anchor sets were defined by grouping SoundWel contexts associated with affiliative/homeostatic emission conditions versus distress/aversive conditions, respectively [13]. Contexts not included in these anchor sets were treated as non-anchor categories (Table 3).

In the present implementation, the positive and negative reference centroids were derived from vocalization contexts with documented affiliative or distress-related emission conditions in the SoundWel corpus. Positive anchors included contexts such as Huddling, Run, Enriched, AfterNursing, and BeforeNursing, whereas negative anchors included contexts associated with nociception, restraint, social isolation, and agonistic interaction (e.g., Castration, Crushing, Restrain, Long-Isolation, ShortIsolation, MissedNursing, Fighting, and Barren). These anchors provide biologically grounded reference configurations that structure the OAAS without imposing categorical affect labels on the space itself.

2.4.4. Distance-based diagnostic metrics

For any projected stimulus (music or reference signal), Euclidean distances to the POS and NEG anchor centroids were computed in 3D:

$$d_{POS} = \|z - \mu_{POS}\|_2$$

$$d_{NEG} = \|z - \mu_{NEG}\|_2$$

and the signed proximity index:

$$\Delta = d_{POS} - d_{NEG}$$

where $\Delta < 0$ indicates closer proximity to the POS anchor region, and $\Delta > 0$ indicates closer proximity to the NEG anchor region. These metrics provide an operational proxy of directed acoustic proximity, not a direct measurement of emotional state.

Table 3
Conceptual alignment between SoundWel context categories and QBA-derived affective adjective labels used in OAAS intervention research [3].

SoundWel vocalization context (emission conditions)	Indicative affective context (SoundWel descriptors)	Closest QBA-oriented affective adjective labels (OAAS framework)	Interpretation notes for OAAS anchoring
Social affiliative / contact contexts	Positive / low-moderate activation	Calm, relaxed, content, affiliative	Positive reference core; emphasizes stability and low irregularity [3].
Exploratory / active engagement contexts	Positive or neutral / high activation	Curious, active, engaged, playful	High activation without distress; supports enrichment-oriented anchoring.
Baseline / routine neutral contexts	Neutral / low activation	Neutral, regulated, stable	Regulatory baseline; neutrality is treated as a meaningful state, not absence of affect.
Social separation / transient uncertainty contexts	Negative or mixed / moderate activation	Alert, uneasy, tense, uncertain	Transitional states; anchoring emphasizes ambiguity and context dependence.
Restraint / handling stress contexts	Negative / high activation	Anxious, distressed, agitated	High irregularity + high activation reference; interpreted cautiously as structural alignment only.
Acute nociceptive contexts (e.g., castration)	Negative / very high activation	Acute distress, panic-like agitation	Extreme high activation negative contexts; used as boundary region in OAAS.

Important note: This table represents a conceptual anchoring strategy for interpretability, not categorical equivalence. OAAS proximity indicates structural alignment of acoustic configuration rather than direct emotional identity.

Table 4
OAAS geometric descriptors used for quantitative comparison and operability assessment.

Descriptor	Definition	OAAS interpretation	Use in this study
OAAS coordinate (x)	Stimulus position in (PC1–PC3)	Acoustic–affective configuration	Represents vocalization clusters, musical stimuli, and reference signals (challenge stimuli and broadband noise)
Centroid (c)	Mean coordinate of a group/region	Reference region (affective anchor)	SoundWel clusters and program families
Distance $d(x, c)$	Euclidean distance (3D)	Structural proximity	Quantifies similarity to affective reference contexts
Displacement Δx	Vector between configurations	Directed navigation	Used to evaluate OAAS operability and directed transformations
Distance change Δd	Difference in distance across transformations	Improvement/constraint	Captures asymmetry in navigability

Note: Geometric descriptors were computed within the PCA-reduced OAAS space and are intended as operational measures of directional acoustic displacement rather than absolute psychoacoustic quantities.

The vocal-only diagnostic OAAS was constructed using context-

specific macrosegment ensembles instead of isolated vocal events, which reduces sensitivity to local acoustic fluctuations and transient emission variability. Accordingly, the resulting centroids should be interpreted as operational context-level reference regions reflecting aggregated acoustic organization across multiple emissions. Although the present study did not perform formal bootstrap or resampling-based centroid stability analyses, the observed clustering structure and context-level separation remained qualitatively consistent across the analyzed vocal ensembles. Because very short vocal events can generate unstable local geometric organization within the acoustic-affective embedding and greater sensitivity to transient acoustic variability in preliminary analyses, SoundWel calls were aggregated into context-specific macrosegment ensembles. This strategy was intended to obtain more stable context-level acoustic profiles for centroid construction while preserving the contextual specificity and ethological interpretability of the original emission categories.

2.5. Evaluation methodology

All musical stimuli were converted to mono, resampled to 48 kHz, and loudness-normalized to -23 loudness units relative to full scale (LUFS) following the European Broadcasting Union (EBU) R128 standard prior to feature extraction. This ensured acoustic comparability across musical stimuli and with the SoundWel-derived porcine vocalization baseline [13].

Evaluation in OAAS is conceived as a multi-level process assessing synthetic acoustic environments as functional affective systems rather than isolated stimuli. OAAS integrates evaluation into the same acoustic-affective space used for diagnosis and design, enabling coherent comparisons between natural vocal expressions, designed sound environments, and control stimuli.

Evaluation operates across three interrelated levels: (1) acoustic verification, (2) organismal response evaluation and affective anchoring, and (3) OAAS-based geometric analyses (centroids, distances, and displacement).

2.5.1. OAAS joint embedding (multi-domain embedding) and acoustic verification

In preliminary analyses, a valence–arousal (PAD) projection was also tested, but it showed strong collapse/instability under cross-domain scaling; we therefore report OAAS-derived centroid-distance metrics (bias and AAI_{OAAS}) as the primary operational summaries for this manuscript.

Each produced sound environment is analyzed using the same OAAS parameter domains employed during design. Stimuli are projected into the OAAS to confirm their positioning within the intended region. This step supports reproducibility and interpretation of subsequent analyses.

In addition, to define a biologically grounded OAAS reference baseline, porcine vocalizations from the SoundWel database were processed to enable direct comparison with longer-duration designed acoustic environments (e.g., musical stimuli) [13]. Because individual vocalizations are short and locally variable, recordings were aggregated into category-specific macrosegments of approximately 3 min, preserving contextual specificity while providing time-extended acoustic profiles comparable in scale to intervention stimuli.

For each macrosegment, acoustic descriptors spanning the OAAS parameter domains (spectral, temporal, dynamic, and informational) were extracted following the standardized feature pipeline described above. Features were subsequently z-score normalized using the same reference scaling applied to all sound environments included in the OAAS, ensuring comparability between vocalization-derived baselines, musical stimuli, and challenge reference signals.

Standardization was implemented as feature-wise centering and scaling to unit variance (z-score) using a single fitted scaler (e.g., StandardScaler), applied consistently to all items prior to PCA. The scaler and PCA model were fitted on the combined dataset used to

populate OAAS (SoundWel macrosegments + all intervention stimuli + challenge reference signals), and the same fitted transforms were then used to project each item into PC1–PC3 [13,24]. This combined corpus comprised $N = 114$ items (vocalization macrosegments $n = 36$, music program stimuli $n = 72$, challenge reference signals $n = 6$). The challenge control set was included for exploratory calibration rather than as a priori ground-truth “negative” labels; it comprised a small set of reference sounds expected to elicit or encode aversive/non-affiliative structure, including: (1) a stimulus previously evaluated under QBA exposure showing aversive responses, (2) a mechanically textured synthetic piece built from modular synthesis of porcine vocal elements, and (3) a piece embedding SoundWel vocalizations from negative emission contexts (see [Supplementary Table S7](#) for the full list).

PCA is label-blind (unsupervised): POS/NEG anchoring is defined post hoc by selecting vocalization subsets based on documented emission contexts, not by the PCA computation itself.

Normalized feature vectors were projected into the OAAS defined by the first three principal components derived from principal component analysis (PCA) (PC1–PC3). Explained variance ratios for PC1–PC3 and a summary of dominant feature loadings are reported in the [Supplementary Material](#) ([Supplementary Figs. S1–S2](#) are distributed together as a single PDF file for formatting consistency) ([Table S2](#)) to support interpretability. Vocalization categories were represented as regions in the OAAS and summarized by their centroids and dispersion profiles. These vocalization-derived regions function as biologically grounded affective reference anchors for subsequent post-hoc evaluation and distance-based analyses of designed sound environments ([Sections 3.2–3.3](#)).

[Supplementary Tables S1–S7](#) are provided together as a single consolidated PDF file for formatting consistency.

Explained variance ratios for the vocal-only PCA configuration are reported in [Supplementary Table S6](#).

2.5.2. Organismal response evaluation and affective anchoring

In animal models, organismal evaluation relies on behavioral and physiological indicators reflecting affective state without verbal report. Behavioral indicators include QBA-aligned adjective labels, ethological observations, affiliative/agonistic interactions, activity patterns, environmental use, and stress-related behaviors. Physiological indicators may include autonomic measures (e.g., heart rate variability (HRV)), endocrine markers, or production-related outcomes depending on the applied context.

Within OAAS, organismal response evaluation serves two complementary functions: (1) providing an empirical basis for anchoring acoustic regions to biologically meaningful affective contexts, and (2) supporting validation of designed environments when experimental response measures are available. Importantly, the present study does not introduce new behavioral or physiological experiments; instead, OAAS is anchored through biologically grounded affective reference contexts derived from previously validated sources.

Affective anchoring is established primarily through porcine vocalization categories extracted from the SoundWel database, a corpus of vocal signals emitted under well-defined production contexts with affect-related descriptors assigned by the original authors [3,31]. These vocalizations function as natural affective reference regions in OAAS, enabling the space to be populated with biologically meaningful acoustic configurations associated with affiliative/positive contexts, high-activation exploratory states, baseline regulatory conditions, transient uncertainty-related events, and sustained negative contexts.

To reinforce interpretability and maintain methodological continuity with previous intervention studies, SoundWel context categories were conceptually aligned with QBA-based affective descriptors previously employed in veterinary functional music interventions [40], while remaining consistent with established ethological frameworks of vocal affect expression [4]. Although SoundWel provides emission-context categories rather than QBA ratings per se, both frameworks share a

common objective: characterizing affect-related conditions through observable behavioural and contextual indicators. This conceptual linkage enables an interpretable mapping between naturally expressed affective communication and intervention-oriented affect modeling, without implying categorical equivalence.

Here, “adjective labels” refer to qualitative affect terms used for interpretability (e.g., calm, tense), while “ratings” refer to structured observer scores derived from standardized QBA protocols [39,40].

Crucially, OAAS interpretation avoids assuming that acoustic similarity implies equivalence of emotional experience. Instead, proximity within OAAS is interpreted as graded structural alignment: acoustic configurations may resemble those of affectively grounded vocal contexts while differing in emotional meaning due to environment, perception, and functional role. Therefore, acoustic convergence must be interpreted in relation to context descriptors and, when available, behavioral validation.

This anchoring strategy enables the operational identification of acoustic regions associated with positive and negative vocal-context references, supporting centroid-distance computation and directed displacement analysis within the OAAS framework.

The alignment between SoundWel emission contexts and QBA-derived affective descriptors should not be interpreted as a direct equivalence between vocal production events and observer-assigned emotional labels. Rather, the mapping is intended as an operational interpretive bridge linking biologically documented emission contexts with broader affect-related behavioral descriptors frequently used in animal welfare assessment. The correspondence is intended to support operational comparison between vocal-context organization and QBA-based behavioral interpretation within a shared acoustic-affective reference space.

