# Musical Intelligence: Scientific Supplement

This supplement accompanies MI-AE-Architecture-R3-3D. S1–S7 retain the trace, observation comparisons and figure-panel protocols. S8 adds the separately frozen birdsong transformation assay, source audit and revision-stage robustness checks. Source tables and analysis records are indexed with this revision package.

## Supplement S1. Computational trace: scope and results

This publication distinguishes an audit of a recorded execution from a public specification of the engine. The source recording, selected configuration and saved outputs identify the comparison. Equations, coefficients, temporal addresses, exact kernels and executable reconstruction records remain under controlled access. Independent reconstruction from the paper alone is consequently limited.

The separate acoustic assay in S8 has its own adapter and source pins. The shared repertoire count does not imply identical execution or activation of the recursive branches.

### Recorded execution and selection

The Swan Lake input lasts 182.596984 s. The 30–50-s interval was selected before inspecting intervention effects. Acoustic inputs, temporal context, output identities and consumption records belong to the same saved execution. The calculation audit checks that the inspected output follows the recorded implementation; it does not establish a cognitive or physiological interpretation.

### Controlled intervention

One temporal-context requirement is changed while the remaining inputs and calculations are fixed. The shared original temporal arrays remain exactly equal. The comparison uses 31,449 common valid frames after excluding frames that do not provide common support. Change counts use the recorded absolute tolerance of one ten-millionth plus a relative tolerance of one millionth. The published figure uses 100-ms display means, while native-frame values determine the checks.

Seventeen output dimensions change. ICEM, PWUP and CHPI are the three changed mechanisms among eight reachable mechanisms; all 805 dimensions outside the declared reachable graph remain exactly equal. Observer and identity-adapter controls have maximum absolute difference zero. Within 30–50 s, maximum absolute and RMS changes in the selected ICEM output are 0.003632 and 0.000643. The full-record maximum is 0.011401. These are model-output differences, not effects measured in listeners.

### Interpretation and reproducibility boundary

Declared reachability does not establish that a receiving calculation consumes the changing dimension. The audit distinguishes these cases. Counts of mechanisms or outputs alone also cannot establish identity between different historical executions. The archived behavioural cache limitations remain attached to their corresponding protocols in S7.

AAC.F1 is separately evaluated as a dimensionless acoustic-context score against excerpt-level arousal. Its source audit checks agreement with saved records; it is not a heart-rate measurement or an additional human dataset. Physiology, personal memory and other proposed constructs require their own observations. Figure 9, S5 and Extended Evidence E3–E4 specify these distinctions. Detailed source inspection requires access to the retained implementation records under the stated research-access terms.

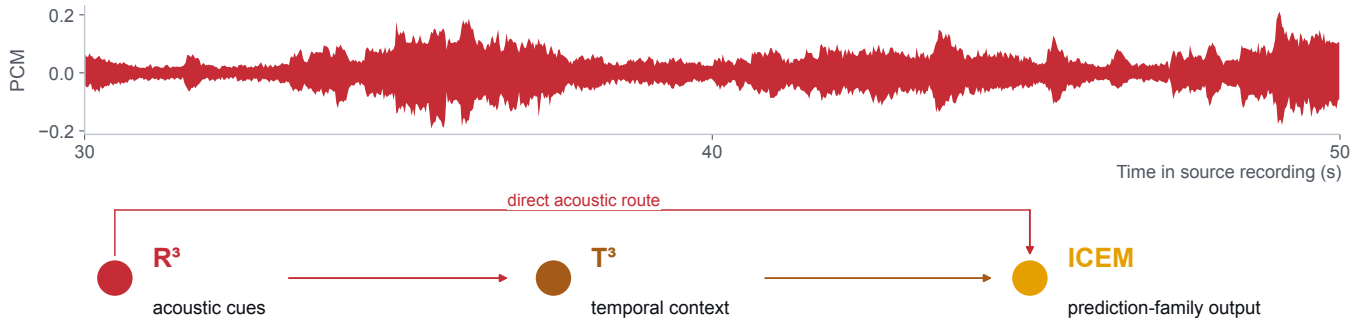
# A mechanism traced through a controlled change

Recorded execution evidence, without publishing the implementation recipe




## a One recording, connected stages, a fixed comparison interval

Swan Lake · 30–50 s







## b Recorded checks separate traceability from a cognitive interpretation

**Input identity**

The compared executions use the same saved acoustic input.

**Calculation audit**

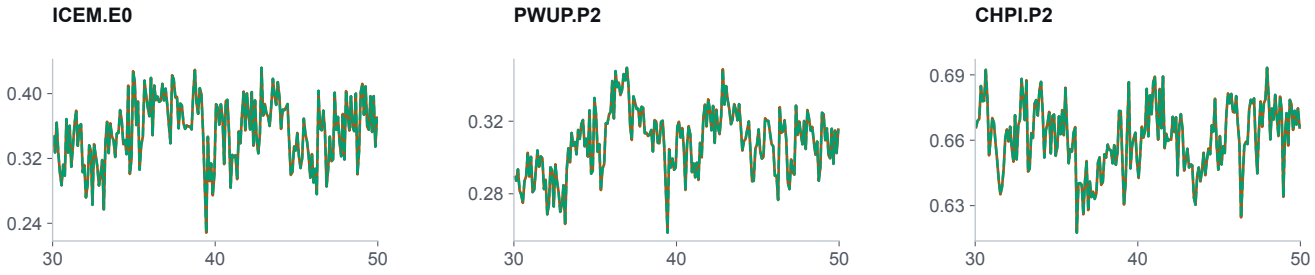
The source audit reconstructs the selected output from its recorded inputs.

**Controlled change**

One temporal requirement is modified; other calculations remain fixed.

## c The modified context changes selected output trajectories

Solid: reference · dashed: modified




Output values in native model units · 100-ms display means · shared time axis in seconds

## d A localized effect, with unchanged controls

**17 / 889**

output dimensions change

**805 / 805**

outside-graph outputs equal

**0**

observer / identity difference

This is a computational intervention, not a participant experiment. Equations, coefficients and request definitions are withheld.

**Figure S1.** Publication-level trace and intervention summary, also shown as Extended Evidence E1. The source waveform and output trajectories are saved observations of the model. The architectural route is schematic. Audit outcomes characterize the recorded computational intervention, not a participant experiment. Implementation equations, parameters and temporal request definitions are not disclosed.

## Supplement S2. Grouped tension prediction

### Sample, grouping and observation model

The film comparison uses 110 entries and perceived tension ratings from the soundtrack material of Eerola and Vuoskoski (2011). Forty source-soundtrack groups keep repeated excerpts and segments together. Five fixed outer GroupKFold splits hold out 22 entries each. All representations use these same outer splits, recordings and targets. Outputs are summarized after the first 172 cached frames, with full-clip context retained. The saved $C^3$ cache contains the channel means needed for the existing fifteen-output tension pool.

The pooled out-of-fold prediction vector supplies RMSE, Pearson r and $R^2$. RMSE is the primary error score. $R^2$ equals one minus the squared-error ratio relative to variation around the pooled observed target mean. The training-mean predictor is fitted separately within each outer training set. Table S2a reports all eight alternatives, and Table S2b retains the same five test-fold RMSEs. Their paired variation describes the observed split sensitivity; no significance or equivalence claim is inferred from five folds.

### Representations and preprocessing

The 97-input representation contains one mean per $R^3$ coordinate. The 970-input representation contains ten descriptors per coordinate: mean, standard deviation, 10th/50th/90th percentiles, mean absolute successive difference, normalized-time slope, and autocorrelations at 0.35, 1 and 4 s. These descriptors characterize the supplied acoustic series; they do not reproduce the complete canonical $T^3$ collection. The 985-input representation appends the fifteen $C^3$ means. The $C^3$-only representation retains all fifteen means without choosing a new winning channel.

The fixed acoustic subset contains mean, standard deviation and mean absolute successive difference for roughness ($R^3$ 0), pleasantness (4), loudness (10), tonalness (14) and spectral flux (21). These five coordinates were selected from the source audit before the compact fits. Their fifteen descriptors form a transparent comparator with the same number of regression inputs as $C^3$. This dimensionality refers to the observation model's inputs; computing a $C^3$ output can involve additional upstream quantities.

Two PCA alternatives reduce the standardized 97- and 970-input representations to fifteen components. PCA is fitted within the relevant inner or outer training partition, without whitening. Components are not rescaled after transformation. The observation models use ridge regression with an intercept. Mean imputation and population-standard-deviation scaling are fitted only to the training partition. Four grouped inner folds choose alpha from 0.1, 1, 10, 100, 1,000 and 10,000 by mean inner-fold squared error. Transformations and model fitting are repeated within each split rather than computed on the pooled sample.

### Comparison scope

The original matched comparison comprised 97 means, 970 descriptors and 970 descriptors plus $C^3$. The fixed extension added $C^3$ alone, the two PCA15 representations, the fixed acoustic subset and the training-mean reference. It retained the same outer folds and tuning grid. A separate NumPy implementation reproduced the original three baseline prediction vectors within $1.06 \times 10^{-13}$ and selected the same penalties. The input columns, complete inner losses, fitted transformations, predictions and chosen penalties are retained in the compact-comparison package.

$C^3$ alone gives pooled RMSE 1.47194; the fixed acoustic subset gives 1.48474. $C^3$ has lower error in three of five folds against that subset and against the 970-descriptor representation. The pooled differences therefore coexist with heterogeneous paired-fold differences. Appending $C^3$ to the 970 descriptors gives 1.53564 versus 1.53508 for those descriptors alone. Main Figure 6a includes these matched comparators alongside the useful $C^3$-only result. Table S2a also retains the PCA and training-mean alternatives.

The comparison is exploratory. The corpus and tension pool had been inspected before the new fitting specification, and parts of the architecture have corpus-guided development histories. Training-only fitting addresses the current regression evaluation; it does not make earlier representation development independent of the film corpus. The present inference is that the specified output representation supports grouped prediction under this observation model. The complete saved comparison is retained without additional feature selection or a revised scoring rule.

| Representation | Inputs | RMSE | Pearson r | $R^2$ |
|---|---|---|---|---|
| 15 $C^3$ means | 15 | 1.47194 | 0.67055 | 0.43148 |
| 15 fixed acoustic descriptors | 15 | 1.48474 | 0.65325 | 0.42155 |
| 97 acoustic means | 97 | 1.60410 | 0.58347 | 0.32480 |
| 970 acoustic/temporal descriptors | 970 | 1.53508 | 0.61895 | 0.38165 |
| 970 descriptors + 15 $C^3$ means | 985 | 1.53564 | 0.61868 | 0.38121 |
| PCA15 of 97 acoustic means | 15 | 1.65161 | 0.54102 | 0.28421 |
| PCA15 of 970 descriptors | 15 | 1.64603 | 0.55121 | 0.28904 |
| Training-target mean | 0 | 1.96593 | -0.15862 | -0.01415 |

Table S2a. Complete pooled held-out film-tension comparison.

| Representation | Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 |
|---|---|---|---|---|---|
| 15 $C^3$ means | 1.37009 | 1.69754 | 1.87579 | 1.12379 | 1.13699 |
| 15 fixed acoustic descriptors | 1.30096 | 1.80774 | 1.51696 | 1.24927 | 1.48323 |
| 97 acoustic means | 1.33345 | 1.74878 | 1.93541 | 1.15757 | 1.71571 |
| 970 acoustic/temporal descriptors | 1.20187 | 1.80417 | 1.86237 | 1.12779 | 1.53052 |
| 970 descriptors + 15 $C^3$ means | 1.23221 | 1.80374 | 1.85695 | 1.12480 | 1.51843 |
| PCA15 of 97 acoustic means | 1.29656 | 1.81932 | 1.97706 | 1.19733 | 1.81816 |
| PCA15 of 970 descriptors | 1.30035 | 1.77968 | 1.95197 | 1.29779 | 1.78733 |
| Training-target mean | 1.89024 | 2.34042 | 2.13564 | 1.41956 | 1.92294 |

Table S2b. RMSE on the five fixed soundtrack-grouped folds; each test fold contains 22 entries.

## Supplement S3. Arousal and valence correspondence

### Observation targets and samples

The transfer comparison relates clip-mean model outputs to perceived arousal and valence ratings in MUSIFEAST-17 (van der Walle et al., 2025). The historical film column refers to the 360-entry energy/arousal and valence tables from the earlier soundtrack evaluation. Those film targets and the new perceived-emotion targets are not identical measurements. Table S3a preserves the eight selected channel identities and their signed Spearman associations across both datasets and the recorded MUSIFEAST subset.

The full MUSIFEAST analysis contains 356 thirty-second excerpts. Its recorded partition comprises seventeen consumed pilot, 274 development and 65 fixed test excerpts. Connected artist/composer, source and identical-audio identities are grouped during partition construction. The 65-excerpt subset is contained in the full sample; the two columns therefore describe overlapping observations. Scores use the channel mean after the first 172 cached frames and the human rating mean for the same excerpt. No multivariate regression, acoustic adjustment or fitted calibration enters these correlations.