2.5.3. OAAS-based geometric analyses (centroids, distances, and displacement)

Quantitative evaluation within OAAS is performed using geometric analyses in the acoustic-affective space rather than hypothesis-driven inferential statistics.

Each acoustic segment or stimulus is represented by its OAAS coordinates in three dimensions:

$$\mathbf{x} = (PC1, PC2, PC3)$$

For a region (cluster) containing N segments, the centroid is computed as:

$$\mathbf{c} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i$$

Similarity and separation between configurations are quantified using Euclidean distance in three dimensions:

$$d(\mathbf{x}, \mathbf{c}) = \sqrt{(x_1 - c_1)^2 + (x_2 - c_2)^2 + (x_3 - c_3)^2}$$

To operationalize relative positioning with respect to positive and negative vocal reference regions, we compute the distances to the POS and NEG vocal centroids:

$$d_{POS} = d(\mathbf{x}, \mathbf{c}_{POS})$$

$$d_{NEG} = d(\mathbf{x}, \mathbf{c}_{NEG})$$

and their differential distance:

$$\Delta = d_{POS} - d_{NEG}$$

where $\Delta < 0$ indicates closer proximity to the POS vocal reference region, and $\Delta > 0$ indicates closer proximity to the NEG region. For a normalized, bounded index that is comparable across datasets and avoids scale effects, we additionally report an OAAS-derived Affective

Acoustic Index:

$$AAI_OAAS = \frac{d_{NEG} - d_{POS}}{d_{NEG} + d_{POS}}$$

which ranges from -1 (maximal proximity to NEG) to $+1$ (maximal proximity to POS). Distances, Δ , and AAI_OAAS are computed in the full three-dimensional OAAS space (OAAS1–OAAS3); two-dimensional plots are used only for visualization. This centroid-based formulation yields an OAAS-derived Acoustic–Affective Index (AAI_OAAS), which provides a scalar operational summary of a stimulus’ relative positioning within the OAAS space. Rather than representing an affective state per se, the AAI_OAAS functions as a decision-oriented metric used throughout the Results to support stimulus comparison, ranking, and acoustic design.

Directional change resulting from OAAS-guided design is expressed as a displacement vector:

$$\Delta \mathbf{x} = \mathbf{x}_{final} - \mathbf{x}_{initial}$$

Two-dimensional projections are used exclusively for visualization and do not enter quantitative evaluation.

2.5.4. Directed displacement procedure

For each displacement sequence, a starting stimulus was iteratively modified under explicit constraints (e.g., preserving overall duration and avoiding clipping; bounded changes in spectral balance, temporal modulation, and dynamic profiles). The directional transformation process was implemented through a sequence of digital signal processing (DSP) operations including frequency-dependent spectral reshaping, band-specific amplitude weighting, temporal-envelope adjustment, loudness normalization, and iterative evaluation of OAAS displacement. After each modification, acoustic features were re-extracted, projected into the fixed joint OAAS model (using the same standardization and PCA transformation), and evaluated by the change in Euclidean distance to the target vocal centroid(s). Iteration continued until distance improvement saturated or a predefined tolerance/iteration limit was reached. Displacement progress was additionally summarized using $\Delta = d_{POS} - d_{NEG}$ and the bounded index AAI_OAAS, reported for each step in the transformation sequence.

To facilitate reproducibility, the original stimuli (W7_01 and W8_03) and the selected POS- and NEG-directed transformation variants analyzed in this study are provided as audio files in the public project repository.

A schematic overview of the computational and operational architecture of the OAAS framework is provided in [Supplementary Methods Figure SM1](#).

3. Results

We first introduce the joint-embedding OAAS configuration used for multi-domain positioning and visualization (Fig. 2), then report the vocalization-derived diagnostic baseline that defines the anchor topology (Fig. 3). We subsequently evaluate OAAS operability via directed stimulus displacement (Fig. 4) and conclude with centroid-distance positioning and OAAS-derived indices supporting ranking and design decisions (Fig. 5; Table 4).

As part of the present implementation, a valence–arousal projection previously used in QBA-oriented animal studies [39] was also evaluated, but it showed instability and geometric collapse under cross-domain scaling; it is therefore not used for main-text inference.

This section presents selected applications and case studies applying OAAS to (1) diagnostic acoustic baselines, (2) post-hoc evaluation of previously validated interventions, and (3) inter-species comparison within a shared representational space. Each case study follows the procedures outlined in Section 2.5 and is based on empirically grounded datasets, enabling examination of affect-relevant acoustic structure across distinct sound sources and temporal scales.

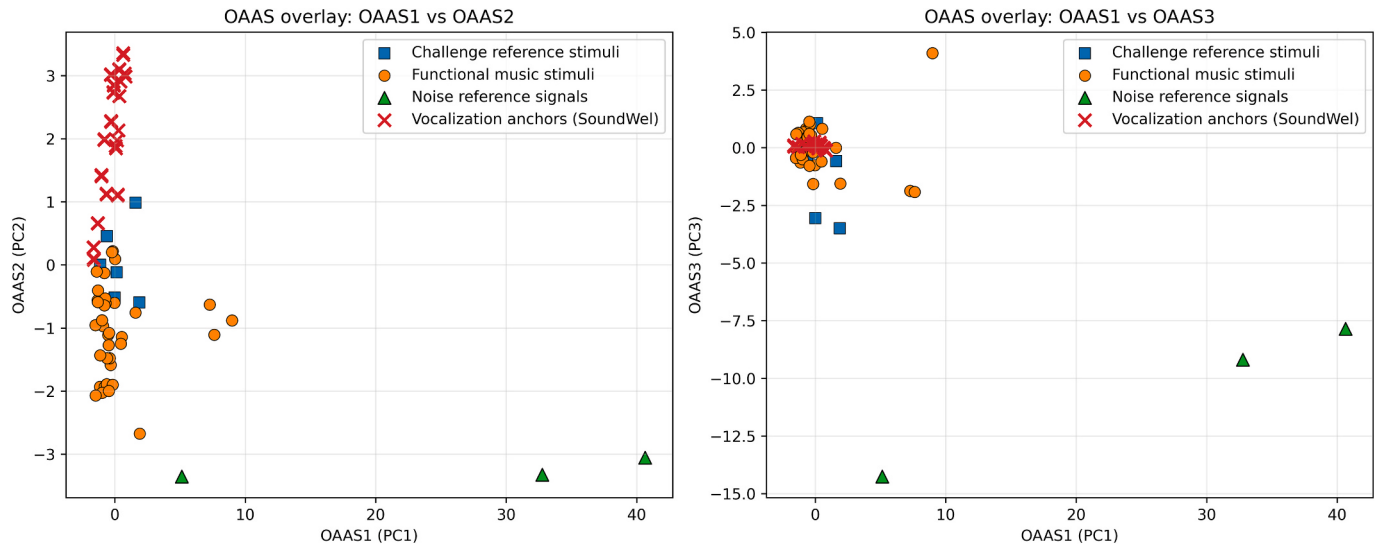


Fig. 2. Dual projection of the joint OAAS embedding (all stimulus domains). (A) PC1–PC2 cartographic overview. The joint OAAS embedding populated with vocalization anchors, functional music stimuli, challenge reference stimuli, and noise reference signals. (B) PC1–PC3 complementary projection. Alternative view of the same joint embedding to facilitate interpretation of separations that may be compressed in PC2.

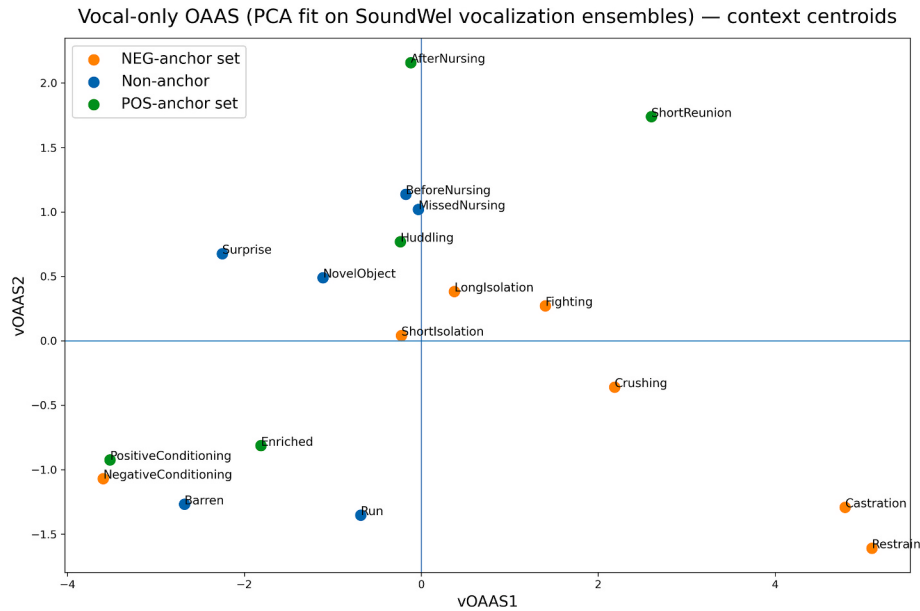


Fig. 3. Vocal-only diagnostic OAAS ("proto-atlas") of porcine vocalization categories.

3.1. Porcine vocalizations as an OAAS diagnostic baseline

This section establishes the biologically grounded reference topology used throughout the remaining analyses. Specifically, vocalization-derived regions provide the anchor geometry for subsequent post-hoc positioning of music, challenge reference signals, and noise reference signals in the joint OAAS embedding.

3.1.1. SoundWel vocalizations as biologically grounded anchors

As a first case study, the OAAS framework was applied to porcine vocalizations to establish an acoustic–affective reference baseline derived from natural emission contexts. The SoundWel pig vocalization database was used as a publicly documented, context-annotated corpus to populate OAAS with vocal-derived anchors spanning diverse production conditions [3,31]. In the present study, we do not replicate or validate the SoundWel classification strategy; instead, SoundWel labels

are used exclusively as reference emission contexts to define biologically grounded anchor regions for OAAS positioning and comparative geometry.

Porcine vocalizations were projected into the vocal-only diagnostic OAAS following the procedures described in Section 2. Detailed centroid-distance and bias metrics for the vocal-only OAAS projection (including vOAAS coordinates, d_{POS} , d_{NEG} , and bias values) are reported in Supplementary Table S5. The resulting OAAS baseline reveals structured organization of vocalization-derived regions within the acoustic–affective space (Fig. 3). Distinct regions emerge that correspond to affiliative/positive contexts, high-activation exploratory states, baseline regulatory conditions, transient uncertainty-related events, and sustained negative contexts. Importantly, these regions are not imposed through ad hoc assumptions, but emerge from the acoustic structure of the vocalizations in conjunction with independently defined emission contexts.

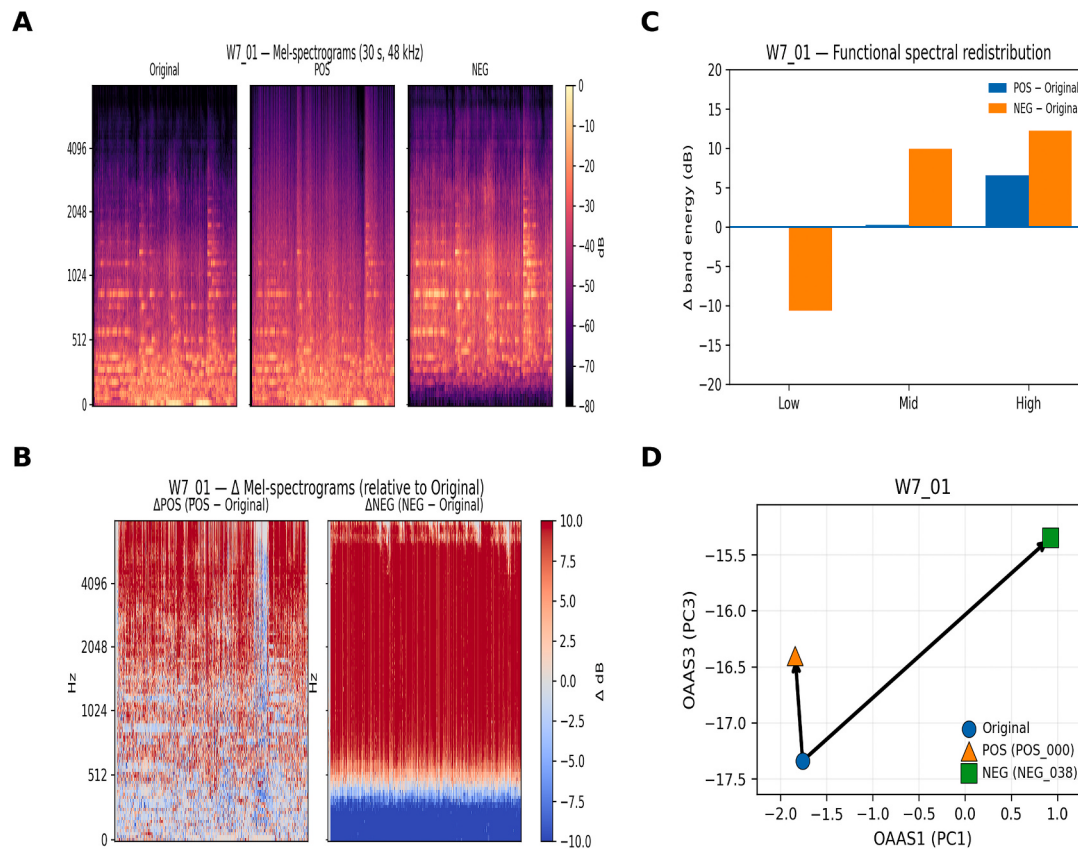


Fig. 4. OAAS-guided directed displacement of musical stimulus W7_01. Multi-level visualization of controlled OAAS-directed transformations toward vocal-derived reference regions: (A) mel-spectrograms of the original stimulus and selected positive (POS) and negative (NEG) variants (30 s, 48 kHz); (B) Δ mel-spectrograms relative to the original; (C) band-wise spectral energy redistribution across low-, mid-, and high-frequency bands; and (D) trajectory of the stimulus and its POS/NEG variants within the joint OAAS embedding (PC1–PC3 plane). The dominant displacement along PC1 and the unequal magnitude of POS- and NEG-directed shifts illustrate asymmetric navigability within the acoustic–affective space under bounded transformation constraints. Mel-spectrograms were computed using a 2048-point fast Fourier transform (FFT), 512-sample hop size, 128 mel bands, and a maximum frequency of 8 kHz.

This vocalization-derived OAAS baseline serves two complementary functions. First, it provides a biologically grounded coordinate reference for situating designed sound environments intended for porcine contexts. Second, it supports structured comparison between natural vocal expression and synthetic acoustic environments, enabling evaluation of whether sound-based interventions occupy regions associated with regulatory, affiliative, exploratory, or aversive acoustic configurations. These functions form the basis for the post-hoc OAAS evaluation of a veterinary functional music program presented in the following section.

Building on this vocalization-derived baseline, we next position the previously validated veterinary functional music stimuli in the joint OAAS embedding to evaluate their acoustic–affective organization relative to the biologically grounded anchor regions.

3.1.2. Vocal-only diagnostic OAAS and proto-atlas of contextual porcine vocal ensembles

To separate anchor topology from variance introduced by non-vocal stimuli, we additionally computed a vocal-only diagnostic OAAS by fitting the standardization and PCA model exclusively on the SoundWel contextual vocalization ensembles ($N = 36$; 18 contexts, two ~ 180 s ensembles per context) [3]. This yields an ethologically grounded reference manifold that is not influenced by functional music, challenge reference signals, or noise reference signals.

The resulting vocal-only embedding is summarized as a proto-atlas of contextual vocal centroids in Fig. 3 (vOAAS1–vOAAS3). We use the term proto-atlas to emphasize that this representation provides a first operational mapping of contextual vocal ensembles rather than an exhaustive atlas of porcine vocal affect; broader corpora and additional

contexts will be required for atlas-level coverage (see Limitations). Within this proto-atlas, affiliative/contact-related contexts (e.g., *After-Nursing*, *ShortReunion*) occupy regions opposed to sustained aversive contexts (e.g., *Castration*, *Restrain*), illustrating a structured contextual organization that emerges from acoustic features while remaining interpretable through the SoundWel emission metadata. The resulting vocal-only diagnostic embedding is shown in Fig. 3.

The resulting vocal-only proto-atlas (Fig. 3) visualizes this biologically grounded topology.

OAAS embedding constructed exclusively from labeled porcine vocalizations to establish a biologically grounded reference geometry. Each point represents an individual vocalization, grouped by emission-context category, allowing visualization of structural organization and relative positioning across the acoustic–affective space. This vocal-only embedding serves as a diagnostic anchor for subsequent joint embeddings, without implying categorical emotional assignment to non-vocal stimuli.

Notably, some conditioning paradigms (*PositiveConditioning* and *NegativeConditioning*) appear proximal in the reduced acoustic space, consistent with the SoundWel reference analysis that reports partial acoustic overlap despite differing contextual labels [3]. This supports a core OAAS principle: the framework does not impose categorical emotional quadrants, but reveals a continuous contextual acoustic topology that should be interpreted using biologically grounded anchors and distance-based operational metrics.

Finally, comparing the vocal-only proto-atlas (Fig. 3) with the vocal anchors visualized inside the joint embedding (Fig. 2) clarifies how anchor geometry is preserved while enabling operational positioning of

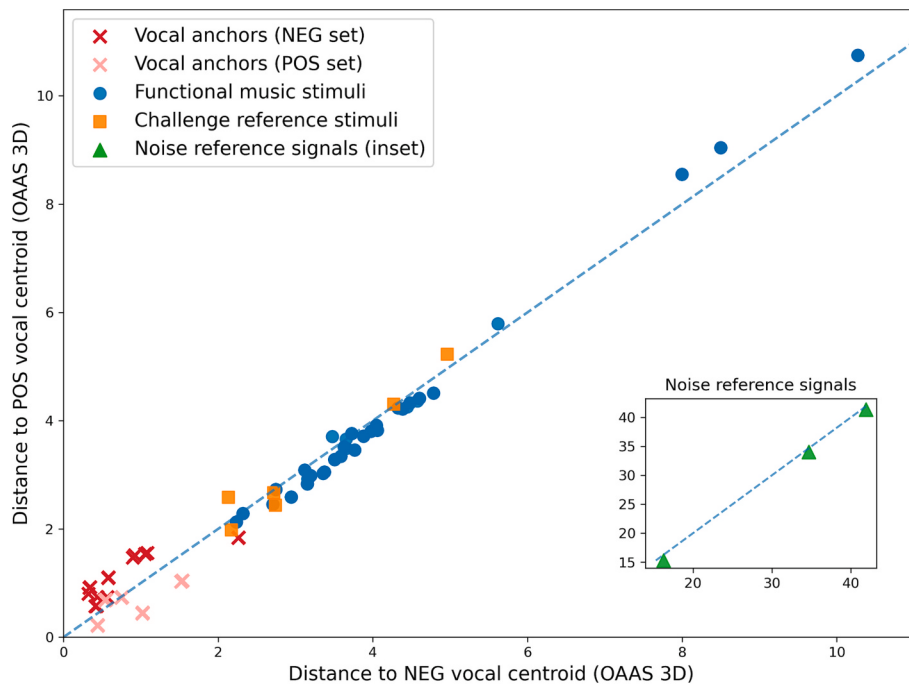


Fig. 5. Distance-based quantification of acoustic-affective proximity in OAAS. Distances between musical stimuli and POS/NEG vocalization centroids computed in the three-dimensional OAAS (PC1–PC3). The centroid-distance plane provides a quantitative representation of structural proximity, supporting stimulus comparison and ranking. These distance measures constitute the geometric basis for OAAS-derived acoustic-affective indices and subsequent operational analyses.

designed and reference stimuli in a shared OAAS coordinate system.

3.2. OAAS-based post-hoc evaluation of a veterinary functional music program

As a second case study, OAAS was applied as a post-hoc analytical framework to a previously validated veterinary functional music program developed for pigs [40]. The analysis focuses on the geometric positioning of the musical stimuli within the three-dimensional OAAS (PC1–PC3) defined by the porcine vocalization diagnostic baseline, enabling characterization of how individual stimuli and the program as a whole relate to vocalization-derived reference regions.

The OAAS was defined by the first three principal components derived from the standardized core acoustic feature set, which together explained 84.0% of the total variance (PC1 = 45.3%, PC2 = 28.7%, PC3 = 10.0%; cumulative PC1–PC2 = 74.0%). The retained PC1–PC3 representation was considered operationally sufficient because it preserved the dominant variance structure distinguishing the analyzed emission contexts while maintaining a globally interpretable Euclidean geometry suitable for centroid-distance computation, trajectory analysis, and constrained acoustic displacement operations. The proportion of explained variance for PC1–PC3 and the cumulative variance are reported in Supplementary Table S2, and the corresponding variable loadings are provided in Supplementary Table S3.

PC1 was primarily driven by informational and spectrotemporal complexity metrics, including temporal entropy of the amplitude envelope, multiscale entropy, spectral entropy, and harmonic ratio. Higher PC1 values were associated with increased acoustic irregularity and reduced harmonic structure, whereas lower values reflected more predictable and harmonically organized configurations. PC2 was dominated by spectral flatness and entropy-dispersion descriptors, capturing differences in broadband noise content and spectral organization. Higher PC2 values corresponded to noisier, less tonally defined environments, while lower values indicated greater spectral organization. PC3 captured finer-grained variation related to harmonic balance and entropy distribution, supporting discrimination among acoustically

similar stimuli occupying adjacent OAAS regions.

Audio materials corresponding to the musical stimuli included in the functional program were analyzed using the same OAAS parameter domains applied to porcine vocalizations. Feature extraction and normalization followed the procedures described in Section 2, ensuring that natural vocal expression, designed musical environments, and control stimuli were represented within a common analytical space.

Several musical stimuli designed as progressive variations of the same compositional strategy (W3_05, W5_05, W9_05) occupied virtually identical positions within OAAS. These stimuli were positioned near the vocalization-derived negative reference core in the joint embedding, indicating proximity in the reduced acoustic geometry. Distance-based positioning (Supplementary Table S7) was used to interpret this proximity with respect to the POS and NEG vocal centroids.

Noise reference signals were positioned far from both vocal-derived POS and NEG centroids, yielding very large absolute distances and near-zero POS–NEG bias (Supplementary Table S7; Table 5). These controls therefore functioned as out-of-range references in the joint embedding and did not systematically align with the sustained negative-context anchor region.

The positive anchor stimuli (W1_08, W10_08) were compositionally and timbrally related to the negative anchor family (W3_05, W5_05, W9_05), while differing in parameter configuration. Their OAAS positioning reflects displacement within a shared compositional strategy rather than a categorical change in sound type. This pattern is consistent with the intended design of the program and supports interpretation of OAAS as capturing parameter-level transformations across related stimuli.