### Channel selection and chronology

Four arousal identities are AAC.F1, PUPF.U0, AAC.A1 and CSG.P0. Four valence identities are SRP.P1, CMAT.S0, SRP.P2 and VMM.V0. The mapping was revised after the initial pilot using prior film tables. There was no new winning-channel selection on the 65-excerpt test subset. Raw normative labels had already been accessed for curation and quality control. Different engine snapshots, populations and target definitions make this an exploratory cross-dataset comparison.

The earlier film analyses selected channels within specified target pools using whole-set absolute association. The 110-entry and 360-entry film sets overlap, with the smaller set developed from the larger pilot. Their results document the channel-selection context, rather than providing two independent replications. The separate film residual procedure subtracted out-of-fold ridge predictions based on 97 $R^3$ clip means before examining mechanism associations; that procedure does not enter the direct transfer scores in Table S3a. Full film pools and their target identities remain in the source index.

### Complete transfer family

All four transferred arousal readouts are positively associated with MUSIFEAST ratings. The full-sample correlations range from 0.531 to 0.691; in the 65-excerpt subset they range from 0.478 to 0.770. AAC.F1 and PUPF.U0 provide the leading full-sample arousal associations. The valence readouts are weaker, spanning −0.086 to 0.231 in the full sample and −0.049 to 0.211 in the subset. Table S3a keeps this complete family visible while the main article focuses on arousal correspondence.

Spearman ρ describes rank correspondence between one model mean and one human excerpt mean. These coefficients are neither percentages of correct predictions nor incremental variance beyond acoustic controls. Their observation unit is the excerpt, not a model frame or an individual rating trial. The channel keys identify the evaluated computations; the human targets specify the empirical interpretation. S1 summarizes the AAC.F1 audit and its interpretation boundary; the implementation expression is not published.

The source record includes the direct score table, historical film tables, partition definitions, transfer chronology and input hashes. The main four-channel table and Table S3a use the same primary mean window. No alternative trimming, target transformation or test-subset channel replacement is introduced for presentation. The displayed precision follows the retained score tables, with complete machine-readable values available through the source index.

| Channel / target | Film, n = 360 | MUSIFEAST, n = 356 | Test subset, n = 65 |
|---|---|---|---|
| AAC.F1 / arousal | +0.652 | +0.691 | +0.770 |
| PUPF.U0 / arousal | +0.698 | +0.676 | +0.735 |
| AAC.A1 / arousal | +0.646 | +0.533 | +0.530 |
| CSG.P0 / arousal | +0.631 | +0.531 | +0.478 |
| SRP.P1 / valence | +0.622 | +0.231 | +0.211 |
| CMAT.S0 / valence | +0.653 | +0.172 | +0.155 |
| SRP.P2 / valence | +0.623 | +0.174 | +0.160 |
| VMM.V0 / valence | +0.643 | -0.086 | -0.049 |

Table S3a. Complete direct transfer family; the subset is nested within the total.

## Supplement S4. Consonance comparisons

### Acoustic routes and perceptual targets

The consonance comparisons evaluate explicit waveform-derived coordinates against interval or chord ratings. On the inspected route, detected spectral peaks contribute to pairwise beating and fusion calculations. Sensory pleasantness is the fixed combination a fixed model-specific combination of acoustic consonance cues. Inharmonicity equals one minus fusion on this route. Their signed correlations consequently describe related quantities within one computation. The mel-only fallback uses a different estimator; route identity is recorded with each evaluation.

The harmonic-dyad comparison uses thirteen semitone-bin rating means from Marjieh et al. (2024). The reported 0.8901 pleasantness coefficient, 0.8132 fusion coefficient and −0.8352 roughness coefficient use these interval means as observations. The pooled analysis uses 617 chord rows from Eerola and Lahdelma (2021), with pleasantness 0.5177, roughness −0.5825 and scaled Sethares dissonance −0.4337. These are different stimulus-level evaluations; individual ratings are not substituted for the aggregated observation counts.

### Nine-reference panel

Main Figure 8G contains seven behavioural target rows and two analytical anchors. The behavioural rows concern harmonic dyads, flute/guitar/piano rating conditions, pure tones, Indian tension and the behavioural ratings from Bidelman and Krishnan (2009). The extended extractor synthesizes harmonic dyads with six partials of amplitude 1/h. Instrument names in the target columns therefore denote rating conditions; they do not indicate reconstruction of the original instrument recordings. The Indian tension target retains its recorded polarity correction (Lahdelma et al., 2024).

The speech-harmonics row (Schwartz et al., 2003) and analytical Sethares reference (Sethares, 1993) are plotted separately from the behavioural targets. Table S4a retains all 63 signed descriptor/reference correlations. Intervals or bins provide the n values. The Bidelman row uses six behavioural interval targets, not a neural frequency-following-response endpoint. The other harmonic, instrumental and pure-tone conditions use thirteen bins, and the Indian tension condition uses twelve.

The source records distinguish direction agreement, magnitude criteria and selected interval-subset ranking. The full signed associations are presented here rather than combining those criteria into one accuracy percentage. The opposite signs for fusion and its algebraic complement are preserved. The source index retains the target conventions, extraction inputs and statistical tables for readers examining a particular rating condition.

### Carillon register and synthesis conditions

The Carillon comparison follows published bell spectra and rating targets from Harrison and MacConnachie (2024). The saved sweep creates additive dyads at approximately 66, 261, 440 and 880 Hz, each under decaying, sustained and idealized spectral constructions. One extractor and its coefficients are retained across the twelve combinations. Inputs are synthesized rather than original bell recordings. The recorded MI sweep uses 1.5-second signals at 44.1 kHz and normalizes the component bells separately before combination.

Table S4b retains all twelve fusion and roughness associations. Fusion ranges from 0.3187 to 0.8297, with the largest value at 880 Hz sustained. Roughness is positive at the lowest register and negative at higher registers. These condition-specific results characterize the same computation under changes in spectral construction and register. The highest cell is not substituted for performance averaged across the design. Main Figure 8J presents the full fusion grid with its original scale.

### Reproduction records

The source tables retain full-precision associations, their recorded probabilities, synthesis conditions and input hashes. The general $R^3$ definitions specify waveform availability, sample rate, support and normalization. Clip-maximum normalization and batch-dependent groove normalization belong to the relevant feature definitions; they are not additional observations. The consonance comparisons reported here use their recorded waveform routes and target conventions. The source index separates these perceptual evaluations from earlier fitted baselines and other historical tasks.

Tables S4a–S4b and the main consonance figure are presentation views of existing numeric records. Together with the synthesis and route definitions, they identify the quantities that must be retained when reproducing the reported comparisons. The complete extended archive remains available for the additional analysis families, whose methods and conclusions are separate from the primary comparisons in this article.

| Reference (n) | R | S | H | F | P | I | D |
|---|---|---|---|---|---|---|---|
| Harmonic dyads (13) | -0.8352 | -0.8132 | +0.3681 | +0.8132 | +0.8901 | -0.8132 | +0.1264 |
| Flute ratings (13) | -0.7747 | -0.7967 | +0.2582 | +0.5330 | +0.7692 | -0.5330 | +0.4505 |
| Guitar ratings (13) | -0.5330 | -0.5495 | +0.2253 | +0.4396 | +0.5989 | -0.4396 | +0.2967 |
| Piano ratings (13) | -0.7363 | -0.5824 | +0.1484 | +0.6154 | +0.6923 | -0.6154 | +0.4725 |
| Pure tones (13) | -0.7253 | -0.3242 | +0.1703 | +0.4505 | +0.3571 | -0.4505 | -0.5220 |
| Indian tension* (12) | -0.5245 | -0.3986 | +0.3636 | +0.8042 | +0.5804 | -0.8042 | +0.4056 |
| Bidelman ratings (6) | -0.8857 | -0.9429 | +0.0857 | +1.0000 | +1.0000 | -1.0000 | +0.0286 |
| Speech anchor (13) | -0.8132 | -0.7692 | +0.2857 | +0.9231 | +0.8791 | -0.9231 | +0.2253 |
| Sethares anchor* (13) | -0.8132 | -0.7692 | +0.2857 | +0.9231 | +0.8791 | -0.9231 | +0.2253 |

Table S4a. All 63 signed correlations. R: roughness; S: Sethares; H: Helmholtz–Kang; F: fusion; P: pleasantness; I: inharmonicity; D: harmonic deviation. Last two rows are analytical anchors; * marks target inversion.

| Register | Synthesis | Fusion ρ | Roughness ρ |
|---|---|---|---|
| ≈66 Hz | Decay | +0.3187 | +0.2198 |
| ≈66 Hz | Sustained | +0.3901 | +0.2198 |
| ≈66 Hz | Idealized | +0.3571 | +0.2363 |
| 261 Hz | Decay | +0.5055 | -0.6264 |
| 261 Hz | Sustained | +0.4341 | -0.6264 |
| 261 Hz | Idealized | +0.3846 | -0.6264 |
| 440 Hz | Decay | +0.5385 | -0.7253 |
| 440 Hz | Sustained | +0.3901 | -0.7253 |
| 440 Hz | Idealized | +0.4066 | -0.7198 |
| 880 Hz | Decay | +0.8022 | -0.6648 |
| 880 Hz | Sustained | +0.8297 | -0.6758 |
| 880 Hz | Idealized | +0.8132 | -0.6648 |

Table S4b. All twelve saved Carillon register/synthesis conditions.

## Supplement S5. Mechanism provenance, evidence and revision

This record distinguishes the motivation for a proposal from evidence about its current implementation. A mechanism is an executable computational unit with a stable name, specified inputs and outputs. Its interpretation as a human process remains an outcome-specific hypothesis. Table S5a covers all 84 canonical names; the accompanying record preserves full source paths, 889 output identities, 484 registry citation entries, the 89 ledger definitions and their historical model links. The five repeated names are PNH, STAI, TPRD, IGFE and MAA. Same-name definitions are not assumed to compute the same function.

Supplement S8 adds an executable-input audit for the transformation assay. Its feature-index and horizon checks concern the frozen implementation; the literature roles below do not adjudicate those code–description discrepancies or validate the resulting equations.

### Literature synthesis and the origin of the repertoire

The original Cognitive Consonance Circuit programme organized literature extraction, feature queries, nine-domain aggregation, evidence classification, knowledge synthesis, representation refinement, cross-reference checks and theoretical synthesis. The retained extraction folders contain 448 JSON records and 1,116 claims. These include LLM-assisted extraction and subsequent enrichment; an extracted interpretation is not itself a new empirical observation. The original programme used exploratory meta-analytic summaries to organize candidate relationships before their later computational implementation.

The historical T01C documents contain 96 model headings, of which two are explicitly marked duplicates; the remaining 94 proposals have 27 mechanistic, 38 integrative and 29 speculative labels. A subsequent February catalogue contains 96 distinct names, including CHPI and SSRI. All 84 current canonical abbreviations occur in that catalogue, while 82 occur in the earlier T01C records; some full-name expansions differ across versions. The twelve catalogue names absent from the current canonical selection are AMSC, DSP, EDTA, MDNS, MPFS, MTNE, OMS, ONI, PTGMP, SDD, SDDP and TMRM. An individual scientific elimination decision was not recovered for each difference. The count 84 consequently identifies this implementation's scope; it is not the outcome of an empirical optimization over the historical catalogue.

The extraction record is not a set of 448 independent studies. Simple normalization identifies 340 distinct DOI strings among 426 nonempty DOI fields, and 385 normalized title strings among all 448 records. Neither number is a final adjudicated study count. The historical Core-4 meta-analysis retains 149 effect/claim rows, including multiple rows from the same source and twelve rows with a default n=30. Its inspected code filters labelled Cohen's d values at absolute magnitude below two, uses a common variance expression and does not model within-study effect dependence. The archived estimates therefore are not transferred to the current equations as validated coefficients or model-performance estimates. Protocol drafts and internal quality percentages likewise do not establish prospective registration or independently verified extraction accuracy. This study uses the traceable source lineage as hypothesis motivation; the original quantitative synthesis would require its own study-level re-adjudication for a renewed meta-analytic claim.

### Evidence states and table interpretation

Every row has a located implementation and literature/design provenance. The revised compact anchors in Figure 9 and Table S5a identify one or two sources with an explicit role in Table S5b. The 4 October 2026 source-role review retains all 84 engine identifiers, computations and evaluation flags. It corrects source selection without claiming new implementation validation. Historical anchors, including the unresolved Sares (2023) entry removed from NEWMD, remain in the editorial record. Reviews and editorials are background, not independent experiments. The wider registry is an unchanged historical record, not an adjudicated evidence set. A retained label may extend beyond the cited observation, particularly at the brainstem, synaptic, cross-species, memory or transfer level; these gaps are stated individually in S5b. The selected examples below continue to distinguish source observation, computational translation and current evaluation.