OAAS-based Euclidean distances between the music-derived sound environments and the OAAS-POS/OAAS-NEG reference centroids, along with their differential distance ($\Delta = d_{POS} - d_{NEG}$), are summarized in Supplementary Table S7. For interpretability, each stimulus is additionally visualized in a centroid-distance plane plotting d_{NEG} (x-axis) against d_{POS} (y-axis) computed in the full 3D OAAS. Points below the identity line ($d_{POS} < d_{NEG}$) are closer to the POS vocal centroid, whereas points above the line are closer to NEG. Fig. 5 reports this centroid-

Table 5

Operational centroid-distance summary (OAAS 3D). Distances to the POS and NEG vocal centroids are reported for representative stimuli from each domain. Positive bias values indicate closer proximity to the POS centroid ($\text{pos_bias} = \text{dNEG} - \text{dPOS}$). The complete ranking across all stimuli is provided in Supplementary Table S7.

Stimulus	Domain	dPOS	dNEG	pos_bias	AAI_OAAS
W10_08.wav	Functional music stimuli	1.819	1.848	0.029	0.008
W2_05.wav	Functional music stimuli	1.997	1.963	-0.034	-0.009
W4_01.wav	Functional music stimuli	2.125	2.283	0.158	0.036
W8_03.wav	Functional music stimuli	2.261	2.508	0.247	0.052
Challenger_02_part1.wav	Challenge reference stimuli	1.681	1.769	0.088	0.025
Challenger_02_part2.wav	Challenge reference stimuli	2.114	2.311	0.197	0.045
Run_SW_02.wav	Vocalization anchors	0.233	0.681	0.448	0.489
Run_SW_01.wav	Vocalization anchors	0.237	0.684	0.447	0.485
ShortIsolation_SW_02.wav	Vocalization anchors	0.644	0.249	-0.395	-0.442
Huddling_SW_01.wav	Vocalization anchors	0.414	0.28	-0.134	-0.193
brown_noise_01_180s.wav	Noise reference signals	16.045	15.947	-0.099	-0.003
pink_noise_02_180s.wav	Noise reference signals	31.196	31.148	-0.048	-0.001
white_noise_03_180s.wav	Noise reference signals	38.621	38.585	-0.036	0.000

distance positioning, including a noise inset for out-of-range controls.

A compact view of the most POS- and NEG-biased items is provided in Supplementary Table S1, whereas the complete consolidated ranking across all stimuli (including noise references) is provided in Supplementary Table S7.

The spatial distribution of musical stimuli, porcine vocalizations, and reference signals within the joint OAAS is shown in Fig. 2B. Musical stimuli occupy a constrained region of the acoustic-affective space, partially overlapping with areas populated by porcine vocalization aggregates while remaining clearly separated from broadband noise reference signals. Table 5 summarizes representative centroid-distance values (OAAS 3D), and the complete ranking is provided in Supplementary Table S7.

3.3. OAAS operability: directed displacement of musical stimuli within the acoustic-affective space

Having established baseline topology and comparative positioning in the joint OAAS embedding, we next evaluate operability through controlled stimulus transformations directed toward predefined vocal-derived reference regions. Displacement was quantified using the centroid-distance metrics and bounded index described in Section 2 (Methods), enabling relative movement toward POS and NEG vocal centroids to be assessed within the same coordinate system.

To evaluate operability in practice, two musical stimuli from the veterinary functional music program (W7_01 and W8_03) were subjected to controlled acoustic transformations aimed at directed displacement toward predefined porcine vocalization-derived reference regions. Stimulus labels correspond to internal identifiers within the development program and are used here solely for traceability. An illustrative example of directed displacement (W7_01) is shown in Fig. 4, while comparative trajectories for both stimuli are provided in Supplementary Fig. S3.

3.3.1. Baseline position of musical stimuli in OAAS

For each stimulus, baseline position within the three-dimensional OAAS was quantified by calculating Euclidean distances to vocalization-derived reference centroids representing positively and negatively annotated emission contexts (Section 2.5). These baseline distances served as reference values for evaluating transformation performance (Fig. 4D for W7_01; Supplementary Fig. S3 and Supplementary Table S7 for both stimuli).

The two stimuli exhibited distinct initial configurations. While both occupied intermediate OAAS regions, W8_03 was initially located closer to the negatively annotated reference centroid than W7_01, indicating stimulus-dependent differences in baseline acoustic proximity within the shared embedding.

3.3.2. Directed acoustic transformations and distance modulation

For each stimulus, multiple candidate transformations were generated under controlled parameter constraints targeting displacement toward either the positive (POS) or negative (NEG) vocalization-derived centroid. From each candidate set, the variant yielding the minimal distance to the target centroid was selected for analysis.

Positive-directed displacement (POS). For both W7_01 and W8_03, POS-directed transformations reduced Euclidean distance to the positive reference centroid. For W7_01, distance decreased from 16.73 to 16.61; for W8_03, distance decreased from 14.89 to 14.76. The displacement trajectory for W7_01 is illustrated in Fig. 4D, while comparative centroid-distance outcomes are reported in Supplementary Fig. S3 and Supplementary Table S7.

Negative-directed displacement (NEG). NEG-directed transformations showed stimulus-dependent behavior. For W7_01, distance to the negative reference centroid decreased from 16.67 to 16.52 (trajectory shown in Fig. 4D). In contrast, for W8_03, NEG-directed transformations did not reduce distance; the best candidate increased the distance from 14.83 to 15.66 (Supplementary Fig. S3; Supplementary Table S7).

3.3.3. Asymmetry and saturation effects in OAAS navigability

The contrasting NEG-directed behavior indicates asymmetric navigability within the acoustic-affective space under the applied transformation constraints. While POS-directed displacement produced distance reductions for both stimuli, NEG-directed displacement was constrained for W8_03.

W8_03 was already positioned near the negative reference centroid at baseline, and further distance reduction was not achieved under the imposed acoustic feature bounds. This behavior is consistent with a saturation-like effect within the local OAAS geometry rather than a failure of the transformation procedure. The illustrative trajectory of W7_01 is shown in Fig. 4D, and the comparative pattern across both stimuli is documented in Supplementary Fig. S3.

3.3.4. Structural robustness of vocal anchors and centroid geometry

To determine whether the observed navigability asymmetry reflected a geometric distortion of the joint embedding or an intrinsic property of the vocal anchor structure, we performed a complementary robustness analysis using a vocal-only OAAS configuration.

In this configuration, PCA and centroid definitions were derived exclusively from SoundWel vocalization macrosegments ($n = 36$). Musical stimuli, challenger references, and noise baselines were subsequently projected into this vocal-derived coordinate system, isolating the geometric contribution of biologically grounded anchors from potential cross-domain embedding effects.

Centroid separation and within-cluster dispersion were quantified in the full three-dimensional OAAS space (PC1–PC3). The Euclidean distance between POS and NEG vocal centroids was 1.593. Within-cluster

dispersion (mean Euclidean distance of items to their respective centroid) was 2.115 for POS and 2.678 for NEG; in both cases, centroid separation was smaller than within-cluster dispersion (separation > dispersion = False), indicating substantial geometric overlap between the full POS and NEG context sets. Centroid metrics are therefore interpreted comparatively within a shared embedding rather than as categorical separators.

This overlap is consistent with the graded and heterogeneous nature of affective emission contexts in natural vocal communication, which do not form linearly separable clusters. Importantly, when musical stimuli (including W7_01 and W8_03) were projected into the vocal-only OAAS, their relative ranking and distance-based positioning remained consistent with the joint embedding results, and extreme outliers continued to correspond primarily to broadband noise references.

These findings indicate that the asymmetry observed in NEG-directed displacement is not an artifact of cross-domain embedding bias. Rather, it reflects intrinsic geometric properties of the vocal anchor distribution combined with feature-bound constraints imposed during directed navigation. Together, the vocal-only robustness analysis supports the stability of centroid-based metrics (dPOS, dNEG, pos_bias, and AAI_{OAAS}) while clarifying that apparent asymmetry arises from graded dispersion within biologically grounded affective contexts rather than from embedding distortion.

3.3.5. Preservation of musical identity under OAAS-guided modulation

Across all transformations, global musical organization (macro-structure, duration, and compositional identity) was preserved. Directed displacement was primarily associated with fine-grained acoustic modifications—such as redistribution of spectral energy and temporal micro-structural adjustments—while the overall compositional structure remained unchanged (Fig. 4A–B for W7_01; comparative visualization in Supplementary Fig. S3). These modifications produced measurable changes in OAAS coordinates and centroid-distance metrics following re-projection into the fixed joint OAAS model (Fig. 4D for W7_01; Supplementary Fig. S3 and Supplementary Table S7 for both stimuli). In addition to geometric displacement within OAAS, directed transformations were consistently associated with systematic modulation of microdynamic structure. Under RMS-preserving normalization, POS-directed variants exhibited reduced crest factor and narrower high-percentile amplitude dispersion, whereas NEG-directed variants showed increased crest factor and expanded percentile dispersion (see Supplementary Table S8). These differences indicate that OAAS-guided navigation entails structured reorganization of transient dynamics beyond simple level adjustment.

3.3.6. Summary of OAAS operability

Together, these results show that controlled transformations can produce quantifiable displacement of designed sound environments relative to vocal-derived reference centroids within the joint OAAS embedding. The illustrative case of W7_01 is shown in Fig. 4, demonstrating spectro-temporal modification patterns and directional OAAS displacement. Comparative POS- and NEG-directed outcomes for both W7_01 and W8_03, including centroid-distance modulation and trajectory asymmetry, are provided in Supplementary Fig. S3 and Supplementary Table S7.

3.3.7. Proximity of musical stimuli to porcine vocalization categories

To further characterize acoustic positioning, Euclidean distances were computed between each musical item and the centroids of the porcine vocalization categories defined within OAAS. Distances were computed in the full three-dimensional OAAS (PC1–PC3), enabling identification of which categories were acoustically closest to, and most distant from, each stimulus. These distance relationships are summarized in Fig. 5 and provide the operational bridge between acoustic structure and decision-oriented ranking.

3.3.7.1. Interpretation of centroid-distance patterns. The centroid-distance analysis revealed stimulus-specific patterns. For W7_01, the closest categories corresponded to regulated and exploratory emission contexts, whereas categories associated with sustained negative contexts were among the most distant. In contrast, W8_03 exhibited greater proximity to high-activation categories, consistent with its denser textural and informational configuration.

To resolve ambiguity in two-dimensional OAAS projections and quantify proximity in the full embedding, we computed three-dimensional Euclidean distances (PC1–PC3) between each program stimulus and the porcine vocalization-category centroids. The complete stimulus-by-category distance matrix is reported in Supplementary Fig. S1, and a family-level summary (mean $\pm$ SD across W1–W10) is provided in Supplementary Fig. S2. Consolidated centroid-distance metrics to POS/NEG reference centroids for all stimuli are provided in Supplementary Table S7.

Together, these distance-based analyses establish the operational and geometric consistency of the OAAS embedding, providing the foundation for the broader implications discussed below.

4. Discussion

The present study introduces the OAAS as an applied framework for the analysis, evaluation, and intentional design of affective sound environments. Conceptually, synthetic acoustic environments are treated here as structured acoustic systems whose perceptual organization can be examined and compared within a multidimensional representational space. This perspective is consistent with the ISO soundscape definition and with reviews that frame acoustic environments through perceptual appraisal and contextual interpretation [9,15,17]. OAAS extends this tradition by operationalizing acoustic structure into a reproducible coordinate geometry anchored to biologically grounded vocal reference signals. Rather than reporting new behavioral or physiological experiments, the results demonstrate how affect-relevant acoustic structure constrains and enables design decisions when vocal anchors, noise references, and designed sound environments are positioned within a shared embedding.

Recent studies have explored the mapping of speech signals into dimensional affective spaces by directly predicting activation–valence coordinates from acoustic embeddings using supervised learning approaches (e.g., x-vector representations with regression models; see [32]). Such frameworks aim to infer affect coordinates from acoustic input and have proven valuable within human speech emotion research. In contrast, OAAS does not aim to infer ground-truth affect coordinates from acoustic input. Instead, it constructs a shared acoustic–geometric embedding in which biologically contextualized anchors define interpretable reference regions. The emphasis is placed on comparative positioning, distance structure, and constrained navigability across heterogeneous sound environments—an especially practical approach in cross-species contexts where psychological labels are not directly observable. Accordingly, OAAS should be interpreted as an operational geometry for design and comparison, not as a model that assigns affect labels to stimuli.

OAAS builds on a cumulative empirical trajectory from our group, including: (1) a review framing music as a welfare-relevant enrichment tool in non-human animals and emphasizing the need for affect-informed design/adaptation grounded in behavioral, physiological, and neurobiological indicators [37]; (2) evidence that pigs exposed to original musical constructions exhibit differentiated emotional responses quantified through QBA [39]; (3) mechanistic links between these responses and engineered spectro-temporal configurations [38]; (4) applied evidence that adapted original music can reduce aggression during regrouping [2]; (5) validation of a long-duration functional music program with measurable psychophysiological effects related to chronic stress [40]; and (6) farm-based lactation evidence showing favorable effects on maternal behavior and reduced pre-weaning piglet

mortality [23]. Together, these studies support OAAS as a translational strategy that links affect-oriented design choices to observable behavioral and physiological outcomes under controlled and field conditions.

The explicit separation between acoustic operation (OAAS) and affective interpretation (QBA, here used as an observational affect vocabulary and discussed in relation to reduced valence–arousal framing) supports transdisciplinary use of the framework. OAAS provides an engineering-oriented geometry for positioning and constrained navigation (coordinates, distances, trajectories), while QBA provides a validated ethological assessment layer for describing affect-relevant organismal state in intervention contexts [39]. In this reading, QBA-derived descriptors can inform the interpretation of OAAS regions and trajectories when behavioral evidence is available, without treating OAAS coordinates themselves as affect ratings. This complementary layering preserves biological interpretability while keeping the acoustic geometry operationally well-defined.

In the present implementation, we additionally explored a PAD-oriented projection derived from QBA-aligned anchors as an interpretive layer. Under cross-domain projection, this mapping showed attenuated dimensional differentiation, a behavior expected when dimensional anchors are indirectly specified from categorical emission-context descriptors and when within-anchor dispersion—particularly in high-activation contexts—approaches or exceeds between-anchor separation. For this reason, PAD-based coordinates are retained here as a supplementary interpretive overlay, whereas the primary analyses rely on centroid-distance geometry and the bounded acoustic index AAI_{OAAS} , which demonstrated greater stability under anchor-sensitivity controls.

OAAS findings also align with established auditory constraints in pigs. Behavioral audiogram data indicate that pigs exhibit a broad hearing range with highest sensitivity in mid-frequency bands, overlapping with the dominant energy and modulation structure of many porcine vocalizations [14]. This physiological profile supports the ecological plausibility of vocal-derived anchoring in OAAS and motivates future extensions incorporating audiogram-weighted descriptors and calibrated playback-level safeguards during intervention deployment.

Across the case studies presented in Section 3, OAAS exhibited consistent structural behavior that supports its use as a design- and evaluation-oriented framework. The contribution is not a new affect taxonomy nor a replacement for welfare-oriented classification approaches. Instead, it is a complementary representational strategy: OAAS formalizes a geometric domain in which heterogeneous sound environments can be positioned, compared, and intentionally displaced relative to biologically grounded reference regions. This addresses a recurring limitation highlighted in recent animal-centered music reviews—namely the lack of reproducible, design-oriented formalisms linking intrinsic stimulus structure to interpretable affect-relevant objectives across heterogeneous protocols—by providing an explicit coordinate geometry for comparison and constrained navigation [19,20].

The vocal-only robustness analysis further clarifies the geometric behavior of OAAS. Although full POS and NEG vocal anchor sets exhibited centroid separation values smaller than within-cluster dispersion, this overlap does not indicate instability of the framework. Rather, it reflects the graded and heterogeneous nature of affective emission contexts in natural vocal communication.

In biological systems, affective signaling in vocal communication is rarely organized into strictly separable acoustic categories; instead, it varies along continuous spectro-temporal dimensions shaped by activation level, context, and communicative function.

The proximity observed between some contextually opposed emission conditions in the vocal-only proto-atlas further illustrates this graded organization. In particular, conditioning paradigms labeled as PositiveConditioning and NegativeConditioning appear relatively close in the reduced acoustic space. This pattern is consistent with the original SoundWel analyses, which also reported partial acoustic overlap between these contexts despite their different behavioral interpretations.

Such proximity reflects the fact that conditioning paradigms share common acoustic substrates related to arousal and communicative activation, even when the contextual valence differs. Within OAAS, these relationships reinforce the interpretation of the space as a continuous acoustic topology rather than a categorical emotional map [4]. The observed overlap between POS and NEG anchor regions is therefore consistent with the expectation of graded acoustic coding across contexts and suggests that OAAS captures realistic structural distributions rather than imposing artificial categorical boundaries.

Importantly, the persistence of geometric relationships when projecting stimuli into a vocal-only OAAS confirms that the asymmetry observed during NEG-directed displacement is not a cross-domain embedding artifact. Instead, it reflects intrinsic properties of the anchor distribution combined with bounded acoustic transformation constraints. This indicates that OAAS behavior is driven by the vocal-anchor structure rather than by embedding bias.

The larger dispersion observed within the NEG vocal anchor set compared to the POS set may suggest internal substructure within negatively annotated emission contexts. Rather than constituting a single homogeneous region, negative vocalizations can encompass distinct acoustic configurations associated with acute high-activation distress (e.g., nociceptive contexts) and sustained or socially mediated stress states (e.g., isolation- or frustration-related emissions), as reflected in context-annotated porcine vocal corpora used here for anchoring [3]. While the present study does not formally partition these regions, the geometric dispersion pattern indicates that negative affective anchoring may be multidimensional rather than unipolar. Future work could test whether refined anchor subsets yield increased centroid separability and improved navigability symmetry.

More broadly, animal vocal emotion research reports that acoustic signatures associated with acute, high-activation distress can differ from those linked to sustained social stress or chronic activation, supporting the possibility that negatively annotated emission contexts are organized along more than a single acoustic axis [3,4]. The framework therefore offers a potential platform for disentangling distinct forms of negative activation in future work, without presupposing categorical boundaries.

Despite substantial progress in soundscape research and affective approaches to environmental appraisal—particularly in defining perceptual frameworks and emotional dimensions of environmental sound [15,17]—the field has comparatively fewer operational tools for intervention-oriented design grounded in explicit acoustic geometry. In practice, heterogeneous sound environments are rarely evaluated within a shared operational space that supports standardized geometric descriptors—such as centroids, distances, trajectories, and displacement—under common acoustic representations. This limits cross-study comparability and complicates the formalization of sound-based interventions as controllable affective systems with defined target regions and quantifiable directed change. OAAS is proposed to address this gap by enabling interpretable positioning and constrained navigation across biologically grounded reference regions and designed sound environments.

4.1. OAAS as a framework for affective sound environment design and evaluation

OAAS operationalizes synthetic acoustic environments as affective systems by treating time-extended sound environments as structured, measurable, and comparable configurations—rather than as isolated auditory exposures. Within this framing, synthetic acoustic environments are understood as intentionally constructed soundscapes designed to shape affective and regulatory conditions over time through integrated spectral, temporal, dynamic, and informational organization. This definition is consistent with soundscape standards that emphasize context-dependent perceptual appraisal [17] and with ecological approaches that treat acoustic environments as integrated systems in

which structure and function co-vary [8,27].

Recent interdisciplinary reviews explicitly call for improved reproducibility, protocol harmonization, and transparent reporting of intrinsic stimulus properties and playback parameters in animal-centered music research [19,20]. OAAS can be read as an operational response to these priorities: by defining a standardized acoustic representation, explicit coordinate geometry, and biologically grounded anchoring, it enables consistent positioning and quantitative comparison across heterogeneous sound environments. OAAS-based design and verification workflows also encourage systematic characterization of intrinsic acoustic properties and transformation effects, aligning with reporting priorities emphasized in these reviews while extending them into a navigable design space.

Rather than prescribing what these environments should sound like (music, noise, hybrid textures, or other forms), OAAS provides a unified basis for quantitative comparison and evaluation across sound sources. This supports interpretable auditing of intervention structure and enables operational analyses based on geometric relationships (e.g., centroids, distances, displacement), linking acoustic characterization directly to design-oriented reasoning under real-world constraints [7,20]. This framing also clarifies why OAAS can incorporate biological anchoring without reducing affect to categories: affective plausibility is approached through graded structural alignment in a shared coordinate space rather than categorical assignment [3,37].

4.2. Biologically grounded anchoring without category reduction

Using porcine vocalizations as an anchoring reference grounds OAAS in natural sound production rather than in purely conceptual affect labels. In this study, anchoring is used for interpretability and operational reference—rather than re-validation of SoundWel. SoundWel functions as a biologically grounded substrate that populates OAAS with reference regions derived from documented emission contexts and valence-related descriptors, enabling interpretation without imposing categorical affect structure on the space [3,4,21].

This distinction matters because classification-oriented approaches and OAAS address different questions. Classification strategies target context-aware recognition and welfare-oriented monitoring of vocal outputs, often emphasizing predictive performance and robustness under field conditions [3,5,36]. Robustness across farms can remain limited by environmental variability and behavioral diversity [25], so predictive accuracy alone does not resolve the representational need for cross-context comparability when interventions must be evaluated across heterogeneous acoustic conditions. OAAS complements these perspectives by providing an operational geometric representation: rather than predicting categories, it supports comparison and design by quantifying proximity, separation, and displacement between heterogeneous sound environments and vocalization-derived reference regions under a standardized acoustic representation [19,20].

A related methodological point concerns the role of QBA alignment. The purpose of aligning vocal emission contexts with QBA-aligned adjective labels is interpretability and methodological continuity—not categorical equivalence. QBA provides a validated observer-based bridge between acoustic interventions and organismal state in pigs [38,39]. Vocal emission contexts provide biologically grounded reference signals, and their conceptual alignment supports coherent interpretation across datasets and methods without treating acoustic proximity as a stand-alone affect inference mechanism [40].

OAAS interpretation is strongest when geometry is read together with emission-context descriptors and, where available, external behavioral or physiological validation, rather than treated as a stand-alone mechanism [3,19]. Comparative neurophysiological evidence also indicates that mammalian brains can exhibit both general and conspecific sensitivity to vocal sounds, supporting the use of vocalizations as biologically meaningful anchors for acoustic modelling [24].

Consistent with this non-inferential stance, OAAS coordinates are

defined by quantitative acoustic descriptors and biologically grounded anchoring rather than by affect ratings. The framework therefore locates sound environments in a reproducible feature geometry that supports structural comparison and constrained navigation across reference regions. Affective interpretation should be conducted through joint reading of OAAS geometry, emission-context descriptors, and independent behavioral or physiological evidence, consistent with soundscape frameworks where appraisal depends on listener and context [9,17].