The evaluation column uses nonexclusive links: **C**, a selected calculation or intervention inspected here or in the linked source audit; **P**, selected outputs included jointly in the fixed fifteen-output film-tension prediction comparison; **A**, selected outputs in the complete eight-channel film-to-MUSIFEAST association family; **H**, a specified historical comparison retained in S7 or the linked technical archive. **U** means no additional mechanism-specific outcome judgement is assigned by this compact crosswalk. U does not mean that the literature is empty, that the mechanism is false, or that every historical analysis involving it has been ruled out. C establishes arithmetic/access, P a joint representation test, and A/H a particular observation or dataset relationship; none establishes every output or the necessity of an entire mechanism.

For P, the fifteen-output representation gave r=0.67055 and $R^2$=0.43148 in the grouped film-tension comparison, while appending those outputs to the 970-descriptor baseline did not improve pooled error. For A, the full channel-specific arousal and valence associations are retained in Table S3a, including weaker or negative transfer. H preserves its original selection, timing and observation-model restrictions. A statistically weak comparison is not recoded as an unavailable test.

**Table S5a. Complete canonical mechanism crosswalk**

| ID / function | Registry name | Reviewed anchors | Implemented operation | Evaluation links |
|---|---|---|---|---|
| BCH / F1 | Brainstem Consonance Hierarchy | Bidelman 2009; Bidelman 2013 | Combines consonance and pitch cues into brainstem-labelled consonance scores. | H |
| CSG / F1 | Consonance-Salience Gradient | Bravo 2017; Sarasso 2019 | Maps acoustic consonance and change cues to salience and affective-evaluation scores. | A |
| MIAA / F1 | Musical Imagery Auditory Activation | Kraemer 2005; Halpern 2004 | Uses the supplied audio to form imagery-labelled, retrieval and recognition proxies. | U |
| MPG / F1 | Melodic Processing Gradient | Rupp 2022; Patterson 2002 | Separates onset, sequence and contour cues into a melodic-processing gradient. | U |
| PNH / F1 | Pythagorean Neural Hierarchy | Bidelman 2009; Tabas 2019 | Maps interval-complexity cues to conflict and consonance-preference scores. | U |
| SDNPS / F1 | Stimulus-Dependent Neural Pitch Salience | Cousineau 2015; Penagos 2004 | Constructs pitch-salience and consonance proxies with an explicit stimulus-dependence interpretation. | U |
| SDED / F1 | Sensory Dissonance Early Detection | Crespo-Bojorque 2018; Fishman 2001 | Transforms roughness and harmonic cues into early-detection and mismatch-labelled scores. | U |
| TPIO / F1 | Timbre Perception-Imagery Overlap | Halpern 2004; Zatorre 2005 | Constructs shared perception/imagery-labelled timbre scores from the supplied audio. | U |

Continued on next page

Table S5a (continued)

| ID / function | Registry name | Reviewed anchors | Implemented operation | Evaluation links |
|---|---|---|---|---|
| TPRD / F1 | Tonotopy-Pitch Representation Dissociation | Allen 2022; Norman-Haignere 2013 | Separates frequency-pattern and pitch-labelled scores within the model. | U |
| HTP / F2 | Hierarchical Temporal Prediction | Norman-Haignere 2022; Bonetti 2024 | Combines acoustic cues at multiple horizons into hierarchical prediction scores. | H |
| SPH / F2 | Spatiotemporal Prediction Hierarchy | Bonetti 2024; Fernandez-Rubio 2022 | Constructs memory-match and deviation scores with oscillatory and hierarchical labels. | U |
| ICEM / F2 | Information Content Emotion Model | Egermann 2013; Cheung 2019 | Combines acoustic change and dispersion into unexpectedness and emotion-labelled proxies. | C, H |
| SNEM / F3 | Sensory Novelty and Expectation Model | Nozaradan 2011; Nozaradan 2012 | Transforms rhythmic and acoustic change cues into novelty and expectation scores. | U |
| IACM / F3 | Inharmonicity-Attention Capture Model | Basinski 2025; Koelsch 1999 | Maps inharmonicity and change cues to attention-capture and segregation proxies. | U |
| MEAMN / F4 | Music-Evoked Autobiographical Memory Network | Janata 2009; Sakakibara 2025 | Maps audio-derived familiarity and tonal cues to autobiographical-memory-labelled proxies. | U |
| MMP / F4 | Musical Mnemonic Preservation | Jacobsen 2015; Fang 2017 | Maps acoustic cues to preserved-memory, recognition and scaffolding-labelled proxies. | C, H |
| SRP / F5 | Striatal Reward Prediction | Salimpoor 2011; Salimpoor 2013 | Constructs wanting, liking, tension and reward-forecast proxies from audio-derived cues. | P, A |
| AAC / F5 | Autonomic-Arousal Circuit | Egermann 2013; Salimpoor 2011 | Maps acoustic intensity and change to autonomic-labelled arousal and chills proxies. | C, P, A |
| VMM / F5 | Valence-Mode Mapping | Mitterschiffthaler 2007; Koelsch 2006 | Combines mode-related and spectral cues into valence and happy/sad-labelled scores. | A |
| DAED / F6 | Dopamine Anticipation-Experience Dissociation | Salimpoor 2011; Gold 2023 | Separates anticipatory and consummatory reward proxies with caudate and accumbens labels. | H |
| MORMR / F6 | mu-Opioid Receptor Music Reward | Putkinen 2025; Salimpoor 2011 | Transforms acoustic and temporal cues into opioid-inspired reward proxies. | H |
| RPEM / F6 | Reward Prediction Error in Music | Gold 2019a; Gold 2023 | Combines surprise and liking proxies into signed reward-prediction-error scores. | P |
| PEOM / F7 | Period Entrainment Optimization Model | Thaut 2015; Thaut 1998 | Transforms rhythmic cues into entrainment and motor-timing-labelled scores. | U |
| MSR / F7 | Musician Sensorimotor Reorganization | Bangert 2006; Lahav 2007 | Constructs musician-labelled precision, inhibition and sensorimotor scores. | U |
| GSSM / F7 | Gait-Synchronized Stimulation Model | Yamashita 2025 | Constructs gait/stimulation-labelled synchronization scores from acoustic timing cues. | U |
| HMCE / F7 | Hierarchical Musical Context Encoding | Mischler 2025; Norman-Haignere 2022 | Combines local and extended temporal cues into hierarchical musical-context scores. | U |
| EDNR / F8 | Expertise-Dependent Network Reorganization | Paraskevopoulos 2022; Leipold 2021 | Maps acoustic cues to expertise, network-separation and efficiency-labelled proxies. | U |
| PSCL / F1 | Pitch Salience in Cortical Lateralization | Penagos 2004; Allen 2022 | Combines spectral cues and BCH outputs into pitch-prominence and cortical-response-labelled scores. | U |
| PWUP / F2 | Precision-Weighted Uncertainty Processing | Mencke 2019; Cheung 2019 | Uses HTP and ICEM outputs to weight tonal and rhythmic prediction-error proxies. | C |
| BARM / F3 | Brainstem Auditory Response Modulation | Rathcke 2024; Grahn 2007 | Combines rhythmic context and upstream scores into beat-alignment and regularization proxies. | U |
| STANM / F3 | Spectrotemporal Attention Network Model | Haiduk 2024; Zatorre 2022 | Partitions acoustic cues into temporal and spectral attention-allocation scores. | U |
| AACM / F3 | Aesthetic-Attention Coupling Model | Sarasso 2019; Sarasso 2021 | Combines acoustic and sustained-response cues into aesthetic-engagement scores. | U |
| AMSS / F3 | Attention-Modulated Stream Segregation | Hausfeld 2021; Wikman 2025 | Combines harmonic and temporal cues into auditory-stream segregation scores. | U |
| ETAM / F3 | Entrainment, Tempo and Attention Modulation | Hausfeld 2021; Doelling 2015 | Combines rhythmic timing cues into entrainment and attention-gain scores. | U |
| RASN / F4 | Rhythmic Auditory Stimulation Neuroplasticity | Thaut 2015; Grahn 2007 | Constructs rhythmic entrainment, motor-facilitation and plasticity-labelled proxies. | U |
| PMIM / F4 | Predictive Memory Integration Model | Koelsch 2000; Koelsch 2009 | Combines harmonic and temporal deviance cues into memory-updating proxies. | U |
| OII / F4 | Oscillatory Intelligence Integration | Bruzzone 2022; Samiee 2022 | Combines acoustic and upstream cues into integration, segregation and encoding scores. | U |
| HCMC / F4 | Hippocampal-Cortical Memory Circuit | Squire 1995; McClelland 1995 | Maps acoustic and memory cues to binding, segmentation and storage-labelled scores. | U |
| RIRI / F4 | RAS-Intelligent Rehabilitation Integration | Thaut 2015; Harrison 2025 | Combines rhythmic cues and available upstream averages into rehabilitation-labelled motor proxies. | U |
| MSPBA / F4 | Musical Syntax Processing in Broca's Area | Maess 2001; Kim 2021 | Transforms harmonic-pattern cues into syntax, violation and Broca-labelled scores. | U |
| PUPF / F5 | Pleasure-Uncertainty Prediction Function | Cheung 2019; Egermann 2013 | Combines uncertainty and surprise proxies into a pleasure response function. | P, A |
| CLAM / F5 | Closed-Loop Affective Modulation | Ehrlich 2019; Daly 2016 | Computes affect-tracking and control-labelled scores without an external closed-loop intervention. | U |
| MAD / F5 | Musical Anhedonia Disconnection | Loui 2017; Martinez-Molina 2016 | Constructs auditory/reward coupling and disconnection-labelled proxies from acoustic cues. | U |

Continued on next page

Table S5a (continued)