4.3. From music to functional sound environments

A persistent challenge in animal-centered sound research is the conceptual and terminological ambiguity between “music,” “soundscape,” and broader categories of synthetic acoustic environments. Recent interdisciplinary reviews explicitly note that the field still lacks an operational definition of “music” in non-human contexts and, as a consequence, struggles to establish reproducible standards for stimulus design and evaluation [19,20]. OAAS addresses this limitation by treating music not as a privileged category, but as one possible implementation of an affective sound environment.

Within the present framework, music is understood as an intentional soundscape: a constructed acoustic micro-ecosystem in which spectral, temporal, dynamic, and spatial relations interact over time to generate coherent perceptual and affective contexts. This interpretation is consistent with acoustic ecology perspectives that frame soundscapes as organized communicative environments and treat music as an intentional reorganization of environmental and communicative sound patterns [29,33]. Accordingly, within OAAS the relevant distinction is not whether a stimulus qualifies as “music” in a cultural sense, but whether its acoustic organization functions as a coherent affect-oriented system aligned with regulatory objectives and how that organization maps onto biologically grounded reference regions.

This framing also addresses a practical limitation repeatedly highlighted in the literature: many animal music studies rely on pre-existing human repertoires selected without systematic control of acoustic structure, species-perceptual accessibility, or functional objectives [19,20]. In response, our previous work introduced the concept of “veterinary functional music”: music composed or selected for a defined welfare-related function, empirically validated in a target species, and engineered through acoustic and musical adjustments adapted to the sensory and perceptual characteristics of that organism [2,39,40]. This concept prioritizes function, empirical validation, and acoustic design over cultural origin or human aesthetic convention.

Importantly, the musical material analyzed in the present OAAS evaluation does not represent arbitrary musical repertoires. The stimuli belong to a veterinary functional music program previously designed and experimentally evaluated under animal-welfare objectives. Within this program, compositional and production strategies were intentionally engineered to attenuate spectro-temporal patterns associated with high-distress vocalizations, including abrupt temporal irregularity, excessive broadband spectral energy, and unstable microdynamic profiles. Consequently, the positioning of these stimuli within OAAS should be interpreted as the acoustic footprint of a deliberately engineered sound environment rather than as a property of “music” as a general stimulus category. This distinction demonstrates that affect-oriented positioning in OAAS emerges from controlled acoustic design decisions rather than from the mere presence of musical structure.

OAAS extends this functional perspective into a reproducible design and evaluation geometry. By positioning both vocalizations and constructed sound environments within a shared acoustic-affective coordinate space, OAAS enables functional music—and other engineered sound environments—to be analyzed and calibrated as systems with measurable trajectories, distances, and constrained transformations. This approach supports the development of affective sound environments suitable for enrichment, interspecies-compatible shared soundscapes, or aversive acoustic configurations when deterrence or risk

avoidance is the objective [19,20].

This perspective is particularly relevant in animal-centered work because auditory perception is adapted to ecological contexts and to the acoustic structure of environments in which vocalizations naturally occur. Vocalizations therefore serve as ecologically grounded anchors reflecting affect-relevant communication under real-world constraints [3,4]. Accordingly, OAAS supports a conceptual shift away from categorical assumptions about “species-specific music” toward an operational view of sound as an engineered ecological medium whose organization can be designed to support regulation, engagement, resilience, or avoidance depending on context and objective.

4.4. Noise, roughness, entropy, and the limits of broadband reference signals

A notable outcome concerns the relationship between broadband noise and affective interpretation. The PCA loading structure (Supplementary Table S3) shows that spectral flatness, entropy-related variables, and reduced harmonic organization contribute strongly to the principal dimensions associated with peripheral broadband noise positioning. Although noise-based reference signals are often recommended in animal-centered sound research—partly because they may mask environmental noise and reduce confounds when comparing musical stimuli across settings [19]—our OAAS analyses suggest that broadband noise should not be assumed to be a perceptually neutral control condition. Instead, broadband noise can function as an active acoustic stimulus with its own perceptual and regulatory effects.

Evidence from human studies supports this interpretive caution: white noise has been associated with changes in autonomic markers (including HRV) during short listening exposures, and stochastic noise embedded into slow-tempo music has been reported to enhance autonomic coherence through mechanisms discussed in relation to stochastic resonance [6,26]. Together, these findings motivate treating broadband noise not only as a masking baseline, but as a candidate functional acoustic intervention whose effects should be empirically verified.

In our lactation study, pink noise and veterinary functional music were evaluated as separate treatments relative to a no-stimulation control condition, and both interventions showed favorable effects on maternal behavior and piglet outcomes [23]. These results further support the view that noise-based stimulation can produce measurable functional effects independent of any masking role, and therefore merits treatment as an experimental condition rather than a trivial baseline.

Porcine vocalizations—particularly in high-activation contexts—often contain noise-like components, high spectral density, and irregular temporal structure. However, such “noisy” biological signals are not equivalent to unstructured broadband noise. Vocalizations exhibit informational organization shaped by production constraints and communicative function, including non-random temporal patterning and structured variability across call types and emission contexts [3].

Consequently, acoustic roughness or broadband energy in vocal signals should not be equated with unstructured noise. Instead, such features typically emerge within constrained spectro-temporal organizations shaped by biomechanical production mechanisms and communicative function. In this sense, affect-relevant roughness is not only a matter of spectral density, but of structured irregularity.

Such irregularity can be operationalized through entropy-related descriptors included in the OAAS operational feature set (e.g., temporal entropy, spectral entropy, multiscale entropy, dispersion profiles). This emphasis on multi-parameter structure aligns with evidence that combinations of acoustic features encode robust communicative contrasts across contexts in mammalian vocal communication [12]. OAAS captures this distinction by integrating descriptors that reflect both spectral texture and informational structure, enabling unstructured broadband noise to be differentiated from biologically grounded irregularity.

From an applied perspective, this implies that noise-based sound environments should not be evaluated solely by spectral profile; temporal and informational organization are also critical for interpreting likely perceptual impact and for assessing proximity to vocalization-derived reference regions. The present study did not systematically differentiate between broadband profiles (e.g., white vs. pink vs. brown), which remains a limitation and an opportunity for future work. Within this context, OAAS provides a structured analytical framework for examining how different broadband structures—and their entropy-related signatures—populate the acoustic-affective space and how they may be operationally used, calibrated, or avoided in the design of functional affective sound environments.

4.5. Post-hoc evaluation: functional organization over stylistic identity

OAAS-based post-hoc evaluation showed that stimuli sharing compositional and production strategies converge within the acoustic-affective space despite differences in musical surface features. This suggests that OAAS captures functional acoustic organization rather than stylistic identity, which is particularly relevant in applied settings where interventions may be implemented as music, hybrid textures, or engineered noise-like structures, yet remain comparable in their spectral-temporal and informational organization.

Stimuli positioned near vocalization-derived regions linked to high activation should not be interpreted as inherently aversive. Instead, proximity reflects shared structural properties such as spectral density, temporal irregularity, and informational load, which can occur both in affect-relevant vocal communication and in designed environments that increase acoustic complexity. This motivates a core interpretive rule of OAAS: geometric proximity does not imply identity of emotional experience. Rather, it indicates graded structural alignment within a shared coordinate system that can inform design decisions and interpretation—particularly when read alongside emission-context descriptors, listener constraints, and external validation evidence.

This interpretive stance is consistent with soundscape and affective appraisal approaches, in which acoustic features contribute to affective plausibility but do not fully determine perceived emotion or welfare outcomes; appraisal remains context-dependent and mediated by listener characteristics and playback conditions [9,17]. Accordingly, OAAS functions as a representational geometry that supports reproducible structural comparison across heterogeneous interventions—while preserving conceptual openness regarding affective interpretation—responding to reproducibility and comparability priorities emphasized in recent animal-centered music reviews [19,20].

4.6. Operability and constraints: OAAS as a constrained control space

The directed displacement experiments demonstrate that OAAS is not merely descriptive but operational. Controlled transformations enabled intentional navigation toward predefined reference regions, quantified geometrically within the acoustic-affective coordinate space. Positive-directed displacement was consistently achievable, whereas negative-directed displacement showed stimulus-dependent saturation when stimuli were already positioned near NEG reference regions.

The observed saturation of NEG-directed displacement should therefore be interpreted as a constrained navigability effect rather than as evidence of a purely topological limit. In OAAS, navigability depends jointly on the local distribution of biological anchors, the baseline acoustic organization of the stimulus, and the explicit transformation constraints imposed to preserve acoustic identity and level stability. Consequently, W8_03 saturation may reflect the interaction between its initial proximity to the NEG region, the heterogeneous dispersion of negative vocal-anchor contexts, and the bounded RMS-preserving, peak-safe transformation regime.

This asymmetry is informative about both the stimulus set and the feasible transformation manifold. The functional music program was

authored under welfare-oriented constraints, intentionally avoiding strongly aversive acoustic extremes. As a consequence, stimuli concentrate within a comparatively “safe” subregion of the joint OAAS, and local navigability becomes anisotropic under identity-preserving production constraints. NEG-directed displacement may therefore saturate near manifold boundaries because further movement would require coupled shifts in entropy, spectral balance, and microdynamic structure that exceed admissible ranges or compromise compositional coherence.

From a geometric perspective, this behavior can also be interpreted as the emergence of local stability gradients within the acoustic-affective space. Regions associated with regulatory or affiliative vocal contexts tend to exhibit relatively compact acoustic distributions, whereas negatively annotated contexts often display broader dispersion associated with heterogeneous high-activation states. Consequently, navigability within OAAS may not be isotropic: displacement trajectories depend not only on transformation constraints but also on the local topology of the anchor distribution. This property supports interpreting OAAS as a structured acoustic landscape with variable navigability rather than as a uniform metric space.

These observations provide a geometry-based audit of feasible transformations under identity-preserving production constraints, consistent with prior physiological and behavioral validation evidence for the authored program [40]. They also motivate future validation studies using broader stimulus sets spanning a larger portion of the OAAS and tested under controlled playback conditions (e.g., QBA-linked outcomes), enabling direct assessment of navigability symmetry and operational predictiveness.

Comparable displacement-based control spaces remain uncommon as explicit formalisms in animal-centered affective sound environment design. Accordingly, OAAS operability is best interpreted as a testable, assumption-bound control notion: it specifies a constrained set of transformations and measurable displacement predictions whose empirical adequacy depends on stated acoustic constraints, playback conditions, and listener context—an approach consistent with experimental sound-use frameworks emphasizing hypothesis discipline and falsifiability under deployment limits [28]. Related analogies also exist in applied acoustics and engineering, where compact representations paired with constraint-aware mechanisms are used to maintain reliable behavior under bounded operating regimes [13].

In this sense, OAAS operability should be understood as “constrained navigability”: measurable displacement within a feature geometry while respecting acoustic plausibility, listener constraints, and parameter coherence. Saturation effects therefore function as empirical indicators of the feasible transformation manifold, supporting OAAS as a structured control space for functional affective sound environment design rather than as an unconstrained generative system.

4.7. Texture-level modulation with structural preservation

Across transformations, OAAS-guided modifications preserved higher-level musical organization while modulating fine-grained acoustic texture (spectral redistribution, granularity, and temporal micro-structure). This supports an applied principle: affect-oriented sound environment design does not necessarily require reauthoring musical content, but can be approached through controlled modulation of acoustic materiality while maintaining structural coherence.