| ID / function | Registry name | Reviewed anchors | Implemented operation | Evaluation links |
|---|---|---|---|---|
| NEMAC / F5 | Nostalgia-Evoked Memory-Affect Circuit | Sakakibara 2025; Barrett 2010 | Combines memory and affect cues into nostalgia and memory-reward scores. | U |
| STAI / F5 | Spectral-Temporal Aesthetic Integration | Kim 2019; Alluri 2012 | Combines spectral and temporal integrity cues into aesthetic-response scores. | U |
| IUCP / F6 | Inverted-U Complexity Preference | Gold 2019b; Gold 2023 | Applies complexity/preference functions to information-content and entropy proxies. | U |
| MCCN / F6 | Musical Chills Cortical Network | Chabin 2020; Putkinen 2025 | Combines arousal and reward cues into chills and cortical-network-labelled scores. | U |
| MEAMR / F6 | Music-Evoked Autobiographical Memory Reward | Janata 2009; Salimpoor 2011 | Combines familiarity and tonal cues into autobiographical-memory reward proxies. | U |
| SSRI / F6 | Social Synchrony Reward Integration | Kokal 2011; Bigand 2025 | Transforms rhythmic and reward cues into synchrony, bonding and group-flow-labelled proxies. | U |
| HGSIC / F7 | Hierarchical Groove State Integration Circuit | Potes 2012; Grahn 2007 | Combines beat, metre and motor-labelled cues into a groove-state score. | U |
| ASAP / F7 | Action Simulation for Auditory Prediction | Ross 2022; Patel 2014 | Constructs motor-simulation-labelled timing predictions from acoustic and upstream cues. | U |
| DDSMI / F7 | Dyadic Dance Social Motor Integration | Bigand 2025; Kohler 2025 | Combines acoustic and upstream timing cues into partner-coordination-labelled scores. | U |
| VRMSME / F7 | VR Music Stimulation Motor Enhancement | Liang 2025 | Constructs virtual-reality-labelled motor and connectivity proxies from acoustic and upstream cues. | U |
| SPMC / F7 | SMA-Premotor-M1 Motor Circuit | Grahn 2007; Hoddinott 2024 | Combines timing cues into sequence-planning, motor-preparation and execution-labelled scores. | U |
| TSCP / F8 | Temporal Synaptic Consolidation Processor | Pantev 2001; Bellmann 2024 | Constructs timbre-specific recognition and plasticity-labelled scores from spectral cues. | U |
| CDMR / F8 | Context-Dependent Mismatch Response | Rupp 2022; Crespo-Bojorque 2018 | Combines acoustic context and deviance cues into mismatch and expertise-labelled scores. | P |
| SLEE / F8 | Synaptic Long-term Encoding Engine | Paraskevopoulos 2022; Porfyri 2025 | Constructs regularity, detection and learning-labelled scores from acoustic and temporal cues. | U |
| PCCR / F1 | Pitch Chroma Cortical Representation | Briley 2013; Allen 2022 | Combines chroma and upstream pitch scores into pitch-class identity and octave-equivalence proxies. | U |
| WMED / F2 | Working Memory-Entrainment Dissociation | Noboa 2025; Yuan 2025 | Separates rhythm-entrainment and working-memory-labelled contributions to prediction scores. | U |
| DGTP / F3 | Domain-General Temporal Processing | Rathcke 2024; Grahn 2007 | Combines timing proxies into shared music/speech-labelled temporal scores. | U |
| SDL / F3 | Salience-Dependent Lateralization | Haiduk 2024; Zatorre 2022 | Transforms spectral and temporal salience cues into lateralization-labelled scores. | U |
| NEWMD / F3 | Neural Entrainment-Working Memory Dissociation | Noboa 2025 | Separates entrainment and working-memory-labelled rhythm scores. | U |
| PWSM / F3 | Precision-Weighted Salience Model | Basinski 2025; Millidge 2022 | Combines error and reliability cues into precision-weighted salience scores. | U |
| VRIAP / F4 | VR-Integrated Analgesia Paradigm | Deng 2025 | Combines acoustic and upstream cues into engagement and analgesia-labelled proxies. | U |
| CMAPCC / F4 | Cross-Modal Action-Perception Common Code | Lahav 2007; Bangert 2006 | Combines acoustic and upstream representations into action/perception transfer scores. | U |
| DMMS / F4 | Developmental Music Memory Scaffold | Partanen 2022; Qiu 2025 | Constructs developmental binding and memory-scaffold proxies from acoustic and upstream cues. | U |
| CSSL / F4 | Cross-Species Song Learning | Burchardt 2025 | Combines rhythmic and melodic cues into copying and shared-learning-labelled scores. | U |
| CDEM / F4 | Context-Dependent Emotional Memory | Sachs 2025 | Combines acoustic affect and memory cues into context-dependent binding scores. | U |
| DAP / F5 | Developmental Affective Plasticity | Scholkmann 2024; Trainor 2012 | Maps acoustic and upstream cues to developmental-plasticity and affect-labelled proxies. | C |
| CMAT / F5 | Cross-Modal Affective Transfer | Spence 2011 | Combines acoustic affect cues into supramodal-labelled valence and arousal scores. | A |
| TAR / F5 | Therapeutic Affective Resonance | Kheirkhah 2025; Ehrlich 2019 | Combines acoustic and upstream affect cues into therapeutic-target-labelled proxies. | U |
| LDAC / F6 | Liking-Dependent Auditory Cortex | Gold 2023; Cheung 2019 | Combines liking and unexpectedness proxies into an auditory-cortex-labelled response. | U |
| IOTMS / F6 | Individual Opioid Tone Music Sensitivity | Putkinen 2025; Salimpoor 2011 | Maps reward cues to individual-sensitivity and opioid-tone-labelled proxies. | U |
| SSPS / F6 | Saddle-Shaped Preference Surface | Cheung 2019; Gold 2019b | Combines information-content and entropy proxies on a preference surface. | U |
| NSCP / F7 | Neural Synchrony Commercial Prediction | Leeuwis 2021 | Transforms timing and upstream cues into synchrony and popularity-labelled proxies. | U |
| CTBB / F7 | Cerebellar Theta-Burst Balance | Sansare 2025; Okada 2022 | Constructs cerebellar-stimulation-labelled timing and balance proxies. | U |
| STC / F7 | Singing Training Connectivity | Zamorano 2023; Kleber 2013 | Combines acoustic and upstream cues into singing-related interoceptive and motor scores. | U |
| ESME / F8 | Error-Signal Modulated Encoding | Koelsch 1999; Vuust 2012 | Combines spectral, timing and available expertise cues into modality-specific mismatch proxies. | U |

Continued on next page

Table S5a (continued)

| ID / function | Registry name | Reviewed anchors | Implemented operation | Evaluation links |
|---|---|---|---|---|
| ECT / F8 | Expertise Compartmentalization Trade-off | Paraskevopoulos 2022; Porfyri 2025 | Combines acoustic and upstream expertise cues into efficiency and transfer-limit proxies. | U |
| UDP / F2 | Uncertainty-Driven Pleasure | Mencke 2019; Gold 2019b | Combines uncertainty and prediction confirmation into a pleasure-labelled score. | U |
| CHPI / F2 | Cross-Modal Harmonic Predictive Integration | Moller 2021; Cheung 2019 | Combines acoustic and upstream prediction scores into harmonic-integration proxies. | C |
| IGFE / F2 | Individual Gamma Frequency Enhancement | Yokota 2025; Bolland 2025 | Constructs gamma-labelled matching and enhancement proxies from acoustic and upstream scores. | U |
| MAA / F2 | Multifactorial Atonal Appreciation | Mencke 2019; Gold 2019b | Combines complexity, familiarity and contextual proxies into atonal-appreciation scores. | U |
| PSH / F2 | Prediction Silencing Hypothesis | Millidge 2022; Auksztulewicz 2016 | Combines hierarchical prediction scores into high-level suppression and low-level persistence proxies. | U |

The P pool contains PUPF.E0/E1/U0/U1/U2/P0, SRP.C2/T0/T1/M0, AAC.F0/F1, CDMR.f01/P0 and RPEM.rpe_magnitude. The A family contains AAC.F1, PUPF.U0, AAC.A1 and CSG.P0 for arousal, and SRP.P1, CMAT.S0, SRP.P2 and VMM.V0 for valence. H links BCH to the harmonic/bonang comparison, HTP/ICEM to the recorded pleasure comparison, MMP to the selected chills comparison, and DAED/MORMR to the ChillsDB–DEAM output contrast. The latter compares model outputs across datasets; it does not measure opioid release. These links concern the named outputs and protocols. They are not an exhaustive index of all channels in historical searches or the later full-suite families in S6. C marks the ICEM intervention (with PWUP/CHPI propagation) and the AAC arithmetic example and the DAP/MMP examples retained in the source audit. Full citations, output names, declared falsification criteria and source paths are preserved for every mechanism in the companion record. The crosswalk indexes the source, implemented operation and recorded evaluation link for each mechanism. An equation-level attribution to a particular empirical finding is supplied only where it has been checked, as in the examples below. Elsewhere that translation remains to be adjudicated; the controlled-access mechanism records preserve the declared input roles, output identities and proposed falsification criteria. These criteria are design records requiring prospective specification, not evidence that a test was completed or that its threshold is justified.

## From an empirical finding to an implemented proposal

**BCH.** Bidelman and Krishnan (2009) report neural and behavioural relations for consonance. The MI implementation combines explicit acoustic/temporal cues into consonance-related scores. Its mixture coefficients, four-stage organization and output names are additional modelling decisions; the occurrence of a neural effect in the source is not evidence that the current score measures a brainstem response. The consonance comparisons in this manuscript primarily concern specified acoustic descriptors and stimulus constructions, not a validation of every BCH output. Natural-timbre transfer and a matched mechanistic alternative remain separate tests.

**ICEM.** Egermann et al. (2013) studied expectation and psychophysiological emotion in a live-concert experiment. Their probabilistic information-content account uses the event probability conditional on context. The inspected MI extraction instead derives a score from acoustic context; it does not compute that event probability or return information in bits. The source audit and temporal intervention are summarized in Section 3 and S1. These characterize the recorded calculation and its propagation without publishing the implementation rule. Testing the proposed expectancy interpretation requires probability/violation measures and acoustic-change alternatives. The source history also records TenseMusic-guided changes, so the formula is not described as developmentally independent of that corpus.

**IUCP.** Gold et al. (2019b) examine predictability, uncertainty and musical pleasure using an information-theoretic expectation model. Their quadratic relationships motivate an intermediate-complexity proposal. MI uses a non-monotonic complexity-preference proposal; its response shape and combination rule are additional implementation assumptions that are not specified here. The source finding does not determine those choices. An appropriate test would compare the specified curve with alternative shapes and with the pleasantness input under the same human target and held-out protocol.

**AAC and MMP.** The AAC arithmetic restatement and its arousal association refer to a dimensionless acoustic-context score. The metadata entry assigning PET/raclopride dopamine findings to Salimpoor 2009 conflicts with the primary article (Salimpoor et al., 2011); the historical label is retained in the source audit with the discrepancy, not endorsed as a corrected reference. MMP's inspected familiarity-labelled path uses audio-derived information and does not receive an individual's listening history. Autonomic coupling and personal recognition therefore require their respective measured targets and listener variables. Neither the label nor an association with another target supplies that missing observation.

Other recorded translation issues remain traceable: the ICEM design note combines Cheung's 39-person behavioural and 40-person imaging samples into an incorrect single fMRI n=79; the primary study separates them Cheung et al. (2019). The historical HTP card describes a cross-domain visual-motion analogy, which must not be presented as a directly measured music-specific time constant. These checks delimit source roles without changing the archived code or silently repairing prior numerical results.

## Revision and prospective measurement

A proposed process remains useful only insofar as it produces distinguishable consequences. Sensory proposals require matched stimuli and perceptual targets; expectancy proposals require contextual probabilities or controlled violations; personal-memory proposals require exposure/history and recognition measures; autonomic proposals require synchronized physiological and subjective responses; motor proposals require tapping, phase correction or other appropriate movement outcomes. Regional and neuromodulator mappings require separately specified observation models and measurements. These are requirements for future tests, not completed experiments.

A revision record should identify the new evidence, affected mechanism/output, previous and revised computation, development data and version hash. The next version is frozen for evaluation, with matched alternatives, grouping and temporal access specified before outcome-based selection. Retaining, combining, narrowing, replacing, removing and extending proposals are all admissible outcomes. Parameter estimation and runtime state evolution remain distinguishable from changing the repertoire. Negative or inconclusive results remain attached to the tested version.

## S5b. Source roles and limits of the retained mechanism names

This focused editorial audit compares all 84 full-name expansions with the compact source set and retained naming records. It is not a systematic review, a study-quality meta-analysis or an empirical evaluation of MI. DOI-matched primary abstracts were inspected where available, and accessible full texts were examined for selected ambiguous mappings. The source-access inventory records abstract-only, full-text and unresolved cases. One relevant source is preferable to an unrelated second citation. A source can directly motivate a topic without supporting every word of a label or the corresponding equation. References added in this revision supply retrospective scientific context; they are not claimed to have determined the original implementation.

All current names remain unchanged to preserve correspondence with the engine. Historical expansions help explain source selection, but neither an earlier name nor the presence of a citation establishes the current biological interpretation. The phrases in the right column are explicit limits on interpretation, not changes to the implementation. Future renaming or computational revision would require a versioned change to the engine and its paper together. No model was executed, refitted or empirically re-evaluated for this editorial update.

Gold (2019a) denotes the PNAS reward-prediction-error study; Gold (2019b) denotes the Journal of Neuroscience predictability/uncertainty study. Removed anchors and all old-to-new selections are retained in the accompanying editorial record. The C/P/A/H/U evaluation flags in Table S5a retain their original meaning and are independent of this source-role audit.