Consistent with this texture-level modulation, RMS-preserving transformations revealed systematic changes in microdynamic structure. POS-directed variants exhibited reduced crest factor and narrower high-percentile amplitude dispersion, whereas NEG-directed variants showed increased crest factor and expanded percentile dispersion (Supplementary Table S8). These effects indicate that OAAS-guided navigation reorganizes transient dynamics in a directionally structured manner while maintaining overall energy stability, reinforcing that affective displacement can be achieved through microdynamic reconfiguration rather than large-scale compositional alteration. Whether such

microdynamic reorganization carries cross-modal ecological relevance remains an open empirical question.

Importantly, recent farm-based evidence indicates that both functional music and pink noise can yield favorable outcomes relative to a no-stimulation control in lactating sows, reinforcing that intervention effects should be interpreted against the actual baseline acoustic environment of the farm rather than an implicit “silence” condition [23]. OAAS provides a framework to quantify such baseline-to-intervention shifts and to compare distinct intervention strategies within a shared coordinate space.

A key implication is that the functional positioning of a sound environment may depend critically on production-level processes, even when the underlying musical composition remains unchanged. The same piece can occupy different OAAS regions depending on timbral balance, spectral density, loudness structure, and micro-dynamic organization introduced during synthesis, mixing, and post-production [37,38]. This has methodological relevance for animal-centered sound research, where stimuli are often drawn from existing human repertoires without systematic control of acoustic structure or perceptual accessibility, potentially introducing uncontrolled variance in affective outcomes [19].

From an applied welfare and health perspective, this implies that sound-based interventions benefit from quality-control procedures analogous to those used in other biological treatments, where dosage, formulation, and reproducibility are essential. OAAS can serve as a calibration and verification tool for production-sensitive configurations, enabling sound environments to be treated as operational affective systems rather than static musical artifacts [20,40].

More generally, this supports the operational framing of synthetic acoustic environments as affective systems: the objective is not to produce a categorical type of sound, but to shape how organized acoustic structure supports regulation, engagement, and affective plausibility under ecological constraints [9,17].

4.8. Datasets as research infrastructure and framework generalization

The present work highlights the role of shared audio datasets—such as vocalization corpora and environmental recordings—as research infrastructure. Such resources allow multiple methodological approaches to coexist, including welfare-oriented classification and monitoring, mechanistic feature-based analysis, and design-oriented synthesis. OAAS is intended to leverage these datasets to support design and evaluation workflows in a transparent and reproducible way.

In the specific case of the porcine anchoring model, the SoundWel project provides a labeled reference corpus of pig calls recorded in clearly defined contexts and organized along valence-related descriptors [3]. Its public availability as a versioned dataset resource further supports reproducibility and interpretability by allowing OAAS reference regions to be grounded in biologically meaningful emissions rather than abstract affect labels [31].

Although this study uses porcine vocalizations as a reference model, OAAS is not species-specific in scope. The porcine model provides a controlled test case for evaluating how the framework behaves under realistic acoustic and affective constraints. The methodological logic underlying OAAS is transferable and can be extended to other species and applied domains where objective measures of sound structure and response indicators are available. Ongoing applications to additional vocalization corpora (including bovine and human datasets) further support this cross-domain extensibility and will be reported separately to preserve methodological clarity.

From a methodological perspective, the combination of open datasets and standardized acoustic representations provides a pathway toward cumulative and comparable research. As animal-centered sound interventions expand beyond exploratory demonstrations, shared repositories become critical for benchmarking, replication, and cross-context generalization. In this sense, OAAS can be understood as a

design-oriented complement to dataset-driven monitoring approaches: it offers a geometric infrastructure that enables heterogeneous stimuli (music, noise, environmental recordings, vocalizations) to be compared and calibrated within a common operational space.

More broadly, this dataset-centered paradigm supports framework generalization: as additional corpora become available—across species, contexts, and recording conditions—OAAS can incorporate new anchoring regions and expand its representational coverage. This suggests a future research infrastructure in which vocalization datasets and soundscape repositories function not only as monitoring resources but also as calibration layers for functional affective sound environment design.

Within this framework, OAAS-directed displacement should be interpreted as an operational reorganization of acoustic-affective structure within a constrained geometric space rather than as direct evidence of equivalence in subjective or biological affective experience. Accordingly, centroid convergence represents functional acoustic proximity under the selected descriptor system, while downstream behavioral and physiological validation remains necessary to evaluate species-specific efficacy under real-world conditions.

This operational geometry may also support welfare-oriented acoustic intervention design by enabling identification and controlled navigation of positively anchored acoustic regions associated with affiliative or regulatory vocal contexts.

4.9. Scope, limitations, and future directions

This article is framed as an operational acoustics contribution. OAAS formalizes a reproducible acoustic geometry for positioning, comparison, and constrained transformation of heterogeneous sound environments relative to biologically grounded reference anchors. OAAS is complementary to welfare-oriented vocalization classification and monitoring pipelines (e.g., SoundWel) [3]. Such approaches primarily target diagnostic inference—detecting and classifying affect-relevant state and/or emission context under field variability and noise—whereas OAAS focuses on geometry for acoustic design, evaluation, and operability anchored in biologically meaningful reference sounds. In this sense, OAAS also supports an interspecies acoustic design perspective, i.e., a communication-oriented strategy in which functional sound environments are engineered relative to species-grounded affective vocal reference regions rather than human-centric musical heuristics.

Scope and validation pathway. The present study does not introduce new behavioral or physiological experiments because the functional music program analyzed here belongs to a line of work previously evaluated under production-relevant conditions using behavioral (QBA) and physiological/health-related indicators [40]. OAAS therefore contributes a complementary engineering layer for auditing acoustic organization, comparability, and operability, rather than re-testing intervention effects *in vivo*.

Proto-atlas and dataset extensibility. The vocal-only diagnostic embedding should be read as a first operational proto-atlas. In this work, SoundWel emission contexts are operationalized into time-extended ensembles (~180 s; two per context; $N = 36$ for 18 contexts) to ensure temporal comparability with longer designed stimuli and to obtain stable context-level acoustic profiles for anchoring [3]. This strategy does not reduce the underlying repository to “a handful of calls”; instead, it defines a practical anchoring layer that can be replaced or expanded as new corpora become available.

Centroid stability considerations. The macrosegment aggregation strategy and the vocal-only robustness analyses presented in this study provide indirect support for the stability of the OAAS reference regions under the current dataset configuration. However, no formal bootstrap, jackknife, or resampling-based stability analysis was performed. Consequently, confidence intervals or uncertainty bounds around centroid locations were not explicitly estimated. Future work should

incorporate dedicated resampling frameworks to quantify centroid uncertainty, evaluate sensitivity to corpus composition, and establish confidence regions for OAAS reference anchors under atlas expansion scenarios.

Embedding choice and metric transparency. The present OAAS implementation uses a linear PCA embedding to retain a globally interpretable Euclidean geometry in which centroid distances, displacement, and bounded indices preserve direct metric meaning. While nonlinear manifold learning methods (e.g., UMAP) can be useful for exploratory visualization and local neighborhood structure, they may distort global distances and introduce hyperparameter-sensitive variability that complicates reproducible cross-study comparability and control-oriented interpretation. In addition, the retained PC1–PC3 OAAS representation preserved 84.0% of the total variance, supporting the operational adequacy of the reduced linear embedding for centroid-distance analysis and constrained acoustic navigation. Accordingly, nonlinear embeddings remain a potential future extension for topology exploration, whereas the current study prioritizes metric transparency and operability under explicit geometric constraints.

Although OAAS-directed transformations can be quantitatively validated through centroid displacement, bounded index variation, and geometric convergence within the acoustic-affective space, the present framework does not claim direct inference of subjective emotional states or biological equivalence across individuals or species. The current implementation primarily validates the operational feasibility of constrained acoustic navigation under explicit metric conditions. Accordingly, the functional efficacy of OAAS-generated variants should be further evaluated through controlled *in vivo* behavioral, physiological, endocrine, and species-specific validation paradigms incorporating ecological context and longitudinal response assessment. In this sense, OAAS should be interpreted as a framework for operational acoustic-affective organization rather than as a direct emotion recognition system.

Reporting discipline and deployment constraints. OAAS interpretability depends on well-documented listener and playback conditions (e.g., SPL and loudness targets, background soundscape, speaker configuration) and on explicit reporting of intrinsic stimulus properties and production assumptions. Integrating OAAS with disciplined reporting practices and *in situ* calibration recordings will strengthen cross-study comparability and reduce the risk of over-reading geometry as direct affect inference, consistent with recent calls for improved protocol transparency in animal-centered music research [19,20]. OAAS should therefore be read as a design-and-evaluation geometry that can integrate upstream annotation schemes when available, but is not dependent on any single one.

Future directions. Priority directions include: (1) multi-corpus anchoring and uncertainty-aware atlas expansion (including confidence regions around centroids), with explicit extension to multi-species vocal repertoires and human affective emissions (e.g., laughter, crying, screams) to test interspecies correspondences under a shared operational geometry; (2) definition and evaluation of shared vs species-specific design targets, i.e., whether interspecies-compatible sound environments can be engineered as overlapping or negotiable reference regions while retaining context-appropriate anchoring; (3) systematic evaluation of broadband and textured reference signals (filtered noise families, roughness/entropy-controlled textures) as functional environments rather than trivial baselines; (4) species-appropriate auditory weighting and perceptual constraints in the feature representation; (5) multimodal validation designs in which OAAS-based displacement predictions are tested against behavioral, physiological, and ecological readouts under controlled playback conditions; and (6) evaluation of whether OAAS-derived geometric organization and centroid proximity systematically correspond to independently measured QBA-based affective responses and related organismal outcomes obtained in intervention-oriented sound environments.

5. Conclusions

Animal-centered sound and music research continues to face persistent methodological challenges, including heterogeneous experimental designs, limited comparability across studies, and insufficient reporting of stimulus properties and playback context [19,20]. To address these limitations, this study formalizes OAAS as an applied coordinate framework designed to enable quantitative comparability and constraint-aware navigation across heterogeneous sound environments through biologically grounded anchoring.

Across multi-domain case studies—including vocalizations, functional music interventions, challenge references, and noise baselines—OAAS demonstrates operational feasibility and methodological coherence. The analyses support three complementary capabilities: (1) construction of a biologically grounded acoustic reference space derived from vocal emission contexts; (2) evaluation and ranking of designed sound environments based on explicit geometric relationships (centroids, distances, bias, and bounded indices) rather than stylistic identity; and (3) operability as a constrained control space supporting directed displacement toward predefined reference regions while preserving structural coherence.

A central implication of OAAS is methodological: it makes explicit a set of structural acoustic mechanisms—spectral balance, temporal micro-structure, dynamic regulation, and informational organization—that are relevant to affect-oriented vocal communication and to the construction of functional sound environments. By formalizing these mechanisms within a shared coordinate geometry, OAAS enables systematic measurement, comparison, and controlled modulation, facilitating design and verification workflows grounded in reproducible representations.

This geometric formalization supports a shift from stimulus categorization toward structured acoustic engineering of functional sound environments. By enabling quantifiable positioning and constrained transformation relative to biologically grounded reference regions, OAAS provides a reproducible basis for intervention-oriented acoustics in animal welfare, enrichment, and biomedical contexts, and suggests translational relevance for soundscape research where design goals require comparable operational representations across heterogeneous environments [19,20].

Future work should expand anchoring corpora toward atlas-level coverage, integrate OAAS with disciplined reporting standards and controlled reproduction validation, and evaluate predictive and regulatory utility across species and ecological contexts. In particular, extending the anchoring layer toward multi-species vocal atlases may enable shared sound-environment design as an operational form of interspecies acoustic communication under welfare-oriented constraints.

Ethics statement

This article reports an analytical and post-hoc framework evaluation and does not involve new animal or human experiments. Ethical approvals and welfare procedures for the original intervention studies are reported in the corresponding primary publications cited in the manuscript.