**Table S5b. Source motivation and untested interpretation, by fixed identifier**

| Identifier | Published observation / source role | Limit on the retained label |
|---|---|---|
| BCH | Bidelman (2009) relates frequency-following responses to consonance; Bidelman (2013) reviews brainstem pitch processing. *Sources:* Bidelman 2009; Bidelman 2013. | The computational score is not a brainstem recording; the review is not an independent replication. |
| CSG | Bravo studies interval appraisal; Sarasso relates aesthetic responses to attention. *Sources:* Bravo 2017; Sarasso 2019. | The studies do not establish a universal monotonic salience gradient; intermediate dissonance can have low salience. |
| MIAA | Kraemer and Halpern link musical imagery to auditory cortical activity. *Sources:* Kraemer 2005; Halpern 2004. | An audio-only output does not observe whether a listener is imagining music. |
| MPG | Patterson supplies empirical evidence for cortical progression in pitch/melody processing; Rupp is an editorial. *Sources:* Rupp 2022; Patterson 2002. | The editorial supplies context, not an independent gradient experiment or validation of the implemented gradient. |
| PNH | Ratio-related consonance and early auditory processing motivate the label. *Sources:* Bidelman 2009; Tabas 2019. | Pythagorean is an interpretive label, not an established distinct neural hierarchy. |
| SDNPS | Cousineau examines stimulus dependence of neural pitch salience; Penagos supplies cortical pitch-salience evidence. *Sources:* Cousineau 2015; Penagos 2004. | The natural-sound limitation is part of the motivation; generalization across stimulus classes is not assured. |
| SDED | Early auditory responses differentiate consonance/dissonance under studied contexts. *Sources:* Crespo-Bojorque 2018; Fishman 2001. | Timing and expertise effects do not establish a context-independent dissonance detector. |
| TPIO | Halpern investigates timbre perception/imagery overlap; Zatorre reviews musical imagery. *Sources:* Halpern 2004; Zatorre 2005. | Overlap of activation does not show that an audio-derived score measures ongoing imagery. |
| TPRD | Allen directly compares pitch and tonotopic representations; Norman-Haignere characterizes pitch-sensitive cortex. *Sources:* Allen 2022; Norman-Haignere 2013. | Representations partly overlap; dissociation does not mean wholly separate cortical systems. |
| PSCL | Penagos localizes pitch-salience responses; Allen compares pitch representations. *Sources:* Penagos 2004; Allen 2022. | These anchors do not establish hemispheric lateralization. The historical label used Localization; the current engine label remains unchanged. |
| PCCR | Briley directly examines cortical pitch-chroma mapping; Allen supplies broader pitch-representation context. *Sources:* Briley 2013; Allen 2022. | Pitch chroma and tonotopic frequency are distinct constructs; the sources do not validate this model score. |
| HTP | Temporal integration windows and predictive-memory hierarchy supply complementary motivations. *Sources:* Norman-Haignere 2022; Bonetti 2024. | Combining them into this temporal-prediction model is a design choice; integration is not itself prediction. |
| SPH | The cited work concerns hierarchical spatiotemporal prediction and predictive processing. *Sources:* Bonetti 2024; Fernandez-Rubio 2022. | The implemented temporal/topological organization is not derived in full from those studies. |
| ICEM | Egermann links expectation violations with psychophysiology; Cheung links uncertainty and surprise with pleasure. *Sources:* Egermann 2013; Cheung 2019. | An unexpectedness-labelled acoustic score is not automatically a probabilistic information-content estimate. |
| PWUP | Mencke theorizes atonal appreciation and uncertainty; Cheung examines uncertainty, surprise and pleasure. *Sources:* Mencke 2019; Cheung 2019. | These anchors do not directly test precision weighting. The precision variable, weighting rule and acoustic translation require explicit model definitions and tests. |
| WMED | Noboa examines different contributions of entrainment and working memory to synchronization; Yuan studies auditory-memory load. *Sources:* Noboa 2025; Yuan 2025. | These findings are not a general double dissociation. WMED and NEWMD are overlapping proposals, not independent replications. |
| UDP | Uncertainty and atonal appreciation motivate a pleasure-related proposal. *Sources:* Mencke 2019; Gold 2019b. | The relation is not uniformly increasing; context and individual differences remain relevant. |
| CHPI | Audiovisual integration and harmonic expectation supply separate components. *Sources:* Moller 2021; Cheung 2019. | Their combination is a proposed synthesis. Relations among audio features are not evidence of visual-auditory transfer. |
| IGFE | Individual-frequency auditory stimulation motivates the label; related oscillatory evidence supplies context. *Sources:* Yokota 2025; Bolland 2025. | A personalized stimulation protocol does not establish that ordinary music or this score enhances gamma-mediated cognition. |
| MAA | Theoretical work on atonal appreciation addresses multiple listener/context factors; Gold supplies predictability background. *Sources:* Mencke 2019; Gold 2019b. | The named multifactorial model remains a theoretical synthesis, not a mechanism validated by both sources. |
| PSH | Predictive coding and response suppression motivate the hypothesis. *Sources:* Millidge 2022; Auksztulewicz 2016. | Silencing refers to a proposed reduction of prediction-related responses, not all neural activity. |
| SNEM | Nozaradan's studies concern selective entrainment to beat and meter. *Sources:* Nozaradan 2011; Nozaradan 2012. | They support the historical Selective Neural Entrainment Model, not a direct novelty test; the current expansion is retained as an identifier. |

Continued on next page

Table S5b (continued)

| Identifier | Published observation / source role | Limit on the retained label |
|---|---|---|
| IACM | Basinski directly examines inharmonicity and attention; Koelsch supplies tuning/mismatch context. *Sources:* Basinski 2025; Koelsch 1999. | Spectral inharmonicity and sequential uncertainty must not be treated as equivalent manipulations. |
| BARM | Rathcke concerns beat ability and speech timing; Grahn concerns beat-related motor-system activity. *Sources:* Rathcke 2024; Grahn 2007. | Neither study measures brainstem ABR/FFR modulation. The historical Beat Ability Regulatory Model explains the anchor lineage, not the current brainstem claim. |
| STANM | The sources examine spectrotemporal cues, selective attention and network organization. *Sources:* Haiduk 2024; Zatorre 2022. | Distributed associations do not establish this model as a distinct biological attention circuit. |
| AACM | Aesthetic appraisal and attention/motor inhibition motivate coupling. *Sources:* Sarasso 2019; Sarasso 2021. | Attention, inhibition and memory remain different outcomes; one does not validate every named output. |
| AMSS | Selective-attention studies concern auditory object and stream representations. *Sources:* Hausfeld 2021; Wikman 2025. | The relation is task-dependent and includes nonmusical objects; the model requires its own segregation task. |
| ETAM | Envelope tracking, musical timing and attention supply related motivations. *Sources:* Hausfeld 2021; Doelling 2015. | The current expanded label is broader than the historical envelope-tracking formulation; tempo modulation is not independently established for every output. |
| DGTP | Speech, rhythm and movement timing motivate shared temporal processing. *Sources:* Rathcke 2024; Grahn 2007. | Transfer between the studied domains does not establish unrestricted domain generality. |
| SDL | Spectrotemporal properties and attention relate to hemispheric asymmetry. *Sources:* Haiduk 2024; Zatorre 2022. | Salience is not interchangeable with either acoustic structure or attentional instructions. |
| NEWMD | Noboa examines entrainment and working-memory predictors of synchronization. *Sources:* Noboa 2025. | The unresolved Sares (2023) anchor is excluded. Different predictor associations are not a universal double dissociation or an independent replication of WMED. |
| PWSM | Basinski separates spectral from sequential uncertainty; predictive-processing accounts motivate precision weighting. *Sources:* Basinski 2025; Millidge 2022. | Its inharmonicity effect must not be cited as proof that every salience effect is precision-weighted prediction error. |
| MEAMN | The sources examine music-evoked autobiographical memory and associated networks. *Sources:* Janata 2009; Sakakibara 2025. | An acoustic recording does not supply an individual's history or demonstrate actual autobiographical recall. |
| MMP | Jacobsen examines preserved musical memory in Alzheimer's disease; Fang reviews music-based interventions. *Sources:* Jacobsen 2015; Fang 2017. | Preservation of one memory function does not imply intact general memory or a personal-familiarity measure from audio alone. |
| RASN | Rhythmic stimulation and auditory-motor rehabilitation motivate the proposal. *Sources:* Thaut 2015; Grahn 2007. | Clinical effects and entrainment do not uniquely identify the proposed neuroplasticity mechanism. |
| PMIM | Musical syntax, mismatch responses and stored regularities motivate prediction-memory integration. *Sources:* Koelsch 2000; Koelsch 2009. | The integration rule is a model synthesis; ERAN and MMN are not interchangeable observations. |
| OII | Oscillatory accounts of intelligence and auditory processing supply separate motivations. *Sources:* Bruzzone 2022; Samiee 2022. | Their integration is a proposed extension, not demonstrated transfer from a pitch score to intelligence. |
| HCMC | Complementary-learning-system and hippocampal-neocortical theories motivate the architecture. *Sources:* Squire 1995; McClelland 1995. | These are general memory accounts, not music-specific validation of a circuit or this implementation. |
| RIRI | Rhythmic auditory stimulation and rehabilitation provide the source topic. *Sources:* Thaut 2015; Harrison 2025. | Intelligent integration remains a proposed computational organization; a clinical benefit requires an appropriate intervention study. |
| MSPBA | Maess links music-syntactic processing with inferior frontal regions; Kim provides complementary syntactic-processing evidence. *Sources:* Maess 2001; Kim 2021. | Broca's area and its right homologue should not be reduced to a unique isolated music-syntax centre. |
| VRIAP | Deng directly compares active and passive VR modes, pain ratings and fNIRS measures in 15 healthy participants. *Sources:* Deng 2025. | This supports VR-analgesia motivation, not analgesia from an audio-only MI score. The intervention includes motor and visual differences. |
| CMAPCC | Auditory-motor action/perception studies motivate shared representations. *Sources:* Lahav 2007; Bangert 2006. | Coactivation supports a common-code hypothesis without establishing identical codes or the implemented integration. |
| DMMS | Developmental music exposure and auditory plasticity motivate a scaffold proposal. *Sources:* Partanen 2022; Qiu 2025. | Human and animal findings are distinct; a memory scaffold is an interpretation, not the directly measured endpoint. |
| CSSL | Burchardt studies tutor-song rhythm learning in zebra finches. *Sources:* Burchardt 2025. | The retained anchor is within one species. Cross-species learning/generalization remains unsupported by this compact source set. |
| CDEM | Sachs concerns context-dependent emotional transitions during music. *Sources:* Sachs 2025. | Context-sensitive emotion does not establish emotional-memory encoding or recall. Memory remains an untested part of the retained label. |
| SRP | Striatal music-reward responses and anticipation motivate a prediction-related proposal. *Sources:* Salimpoor 2011; Salimpoor 2013. | The model's prediction computation is an additional design choice, not a measured striatal signal. |
| AAC | Egermann measures psychophysiological expectation responses; Salimpoor relates pleasure, dopamine and autonomic arousal. *Sources:* Egermann 2013; Salimpoor 2011. | The implemented readout remains dimensionless and audio-derived; it is not a heart-rate, skin-conductance or dopamine measurement. |
| VMM | The sources investigate affective responses and musical valence. *Sources:* Mitterschiffthaler 2007; Koelsch 2006. | A unique mode-to-valence mapping is not isolated from other musical and listener factors by this source pair. |
| PUPF | Cheung examines uncertainty/surprise and pleasure; Egermann examines expectation-related psychophysiology. *Sources:* Cheung 2019; Egermann 2013. | Pleasure and arousal are separate targets; an arousal association does not validate a pleasure-prediction function. |
| CLAM | The sources concern closed-loop music/affect regulation protocols. *Sources:* Ehrlich 2019; Daly 2016. | An audio-only calculation is not an operating feedback intervention without sensed state and an adaptive control loop. |
| MAD | Musical anhedonia studies concern atypical auditory-reward connectivity. *Sources:* Loui 2017; Martinez-Molina 2016. | Disconnection does not imply complete absence of anatomical connections or a diagnosis obtainable from audio. |
| NEMAC | Music, nostalgia, autobiographical memory and affect motivate the proposal. *Sources:* Sakakibara 2025; Barrett 2010. | Circuit is an interpretive synthesis; listener history is required to study individual nostalgia. |

Continued on next page

Table S5b (continued)

| Identifier | Published observation / source role | Limit on the retained label |
|---|---|---|
| STAI | Spectral/temporal properties and aesthetic appraisal motivate integration; Alluri supplies acoustic-network context. *Sources:* Kim 2019; Alluri 2012. | The sources do not fix the particular integration equation or establish every aesthetic output. |
| DAP | Development and musical experience motivate plasticity. *Sources:* Scholkmann 2024; Trainor 2012. | Affective plasticity is more specific than the cited developmental auditory effects and needs affect-related developmental outcomes. |
| CMAT | Spence reviews crossmodal correspondences and supplies general conceptual background. *Sources:* Spence 2011. | The retained source does not establish learned music-visual affect transfer. The unrelated infant-speech anchor is removed; transfer remains a hypothesis. |
| TAR | Music-related affect and intervention literature motivate an affect-regulation proposal. *Sources:* Kheirkhah 2025; Ehrlich 2019. | Therapeutic resonance is not a uniquely demonstrated mechanism or evidence of treatment efficacy for this score. |
| DAED | Salimpoor distinguishes reward anticipation from peak pleasure; the second source provides complementary reward evidence. *Sources:* Salimpoor 2011; Gold 2023. | Different phases are supported, but not every cited study measures dopamine, and MI does not measure neurotransmitter release. |
| MORMR | Putkinen directly studies music reward and mu-opioid receptors; Salimpoor concerns dopaminergic reward. *Sources:* Putkinen 2025; Salimpoor 2011. | The second source is not an opioid replication; receptor availability/binding and release require distinct interpretations. |
| RPEM | Gold (2019a, PNAS) models musical reward prediction errors and learning; Gold (2023) reports RPE-like responses. *Sources:* Gold 2019a; Gold 2023. | This motivates the construct without establishing the same latent error or reinforcement-learning process in the MI equation. |
| IUCP | Gold (2019b, Journal of Neuroscience) studies predictability, uncertainty and liking. *Sources:* Gold 2019b; Gold 2023. | An inverted-U is not a universal preference law; the implemented response shape needs comparison with alternatives. |
| MCCN | Chabin supplies cortical EEG evidence related to chills; Putkinen supplies reward/opioid context. *Sources:* Chabin 2020; Putkinen 2025. | Neither establishes that every model output measures chills or a uniquely specified cortical network. |
| MEAMR | Autobiographical-memory and music-reward studies supply complementary components. *Sources:* Janata 2009; Salimpoor 2011. | Their coupling in a single proposed mechanism is a synthesis, not an independently tested integrated process. |
| SSRI | Kokal relates synchronized drumming to reward-localizer caudate activity, conditional on rhythm proficiency; Bigand concerns dyadic dance. *Sources:* Kokal 2011; Bigand 2025. | The integration rule is proposed. Reward-related activity and interpersonal synchronization are distinct observations, not an audio-only social-reward measure. |
| LDAC | The sources relate liking to auditory cortical responses. *Sources:* Gold 2023; Cheung 2019. | Dependent denotes an observed association/modulation, not proven causal necessity for the implemented readout. |
| IOTMS | Putkinen relates individual receptor measures to pleasure-dependent responses; other reward work supplies context. *Sources:* Putkinen 2025; Salimpoor 2011. | A second reward citation is not an independent opioid-tone result; receptor availability is not momentary release. |
| SSPS | Uncertainty-by-surprise interactions motivate a two-dimensional preference proposal. *Sources:* Cheung 2019; Gold 2019b. | The exact saddle geometry is a modelling choice, not an invariant response surface derived from the sources. |
| PEOM | Period entrainment and phase-correction studies motivate timing adjustment. *Sources:* Thaut 2015; Thaut 1998. | Optimization describes a computational proposal; the particular objective and coefficients are not automatically source-derived. |
| MSR | Bangert examines shared auditory-motor networks in musicians; Lahav studies action representations elicited by learned sounds. *Sources:* Bangert 2006; Lahav 2007. | Task/training-specific effects motivate reorganization, without proving a universal sensorimotor transformation or this model's score. |
| GSSM | The main source studies gait-synchronized electrical brain stimulation after stroke. *Sources:* Yamashita 2025. | That external stimulation protocol is not implemented merely by processing or listening to music; motor benefit is not inferred here. |
| HMCE | The source studies hierarchical neural encoding of musical context. *Sources:* Mischler 2025; Norman-Haignere 2022. | Encoding-model performance does not establish the implementation as a distinct discovered biological circuit. |
| HGSIC | Potes relates high-gamma activity to sound intensity; Grahn concerns beat-related motor activity. *Sources:* Potes 2012; Grahn 2007. | These anchors do not establish hierarchical groove-state integration. The historical High Gamma-Sound Intensity Coupling label explains the source lineage. |
| ASAP | Patel and Iversen explicitly propose Action Simulation for Auditory Prediction. *Sources:* Ross 2022; Patel 2014. | MI is an implementation inspired by that published hypothesis; its outputs are not an independent validation of ASAP. |
| DDSMI | Dyadic dance and joint musical-action studies motivate social-motor coordination. *Sources:* Bigand 2025; Kohler 2025. | Dance and piano tasks supply different evidence; the integrated model is not directly tested by both. |
| VRMSME | Liang studies VR music stimulation, motor observation and imagery with fNIRS. *Sources:* Liang 2025. | Network activation/connectivity is not demonstrated motor-performance enhancement. The motion-sickness source is removed. |
| SPMC | Grahn and Hoddinott concern beat-related motor/association systems, including SMA and putamen. *Sources:* Grahn 2007; Hoddinott 2024. | The specified serial SMA-premotor-M1 circuit and its directionality are not established by this pair. |
| NSCP | Leeuwis relates intersubject neural synchrony to subsequent streaming success. *Sources:* Leeuwis 2021. | The removed pupil/groove study is not commercial-prediction evidence. Audio-derived synchrony-labelled values are not measured intersubject EEG synchrony. |
| CTBB | Sansare studies cerebellar theta-burst stimulation and balance; Okada provides different cerebellar timing context. *Sources:* Sansare 2025; Okada 2022. | The monkey saccade study is not a balance/iTBS replication; an audio-derived score is not the external stimulation intervention. |
| STC | Zamorano relates singing training to connectivity; Kleber probes expertise and altered vocal feedback. *Sources:* Zamorano 2023; Kleber 2013. | Training associations and experimental feedback perturbation are distinct designs, not interchangeable causal estimates. |
| EDNR | Paraskevopoulos and Leipold concern musician-related network differences. *Sources:* Paraskevopoulos 2022; Leipold 2021. | Cross-sectional expertise differences do not by themselves establish training-induced causal reorganization. |
| TSCP | Pantev studies timbre-specific cortical responses; Bellmann synthesizes timbre neuroimaging. *Sources:* Pantev 2001; Bellmann 2024. | Neither measures synaptic consolidation. The historical Timbre-Specific Cortical Plasticity label and current identifier differ in explanatory level. |
| CDMR | Crespo-Bojorque examines context/expertise effects on mismatch responses; Rupp is an editorial. *Sources:* Rupp 2022; Crespo-Bojorque 2018. | The editorial is background, not a second mismatch experiment; MMN timing and presence depend on the studied context. |

Continued on next page

Table S5b (continued)

| Identifier | Published observation / source role | Limit on the retained label |
|---|---|---|
| SLEE | Paraskevopoulos examines statistical-learning expertise; Porfyri studies network changes after four weeks of training. *Sources:* Paraskevopoulos 2022; Porfyri 2025. | MEG/EEG network findings do not measure synaptic long-term encoding. The historical Statistical Learning Expertise Enhancement interpretation is narrower. |
| ESME | Koelsch and Vuust show expertise-related mismatch responses to musical features. *Sources:* Koelsch 1999; Vuust 2012. | They do not directly test error-gated encoding or learning. The historical Expertise-Specific MMN Enhancement label is narrower than the current expansion. |
| ECT | Paraskevopoulos reports greater within-network and reduced between-network communication; Porfyri studies training-related connectivity. *Sources:* Paraskevopoulos 2022; Porfyri 2025. | Compartmentalization does not establish a necessary behavioural cost or general trade-off. |

## Supplement S6. Public-data coverage and access decisions

The study uses secondary human data; no new purpose-designed participant experiment was collected under the project's available research conditions. This section documents the practical boundary of the recorded public-data search rather than asserting that the field contains no further usable datasets. Availability was assessed for a particular output, observation and stimulus correspondence. Audio availability, target suitability and completed analysis are separate decisions.

The 30 September 2026 review inventoried 159 existing local dataset/cache directories and searched prior analysis and manuscript records before considering additional releases, prioritizing 2025–2026. This does not prove that a source had never been used anywhere in the project or that music recordings do not overlap across datasets. New release date, data-collection date and new-to-project status were distinguished. Prior-use search records, dataset registries, download logs, mapping decisions and the full-suite method are retained with the source package.

Reuse constraints are documented independently of MI. In Eerola's (2025) sample of music-psychology publications from 2017–2022, only two of ten accessible datasets appeared complete and clearly documented. This historical sample provides context, not a census of 2026 data availability. The ManyMusic-Stim project explicitly provides open audio, metadata and human ratings while addressing stimulus-selection and copyright constraints. The FME-24 repository instead states that original audio is omitted because of copyright. These cases illustrate why missing access and a deferred available resource must receive different entries. Existing consonance, chord-pleasure, film-emotion, TenseMusic, chills, PMEmo and neural comparison records retain their individual stimulus and timing qualifications in S1–S4 and S7, with separate historical families in the technical archive. Table S6a concerns the additional recent-dataset review; it is not a replacement for that prior-use inventory. A missing original recording is also distinct from a historical experiment evaluated using an explicitly disclosed proxy synthesis.

**Table S6a. Recorded access and eligibility decisions (30 September 2026)**

| Dataset / source | Recorded state | Access or mapping boundary | Consequence for the proposed test |
|---|---|---|---|
| MUSIFEAST-17 | Processed; partial response mapping | 356/356 released recordings; 32 response rows lack a released stimulus match. | Excerpt-level affect, enjoyment and familiarity targets; direct transfer and fitted comparisons remain distinct. |
| Lucerne Groove Library | Processed except one file | 443/444 waveforms; one recorded HTTP 500 failure (SheW+4). | Pleasure, energy, movement, regularity and interest; source groups retained. |
| MEET | Partially released; processed | 38/116 rated stimuli have released audio; accuracy inconsistencies and perceived/felt wording remain. | Continuous affect comparisons use the accessible subset; classification accuracy held back. |
| GlobalMood | Original experiment segments unresolved | 1,180 local song identities; exact random 15-s stimuli/start times not recovered. | Ratings audited; full-track links do not identify the rated excerpt, so no MI comparison. |
| FME-24 | Original audio unavailable | 422 excerpts; repository omits recordings for copyright; 78 published features available. | Published-feature comparisons only; no MI inference from unavailable recordings. |
| ManyMusic-Stim | Available; deferred | 30.4-GB archive not acquired in this run. | Identify rated subset and exact timing before processing; not a missing-audio case. |
| AIME | Source card checked; deferred | 15,600 pairwise comparisons; no completed MI comparison. | Separate music preference from text–audio alignment and group by generator/prompt. |
| SongEval | Source card checked; deferred | Whole-song judgements; no completed MI comparison. | Short excerpt is not the same target; memorability rating is not delayed recall. |
| Affective surprise–memory | Separate event-alignment task; deferred | 2026 release not acquired in this run. | Align object events and music for recognition/source-memory hypotheses. |
| Sonic Seasoning | Outside primary target scope | 377-record mixed collection; historical and generated stimuli must be separated. | Music–taste associations require a distinct test; not treated as affect/groove replication. |
| PGGT | Exact audio access not verified | 2025 release of a 2024 survey. | Pop-music groove candidate; resolve stimulus correspondence before MI comparison. |

### Mapping, target and licence qualifications

MUSIFEAST's raw response table contains an additional stimulus identifier, 60s_LOW_08, with 32 rows that do not map to the 356 released audio/metadata records; those rows were retained outside the MI mapping. The Lucerne access failure concerns one identified waveform, not the other 443 successfully processed files. MEET contains ratings for 116 stimuli but publishes 38 corresponding audio files. Among the 7,600 response-code/accuracy pairs examined for those files, 441 disagree; the accuracy target was held back. The repository and workbook also differ between perceived and felt emotion wording. A nominal emotion code was not averaged into a continuous affect score.

GlobalMood's downloaded tables contain 1,180 song identities, 988,925 raw ratings and 117,996 mean-rating rows. The exact random fifteen-second experiment excerpts or their starting times were not available in the inspected releases; a full-track YouTube identifier was not substituted for the rated segment. The local 1,180 count differs from a web summary of 1,182. Four expected mean-label cells are absent, and the number of participant identifiers also differs from the reported rating-phase sample. These unresolved source issues remain in the integrity report. FME-24 supplies 1,784 annotation rows, 91 columns and 78 published features for 422 excerpts, but not the original audio. Its feature-only comparisons are therefore not MI evaluations.

Source licences and redistribution permissions are retained separately from technical accessibility. MUSIFEAST's data licence does not make all commercial audio unconditionally CC0. An explicit redistribution licence was not recovered for Lucerne in the local review; MEET reports CC BY 4.0. The manuscript does not supply new permissions or redistribute these recordings. The broader neural eligibility inventory contains 32 records across modalities; it is not a set of 32 suitable independent fMRI/PET datasets.

### Completed processing and outcome-specific comparisons

The separate recent-dataset pipeline records processing of 837 available audio files: 356 MUSIFEAST, 443 Lucerne and 38 MEET. Each saved record contains the available 1,791-channel collection, including 889 outputs from 84 unique mechanisms. Output production is not equivalent to matched human validation of every channel. This run includes native belief observations and model-derived RAM/neuro outputs, but not Bayesian posterior loops or the context-dependent reward formula.

Table S6b reproduces the fixed-family comparisons from the saved full-suite report. MUSIFEAST uses the 65-excerpt held-out partition; Lucerne and MEET use pilot-excluded nested grouped cross-validation. The saved protocol was fixed after the initial pilots and before these comparisons; it is

a local scope lock, not a prospective registration. Imputation, scaling, categorical encoding and ridge selection use training data only, with four inner group folds and alpha in {0.1, 1, 10, 100, 1000}. MUSIFEAST holds artist/source/audio-related groups together. Lucerne and MEET hold source works together and exclude pilot-related works, but may share artists across folds. Listeners are not separately held out. The intervals use 2,000 source-group bootstrap draws of fixed predictions; they neither refit the model nor quantify new-participant generalization. These fitted comparisons are distinct from the direct-channel associations in Table S3a. Lucerne's movement/regularity/interest families are exploratory selections over 132 F7 channels. The MEET $R^3$-97 baseline is unstable at n=30; a positive delta from a strongly negative baseline does not imply good absolute prediction.

The alternative prespecified four-acoustic-feature MEET comparison gives arousal $R^2$ 0.318 to 0.523 and valence -0.298 to -0.122, illustrating control sensitivity. FME's raw published-feature comparison was disrupted by two feature values of magnitude $10^9$ in one annotation. A separately labelled post-result robustness check used training-fold-only 1st–99th percentile feature limits, leaving observations and targets unchanged. That feature-only check is not evidence for MI. These outcomes demonstrate why unavailable observations, limited datasets and tested weak increments must remain separate categories.