CRediT authorship contribution statement

Berardo de Jesús Rodríguez: Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Juliana Zapata-Cardona:** Writing – review & editing, Visualization, Investigation, Formal analysis, Data curation, Conceptualization.

Funding

This research was supported by the Programmatic Call 2020 (Social Sciences, Humanities and Arts) of Universidad de Antioquia, Medellín, Colombia.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors acknowledge the use of AI-assisted language and technical tools during manuscript preparation, structuring, and formatting. All scientific decisions, interpretations, and conclusions remain the sole responsibility of the authors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.apacoust.2026.111457>.

Data availability

Minimal reproducible outputs (CSV) required to reproduce the main OAAS descriptors and supplementary tables, together with reproducibility scripts and audio examples for Fig. 4 (excluding third-party or proprietary audio such as SoundWel waveforms), are available in the public GitHub repository: <https://github.com/berardorodriguez/oaas-applied-acoustics-reproducibility>

The SoundWel dataset is publicly available through its original releases and should be accessed from the official sources cited in the manuscript.

References

- [1] International Organization for Standardization, 2014. Acoustics — soundscape — part 1: Definition and conceptual framework.
- [2] Lukashevich H, Grollmisch S, Abeßer J, Stober S, Bös J. How reliable are posterior class probabilities in automatic music classification?. In: Proceedings of the 2023 ACM International Conference on Multimedia Retrieval; 2023. p. 45–50. <https://doi.org/10.1145/3616195.3616228>.
- [3] Viveros-Muñoz R, Huijse P, Vargas V, Espejo D, Poblete V, Arenas JP, et al. The SPASS dataset: a new synthetic polyphonic dataset with spatiotemporal labels of sound sources. *Appl Acoust* 2023;214:109665. <https://doi.org/10.1016/j.apacoust.2023.109665>.
- [4] Herranz-Pascual K, García I, Aspuru I, Santander A. Soundscape and emotions through the acoustic environment: a systematic review. *Front Psychol* 2020;11:573563. <https://doi.org/10.3389/fpsyg.2020.573563>.
- [5] Fiebig A, Jordan P, Moshona CC. Assessments of acoustic environments by emotions – the application of emotion theory in soundscape. *Front Psychol* 2020;11:573041. <https://doi.org/10.3389/fpsyg.2020.573041>.
- [6] Kriengwatana BP, Mott R, ten Cate C. Music for animal welfare: a critical review & conceptual framework. *Appl Anim Behav Sci* 2022;251:105641. <https://doi.org/10.1016/j.applanim.2022.105641>.
- [7] Kogan LR, Schoenfeld-Tacher R, Simon AA. Behavioral effects of auditory stimulation on kennel dogs. *J Vet Behav* 2012;7:268–75. <https://doi.org/10.1016/j.jveb.2011.11.002>.
- [8] Wells DL, Graham L, Hepper PG. The influence of auditory stimulation on the behaviour of dogs housed in a rescue shelter. *Anim Welf* 2002;11:385–93. <https://doi.org/10.1017/S0962728600025112>.
- [9] Snowden CT. Animal signals, music and emotional well-being. *Animals* 2021;11:2670. <https://doi.org/10.3390/ani11092670>.
- [10] Briefer EF. Vocal expression of emotions in mammals: Mechanisms of production and evidence. *J Zool* 2012. <https://doi.org/10.1111/j.1469-7998.2012.00920.x>.
- [11] Filippi P, Congdon JV, Hoang J, Bowling DL, Reber SA, Pašukonis A, et al. Affective-acoustic information. *Proc R Soc B Biol Sci* 2017;284:20170990. <https://doi.org/10.1098/rspb.2017.0990>.
- [12] Laurijs KA, Briefer EF, Reimert I, Webb LE. Vocalisations in farm animals: a step towards positive welfare assessment. *Appl Anim Behav Sci* 2021;236:105264. <https://doi.org/10.1016/j.applanim.2021.105264>.

- [13] Briefer EF, Sypherd C-C-R, Linhart P, Leliveld LMC, Padilla de la Torre M, Read ER, et al. Classification of pig calls produced from birth to slaughter according to their emotional valence and context of production. *Sci Rep* 2022;12:3409. <https://doi.org/10.1038/s41598-022-07174-8>.
- [14] Montoya-Zuluaga J, Álvarez-Hernández N, Betancourth-Chaves P, Vallejo-Timarán D, Rodríguez BDJ. Musical stimulation in lactating sows affects pre-weaning piglet mortality and maternal behavior. *Sci Rep* 2026;16:4049. <https://doi.org/10.1038/s41598-025-34051-x>.
- [15] Zapata-Cardona J, Ceballos MC, Tarazona-Morales AM, David-Jaramillo E, de Rodríguez B. Spectro-temporal acoustic elements of music interact in an integrated way to modulate emotional responses in pigs. *Sci Rep* 2023;13:2994. <https://doi.org/10.1038/s41598-023-30057-5>.
- [16] Zapata-Cardona J, Ceballos MC, Tarazona-Morales AM, David-Jaramillo E, de Rodríguez B. Music modulates emotional responses in growing pigs. *Sci Rep* 2022;12:3382. <https://doi.org/10.1038/s41598-022-07300-6>.
- [17] Zapata-Cardona J, Duque-Arias S, David-Jaramillo E, et al. Effects of a veterinary functional music-based enrichment program on the psychophysiological responses of farm pigs. *Sci Rep* 2024;14:18660. <https://doi.org/10.1038/s41598-024-68407-6>.
- [18] Schafer RM. *The soundscape: our sonic environment and the tuning of the world*. Rochester, VT, USA: Destiny Books; 1977.
- [19] Truax B. *Acoustic communication*. 2nd ed. Westport, CT: Ablex Publishing; 2001.
- [20] Fitch WT. Four principles of bio-musicology. *Philos Trans R Soc, B* 2015;370:20140091. <https://doi.org/10.1098/rstb.2014.0091>.
- [21] Honing H, ten Cate C, Peretz I, Trehub SE, Winkler I. Without it no music: cognition, biology and evolution of musicality. *Philos Trans R Soc, B* 2015;370:20140088. <https://doi.org/10.1098/rstb.2014.0088>.
- [22] Aletta F, Kang J, Axelsson Ö. Soundscape descriptors and a conceptual framework for developing predictive soundscape models. *Landsc Urban Plan* 2016;149:65–74. <https://doi.org/10.1016/j.landurbplan.2016.02.001>.
- [23] Zapata-Cardona J, Ceballos MC, de Rodríguez B. Music and emotions in non-human animals from biological and comparative perspectives. *Animals* 2024;14:1491. <https://doi.org/10.3390/ani14101491>.
- [24] The SoundWel Database: A labeled pig vocalization repository (dataset record), 2022.
- [25] Trnka M, Darjaa S, Ritomský M, Sabo R, Rusko M, Schaper M, et al. Mapping discrete emotions in the dimensional space: an acoustic approach. *Electronics* 2021;10:2950. <https://doi.org/10.3390/electronics10232950>.
- [26] Álvarez-Hernández N, Vallejo-Timarán D, de Rodríguez B. Adapted original music as an environmental enrichment in an intensive pig production system reduced aggression in weaned pigs during regrouping. *Animals* 2023;13:3599. <https://doi.org/10.3390/ani13233599>.
- [27] Heffner RS, Heffner HE. Hearing in the domestic pig (*sus scrofa*) and goats (*capra hircus*). *Hear Res* 1990;48:231–40.
- [28] Kriengwatana BP, Nager RG, South A, Ullrich M, Doolittle EL. Playing music to animals: an interdisciplinary approach to improving our understanding of animals' responses to music. *Anim Behav* 2025;221:123074. <https://doi.org/10.1016/j.anbehav.2025.123074>.
- [29] Farina A, Gage SH. *Ecoacoustics: the ecological role of sounds*. Wiley-Blackwell; 2017.
- [30] Pijanowski BC, Villanueva-Rivera LJ, Dumyahn SL, Farina A, Krause BL, Napoletano BM, et al. Soundscape ecology: the science of sound in the landscape. *Bioscience* 2011;61:203–16.
- [31] Farina A. Eco-field theory: the ecological meaning of sounds. *Biosemiotics* 2021;14:365–82. <https://doi.org/10.1007/s12304-021-09435-9>.
- [32] Coutant M, Villain AS, Briefer EF. A scoping review of the use of bioacoustics to assess various components of farm animal welfare. *Appl Anim Behav Sci* 2024;275:106286. <https://doi.org/10.1016/j.applanim.2024.106286>.
- [33] Xie Y, Wang J, Chen C, Yin T, Yang S, Li Z, et al. Sound identification of abnormal pig vocalizations: enhancing livestock welfare monitoring on smart farms. *Inf Process Manag* 2024;61:103770. <https://doi.org/10.1016/j.ipm.2024.103770>.
- [34] Pann V, Kwon K-S, Kim B, Jang D-H, Kim J, Kim J-B. Robustness of CNN-based model assessment for pig vocalization classification across diverse acoustic environments. *Comput Electron Agric* 2026;240:111181. <https://doi.org/10.1016/j.compag.2025.111181>.
- [35] Morvai B, Boros M, Ferrando E, Magyari L, Andics A. Comparative EEG reveals general and conspecific vocalization sensitivities in evolutionarily distant mammal species. *Neuroimage* 2025;317:121355. <https://doi.org/10.1016/j.neuroimage.2025.121355>.
- [36] Díaz-Lozano KA, Becerra-Casillas OA, Huidobro N, Cervantes-L J, Serrano-García DI, Vélez-Pérez H, et al. Stochastic auditory noise enhances cardiac coherence evoked by slow-tempo music: preliminary findings. *Physiol Behav* 2026;304:115181. <https://doi.org/10.1016/j.physbeh.2025.115181>.
- [37] Poppelinkhuizen S, Van Steijn NJ, Marsman EMJ, Ten Berg S, De Haas LM, Kleer PS, et al. Classical beats and white noise: unveiling the effect of different music characteristics on heart rate variability. *Complement Ther Med* 2026;96:103288. <https://doi.org/10.1016/j.ctim.2025.103288>.
- [38] Gábor A, Lehoczki F, Bensaali-Nemes F, Faragó T, Surányi K, Andics A. Cross-species acoustic codes for yes and no in human nonverbal vocalizations. *Cognition* 2026;266:106284. <https://doi.org/10.1016/j.cognition.2025.106284>.
- [39] Pulina G. Opinion paper: the sound use of experimental hypotheses in animal science. *Animal* 2025;19:101634. <https://doi.org/10.1016/j.animal.2025.101634>.
- [40] Gui S, Gao J, Xie Y, Bi D, Li X. Latent deep-kernel filtering for lightweight and robust industrial sound source localization. *Mech Syst Sig Process* 2025;241:113554. <https://doi.org/10.1016/j.ymssp.2025.113554>.

Glossary and abbreviations

- OAAS:** Operational Acoustic–Affective Space
QBA: Qualitative Behavioural Assessment
PAD: Pleasure–Arousal–Dominance
AAI_OAAS: OAAS-based acoustic–affective index (bounded centroid-distance index)
POS / NEG: positive / negative vocal reference centroids (context-annotated anchor regions)
PCA: principal component analysis
SPL: sound pressure level
HRV: heart rate variability
RMS: root mean square
FFT: fast Fourier transform
STFT: short-time Fourier transform
LUFs: loudness units relative to full scale
EBU: European Broadcasting Union
HPSS: harmonic–percussive source separation
vOAAS: vocal-only OAAS projection fitted on vocal anchors