**Table S6b. Saved full-suite primary-window comparisons**

| Dataset / target / family | n | Baseline $R^2$ | +MI $R^2$ | $\Delta R^2$ [95% interval] |
|---|---|---|---|---|
| MUSIFEAST-17 / valence / paper_valence_leaders | 65 | 0.301 | 0.341 | +0.040 [-0.045, +0.109] |
| MUSIFEAST-17 / arousal / paper_arousal_leaders | 65 | 0.586 | 0.569 | -0.016 [-0.087, +0.036] |
| MUSIFEAST-17 / enjoyment / archive_liking_top3 | 65 | 0.700 | 0.703 | +0.003 [-0.028, +0.024] |
| MUSIFEAST-17 / enjoyment / paper_native_pleasure | 65 | 0.700 | 0.698 | -0.002 [-0.011, +0.009] |
| Lucerne / PLE / paper_native_pleasure | 426 | 0.237 | 0.254 | +0.017 [-0.008, +0.049] |
| Lucerne / PLE / archive_liking_top3 | 426 | 0.237 | 0.230 | -0.007 [-0.019, +0.003] |
| Lucerne / ENE / paper_arousal_leaders | 426 | 0.612 | 0.611 | -0.001 [-0.006, +0.005] |
| Lucerne / MOV / exploratory_F7_MOV | 426 | 0.486 | 0.493 | +0.007 [-0.021, +0.035] |
| Lucerne / REG / exploratory_F7_REG | 426 | 0.372 | 0.361 | -0.010 [-0.027, +0.007] |
| Lucerne / INT / exploratory_F7_INT | 426 | 0.295 | 0.296 | +0.001 [-0.023, +0.025] |
| MEET / valence / paper_valence_leaders | 30 | -0.475 | -0.432 | +0.043 [-0.012, +0.109] |
| MEET / arousal / paper_arousal_leaders | 30 | -7.839 | -7.791 | +0.048 [+0.012, +0.118] |

Baseline follows the saved source comparison: MUSIFEAST combines four acoustic features, MIR33, genre and two familiarity ratings; Lucerne combines the four acoustic features, published numeric rhythm descriptors, style and familiarity; MEET combines the four acoustic features and 97 $R^3$ means. The four acoustic features are RMS, spectral centroid, zero-crossing rate and onset-envelope variability. Human familiarity is an explicit covariate, so these are not uniformly audio-only baselines. PLE, ENE, MOV, REG and INT denote pleasure, energy, movement, regularity and interest. Deltas and intervals are copied at the report's precision; differences of displayed rounded scores need not equal the displayed delta. The source records all sensitivity windows and alternative controls. No fitting or significance testing was performed for this table.

## Source record and access date

The linked mechanism–evidence package retains machine-readable rows for every mechanism, historical catalogues, the 448-record extraction index, source hashes and document checks. The dataset record retains DATASET_RESEARCH, FULL_DATASET_RESULTS, FULL_SUITE_METHOD and the full integrity audit, all dated 30 September 2026. Access states are dated observations; they are not promises that a repository will remain unchanged. Detailed calculation paths and unresolved source questions remain inspectable without converting editorial or schema checks into scientific validation.

Detailed evaluation tables and source roles are retained in this supplement. Implementation-level audit records are available through the controlled-access route stated in the article.

## Supplement S7. Protocols for additional retained figure panels

Main Figures 6g,h and 7a–e retain the full plotted families. This section preserves their human targets, channel selection, weaker outcomes and analysis limits. These records are separate from the compact fitted tension comparison in S2 and direct arousal transfer in S3.

### Continuous tension and reported chills

The archived analysis tests changing tension within a piece, rather than a correlation across whole-piece means. Across 38 TenseMusic pieces with 30 raters and 50-Hz continuous ratings (Barchet et al., 2024), `AAC.F1:hr_pred_2s` (autonomic-forecast-labelled output) had an archived Fisher-z mean Spearman correlation of +0.421 with consensus tension in the lag-selected within-piece analysis. All 15 channels in the specified tension pool met its Bonferroni criterion, with positive forecast associations and negative associations for selected prediction-error and mismatch readouts (Figure 7c,d). The forecast was positive in 34 of 38 pieces. The +0.421 and 34/38 values describe that historical procedure; the later timing audit does not establish them as a final, clock-verified positive-tracking estimate (S7).

Lags span ±5 s in 0.5-s steps. The largest absolute correlation selects the lag, while the reported coefficient retains its sign; selection repeats in 200 circular-shift nulls. These are within-piece associations, not held-out forecasts. Fisher-combined significance need not imply consistent direction: SRP harmonic tension averages +0.018, with 19/38 positive pieces, despite corrected $p = 3.04 \times 10^{-26}$. The separate ridge leave-one-piece-out median (+0.101) and human inter-rater agreement (+0.386) use different comparisons.

AAC reconstruction passed tolerance on all 38 cached inputs. Float64 M14 arithmetic changed its aggregate correlation by $-3.96 \times 10^{-8}$, retaining 34 positive pieces; null probabilities were not rerun (this section).

In clean-seven ChillsDB (Schoeller et al., 2023), `MMP.P2` (familiarity-labelled output) was elevated within ±5-s event windows: mean rank-biserial r = +0.2307, Bonferroni p = 0.0072, positive in all seven stimuli. This detailed table supersedes the earlier summary manifest. Nine-stimulus and alternative-denoising effects were +0.1599 and +0.2287, with corrected p = 0.4526 and 0.4775. The latter also changes the 8-s exclusion, event/frame masks, effect aggregation and null p calculation; it is not a denoising-only comparison (this section). `AAC.IO:chills_intensity` gives +0.0378, corrected p = 1. These results support neither a measured familiarity mechanism nor autonomic physiology: neither was independently observed.

The physical audio/annotation clock relation and behavioural response lag are separate. The audit identified a Schoenberg annotation cross-mapping, invalid initial samples and normalized-time stretching in the historical evaluation. A source-verified clock origin and a specified directional or fixed-lag rule are required before accepting a replacement positive-tracking coefficient. The retained numerical-stability check above addresses arithmetic, not these alignment questions.

### Emotion ratings across excerpts and over time

We contrasted emotion judgements about whole excerpts with ratings that vary during listening. In the 110-entry film-soundtrack set of Eerola and Vuoskoski (2011), all eight specified emotion pools contained Bonferroni-significant associations. For selected leading channels, original and acoustic-residual correlations were +0.741 → +0.430 for sadness, +0.722 → +0.447 for tenderness, and −0.683 → −0.504 for tension (Figure 6g). Targets are perceived musical emotion, not felt emotion or GEMS scales; model labels do not measure physiology.

Ratings were predicted from 97 clip-averaged $R^3$ features with five-fold ridge regression; each out-of-fold prediction was subtracted before correlating residuals with mechanism outputs. Clip averaging removes temporal information, and whole-set channel selection limits generalisation. Seven of eight selected leading readouts retained residual $|\rho| \geq 0.10$. This effect-size criterion differs from corrected significance: the selected anger and happiness readouts did not pass residual Bonferroni correction. Every full pool contained at least one residual $|\rho| \geq 0.10$ in the detailed tables. In the 360-excerpt pilot set, three leading channels were identical across sets; the other five leading associations used different members of the specified pools (Figure 6h). The 110-entry Set 2 was developed from the 360-entry Set 1 pilot, so this is overlapping cross-set correspondence rather than independent held-out replication.

The archived dynamic-emotion test was less consistent. Its 2-Hz ratings define within-excerpt trajectories rather than one value per excerpt. In the 200-selected-excerpt PMEmo pilot (Zhang et al., 2018), leading channels had 192 usable excerpts. No arousal channel among 19 met Bonferroni correction, compared with three of 15 valence channels (Figure 7a,b). Leading mean correlations were +0.1619 for the arousal-labelled output and +0.1197 for wanting. The separate human reliability analysis used 740 excerpts and cannot turn these results into matched ceiling estimates.

Positive excerpt-level associations and weak dynamic effects answer different questions. The observation units and timing requirements remain distinct. PMEmo timing and finite-sample limitations are recorded alongside the preserved historical results in S7.

#### Why the chills protocols are not interchangeable

TC005 uses all reported events and frames, ±5-s windows, 500 uniformly placed event-time nulls over the complete record, seed 2026051205, and a Fisher-z mean effect. Its empirical p is (1 + count(null ≥ observed))/(1 + number of valid nulls). TC006 excludes events at or before 8 s and masks the first 8 s of frames, constrains null centres to keep windows within the remaining interval, uses seed 2026051206, and reports an arithmetic mean effect. Its p is a raw tail fraction without the +1 correction; the tail changes for negative observed effects. The shared Fisher helper discards p = 0 values. Thus the number of pieces entering the combined statistic can differ even when the mean effect is similar. Both apply channel-family Bonferroni correction, but they are not otherwise the same test.

The aggregate tables do not identify each discarded p or isolate the contribution of each procedural difference. A matched sensitivity rerun is required to attribute the change quantitatively. The alternative result is retained as a protocol-dependent observation, not a denoising-only robustness test.

#### AAC F1: targeted numerical sensitivity

The preserved audit fixed all 38 pieces, the AAC.F1 output and the original signed effect summary before inspecting the numerical alternative. Conditional reconstruction from the retained inputs agreed with the cached outputs within the recorded tolerance. The targeted precision change left the signed Fisher-z mean association effectively unchanged: 0.4206821003 for the reference and 0.4206820608 for the alternative, with 34 positive pieces in each. This does not settle other numerical regimes or AAC's physiological interpretation. Circular-shift null probabilities were not recomputed; archived corrected probabilities are not assigned to the alternative. Implementation-level numerical definitions and reconstruction records remain under controlled access.

#### Execution identity and retained gaps

For the central Swan Lake example, the audio checksum matches the recorded input, all 852 Python files in the preserved snapshot match the original selection lock, and all 84 selected implementation paths and their ordered inventory agree with recorded execution. Saved $R^3/T^3$ arrays, reference/intervention outputs, successful-read traces, calculation audits, display arrays and figure scripts form a separately hashed chain. The prior executor replay exactly reproduced the saved $C^3$/RAM/proxy outputs from saved $R^3/T^3$ inputs. It was not a repeat of audio extraction. Missing-lookup behaviour is recoverable from locked code and recorded availability, but there is no complete dynamic log of every fallback branch. The controlled $R^3/T^3$ figures have a second chain of 86 verified preserved input/source records, synthesis definition, intermediate arrays and display queries.

The historical TenseMusic cache is less fully identified. All 38 clip checkpoints and the build pin cite d74d3ee760c5…, whereas the dataset manifest gives e537f90e6d14…. The currently available source tree has aggregate 482ade45c50f… under the recorded hashing algorithm. This is a disagreement among identity records, not interchangeable names for one verified execution. The retained builder instantiates 89 mechanisms before name-keyed overwrites produce 84 arrays; the current 84-name inventory cannot substitute for its caller. Successful AAC numerical reconstruction does not determine all historical duplicates, execution paths

or fallback use. The film cache also stores 84 mechanism arrays and 889 columns; these cardinalities alone identify neither implementation selection nor execution order.

The current canonical inventory records duplicate STAI selection differences, same-depth HCMC/MSPBA inputs unavailable to OII, and belief precision requests absent from the canonical mechanism-only demand collection. Those facts cannot fill absent run-level logs. Selected implementation paths, actual execution order and fallback consumption remain explicitly unresolved for archived families where they were not recorded. Future runs need to store these fields with the outputs.

**Timing audit and the status of the archived tension coefficient**

The archived AAC.F1 coefficient +0.421, its 34/38 positive-piece count and the 15-channel significance summary are retained with the historical tables and figures. They used normalized-time interpolation and lag selection by maximum absolute Spearman correlation. The later source audit verified all 38 audio identities against original OSF SHA-256 metadata, identified the Schoenberg/Schoenberg_2 annotation cross-mapping and found invalid initial consensus samples. Dropping an invalid sample must use the same time mask for both aligned series; it must not independently reset either clock. These findings limit the interpretation of the retained coefficient and count.

The original Barchet analysis joins feature and rating time columns and applies a 4.5-s reporting-delay shift. This analysis convention does not establish the physical audio-to-rating clock offset. The inspected public source files did not supply a sound-start event or the experiment clock-reset relation. A duration excess or first valid mouse sample alone cannot identify audio onset. No replacement primary coefficient is accepted until that relation and the directional or fixed-lag rule are specified. Selecting the largest signed correlation would require repeating that same selection in every null draw; probabilities from the historical absolute-maximum test cannot be reused. Post-result corrective choices remain exploratory.

A negative selected correlation can identify an inverse association, but is not evidence for a pre-specified positive-tracking claim. The historical combined absolute-association test and a directional tracking test therefore have different null hypotheses. The small point-estimate change from a finite-value mask also does not imply that missing-value handling is neutral for the null probabilities.

**Dynamic emotion and resolution limits**

The PMEmo values in the Figure 7a,b remain those of the archived 200-selected-excerpt pilot, with 192 usable excerpts per leading channel. The subsequent audit found that target timestamps and the 15-s annotation exclusion must be represented explicitly; short-record eligibility, constant channels and undefined observed or surrogate correlations also require explicit handling. Corrective calculations are separately recorded rather than substituted for the deposited numbers. Weak or null dynamic arousal results cannot be explained away by its 2-Hz grid, and positive valence effects do not establish 172-Hz human validation. Establishing a resolution effect would require the same stimuli, target, filtering, time support, lag rule and dependence-aware evaluation across controlled resolutions.

**Table S7a**. All historical within-piece tension channels (Figure 7c,d).

| Channel | Pieces | Fisher-z mean ρ | Positive pieces | Bonf. p |
|---|---|---|---|---|
| AAC.F1:hr_pred_2s | 38 | +0.42068 | 34 | 1.42e-37 |
| AAC.F0:scr_pred_1s | 38 | +0.38834 | 34 | 1.83e-37 |
| SRP.T0:tension | 38 | +0.31927 | 32 | 1.51e-31 |
| SRP.T1:prediction_match | 38 | -0.18903 | 9 | 8.01e-31 |
| CDMR.P0:mismatch_signal | 38 | -0.16817 | 8 | 4.04e-28 |
| SRP.C2:prediction_error | 38 | -0.13595 | 11 | 1.6e-28 |
| PUPF.U0:entropy_H | 38 | +0.10351 | 29 | 5.86e-31 |
| RPEM.rpe_magnitude | 38 | -0.09866 | 7 | 1.03e-32 |
| CDMR.f01:mismatch_amplitude | 38 | -0.05143 | 17 | 2.73e-26 |
| PUPF.U2:HS_interaction | 38 | +0.03912 | 23 | 1.33e-23 |
| PUPF.E0:prediction_err | 38 | -0.02958 | 17 | 5.82e-27 |
| PUPF.U1:surprise_S | 38 | +0.02266 | 29 | 1.14e-26 |
| PUPF.P0:surprise_pleasure | 38 | -0.02175 | 16 | 6.86e-25 |
| SRP.M0:harmonic_tension | 38 | +0.01803 | 19 | 3.04e-26 |
| PUPF.E1:uncertainty | 38 | -0.01611 | 19 | 5.96e-22 |

**Table S7b**. All chills channels under three overlapping protocols (Figure 7e).

| Channel | Clean 7 effect | Bonf. p | Full 9 effect | Bonf. p | Alternative 7 effect | Bonf. p |
|---|---|---|---|---|---|---|
| AAC.P0:current_intensity | +0.03005 | 0.926 | +0.01677 | 0.044 | +0.02403 | 0.13 |
| AAC.I0:chills_intensity | +0.03783 | 1 | +0.01370 | 0.142 | +0.02833 | 0.419 |
| AAC.I1:ans_composite | +0.08154 | 0.0454 | +0.05585 | 0.0471 | +0.06831 | 1 |
| AAC.F0:scr_pred_1s | +0.10086 | 0.0204 | +0.06441 | 0.168 | +0.07289 | 1 |
| AAC.F1:hr_pred_2s | +0.13269 | 0.0319 | +0.08697 | 1 | +0.10262 | 0.308 |
| AAC.E0:emotional_arousal | +0.16972 | 0.0141 | +0.11255 | 0.472 | +0.14685 | 0.0533 |
| AAC.E1:ans_response | -0.01559 | 1 | -0.02140 | 0.102 | -0.01339 | 0.654 |
| AAC.P1:driving_signal | -0.00263 | 1 | -0.00865 | 0.789 | -0.00789 | 0.706 |
| AAC.P2:perceptual_arousal | +0.11439 | 0.0557 | +0.07355 | 0.0235 | +0.09031 | 1 |
| DAED.f01:anticipatory_da | -0.02741 | 1 | -0.03245 | 0.148 | -0.02632 | 0.222 |
| DAED.f02:consummatory_da | -0.11818 | 1 | -0.14530 | 0.0536 | -0.12420 | 0.234 |
| DAED.f03:wanting_index | -0.11135 | 1 | -0.11664 | 0.0177 | -0.09599 | 0.162 |
| DAED.f04:liking_index | -0.11569 | 1 | -0.14176 | 0.0637 | -0.12085 | 0.162 |
| SSRI.P1:endorphin_proxy | -0.02899 | 1 | -0.04219 | 0.243 | -0.02998 | 0.316 |
| MMP.P2:familiarity | +0.23070 | 0.00716 | +0.15990 | 0.453 | +0.22870 | 0.477 |
| SRP.M0:harmonic_tension | -0.00226 | 1 | -0.01028 | 0.87 | -0.01102 | 0.4 |

Table S7b (continued)

| Channel | Clean 7 effect | Bonf. p | Full 9 effect | Bonf. p | Alternative 7 effect | Bonf. p |
|---|---|---|---|---|---|---|
| SRP.M2:peak_detection | -0.03909 | 1 | -0.04616 | 0.145 | -0.02524 | 0.107 |
| SRP.P2:pleasure | -0.08482 | 1 | -0.10917 | 0.0992 | -0.08255 | 0.217 |
| BCH.P0:consonance_signal | -0.03674 | 1 | -0.05556 | 0.158 | -0.03408 | 0.264 |
| PUPF.P0:surprise_pleasure | -0.00412 | 1 | -0.00299 | 1 | +0.00493 | 1 |
| UDP.M1:pleasure_index | -0.08371 | 1 | -0.10802 | 0.225 | -0.07378 | 0.392 |
| MORMR.f02:chills_count | +0.00079 | 1 | +0.01674 | 1 | +0.01036 | 1 |

The clean-seven column uses the TC005 Fisher-z mean; the alternative seven-stimulus protocol changes denoising, exclusion, masks, aggregation and null handling as described above. The columns are not independent replications and are not a denoising-only sensitivity test. The complete event and per-piece records remain in Source Data S4.2.

**Table S7c**. Complete dynamic-emotion families (Figure 7a,b).

| Target | Channel | Clips | Fisher-z mean ρ | Positive clips | Bonf. p |
|---|---|---|---|---|---|
| Arousal | AAC.E0:emotional_arousal | 192 | +0.16192 | 118 | 1 |
| Arousal | AAC.A1:hr | 192 | +0.13218 | 123 | 1 |
| Arousal | AAC.E1:ans_response | 192 | -0.08508 | 80 | 1 |
| Arousal | AAC.A0:scr | 192 | -0.07061 | 88 | 1 |
| Arousal | CSG.E1:sensory_evidence | 192 | -0.06809 | 84 | 1 |
| Arousal | PUPF.E1:uncertainty | 192 | -0.06319 | 83 | 1 |
| Arousal | AAC.P2:perceptual_arousal | 192 | +0.06265 | 110 | 1 |
| Arousal | AAC.P1:driving_signal | 192 | -0.05878 | 86 | 1 |
| Arousal | AAC.P0:current_intensity | 192 | -0.04893 | 92 | 1 |
| Arousal | PUPF.U0:entropy_H | 192 | +0.04522 | 105 | 1 |
| Arousal | PUPF.E0:prediction_err | 192 | +0.04205 | 107 | 1 |
| Arousal | AAC.F1:hr_pred_2s | 192 | +0.03769 | 106 | 1 |
| Arousal | AAC.F0:scr_pred_1s | 192 | +0.03646 | 104 | 1 |
| Arousal | MCCN.E2:arousal_index | 192 | +0.03274 | 106 | 1 |
| Arousal | AAC.A2:respr | 192 | +0.03015 | 102 | 1 |
| Arousal | CSG.P0:salience_network | 192 | +0.03014 | 104 | 1 |
| Arousal | TAR.T0:arousal_mod_tgt | 192 | -0.02961 | 88 | 1 |
| Arousal | PUPF.U1:surprise_S | 192 | -0.01499 | 93 | 1 |
| Arousal | CSG.E0:salience_activation | 192 | -0.00232 | 99 | 1 |
| Valence | SRP.P0:wanting | 192 | +0.11967 | 121 | 0.00379 |
| Valence | VMM.R0:happy_pathway | 192 | +0.08945 | 104 | 0.0183 |
| Valence | SRP.P2:pleasure | 192 | +0.07485 | 110 | 1 |
| Valence | CMAT.S0:supramodal_valence | 192 | +0.06430 | 108 | 0.515 |
| Valence | CLAM.P1:valence_tracking | 192 | +0.03674 | 101 | 1 |
| Valence | VMM.V0:valence | 192 | +0.03025 | 99 | 0.803 |
| Valence | VMM.V1:mode_signal | 192 | +0.02672 | 96 | 0.0237 |
| Valence | SRP.P1:liking | 192 | +0.02257 | 97 | 1 |
| Valence | CSG.P1:affective_evaluation | 192 | -0.02179 | 38 | 1 |
| Valence | PUPF.G0:pleasure_P | 192 | -0.01778 | 89 | 1 |
| Valence | CSG.M1:rt_valence_judgment | 192 | +0.01656 | 91 | 1 |
| Valence | VMM.R1:sad_pathway | 192 | -0.01654 | 92 | 1 |
| Valence | CSG.E2:consonance_valence | 192 | +0.00986 | 45 | 1 |
| Valence | CSG.F0:valence_pred | 192 | +0.00766 | 76 | 1 |
| Valence | UDP.M1:pleasure_index | 192 | +0.00711 | 97 | 1 |

All eight original film-emotion pools, their residual comparisons and the 360-entry pilot tables are retained in the emotion source records. Main Figure 6g selects the largest absolute original association within each 110-entry pool; Figure 6h permits a different leading channel in the overlapping pilot, identifying changed selections by marker shape. Complete tables, rather than only these maxima, define the comparison family. The R2 source index identifies all files and the figure-panel mapping.

# Supplement S8. Controlled birdsong transformations and computational audit

## S8.1 Design, source and chronology

The accompanying manuscript, *Reference-dependent $C^3$ responses to birdsong-to-music transformations* (Erdem, 13 October 2026; revised abstract edition), reports a computational assay of one 2.449977-s common-blackbird phrase, annotated as eight events. Its five primary paths join bird to synthetic transcription, transcription to flute or piano, and bird directly to the same flute or piano endpoints. Each primary path contains 97 anchors. Three isolated operator paths use 25 anchors each; a $3 \times 3 \times 3$ factorial design uses 27 settings. Overlapping conditions share files: 569 primary waveforms plus 569 common-RMS controls and eight ±6-dB controls give 1,146 recordings. Twenty-four diagnostic endpoint probes are additional follow-ups, not part of that total. These are dependent renderings of one source; there are no new participant or neural observations.

The protocol is exploratory. Initial 49-anchor paths were refined after midpoint-error inspection; direct bird–instrument paths and the 24 endpoint probes were added subsequently. The accompanying amendments preserve this chronology. The 13 mechanisms in Figure S8 are selected after inspecting the data, as the union of the five largest P-layer path-RMS responses on each primary matched-level path. The full 84-mechanism tables define the response set.

Pitch, timing and spectral replacement are renderer operations, not perceptually orthogonal axes. The instrument paths combine sampled waveforms and the transcription with equal-power weights, different attacks and morph-dependent releases (65 ms for flute; 120 ms for piano). Requested pitches span G6–F8, including values outside ordinary instrument ranges. The endpoints are pitch-shifted, sample-derived renderings. The bird source is bandwidth-limited by its 22.05-kHz recording; instrumental renderings contain wider bandwidth. Whole-phrase RMS matching (target 0.0188348729) does not match perceived loudness, event-wise envelopes or spectral energy. No octave folding is used.

## S8.2 Frozen execution and output selection

The strict adapter decodes rendered 44.1-kHz files to float32 arrays and averages stereo channels. It uses a 2,048-point FFT, 256-sample hop, 128 mel bands, power spectra, log1p and per-clip maximum normalization. The full phrase is processed once per condition, without loop padding or state carried between anchors. Every run checks 84 names, 889 dimensions, 423 frames and finite outputs. The adapter uses the canonical executor, retains discovery order within depth and takes the maximum of declared processing depth and role depth.

Of 88 exported classes, first-function-wins selection retains 84 names. Collisions retain F2 IGFE, F1 PNH, F1 TPRD and F2 MAA; later same-name variants are excluded. The P outcome comprises 202 dimensions selected by declared layer ranges. A name-prefix filter would retain only 187: fifteen P-layer dimensions in eight mechanisms lack that prefix. The analysis uses the registry, not that inadequate filter. Raw $C^3$ values may be signed and are not probabilities.

The 131 weighted belief aggregates are retained with clip-wise percentile stretching disabled. Eight separate recursive classes are not executed in this assay. The 26 regional and four neuromodulator-inspired outputs are derived projections of the same computations. Their differences do not constitute additional biological observations. Twenty mechanisms directly request horizons longer than the phrase; 78 use future or centred context. Downstream consumers can inherit these limitations. The outputs are offline representations; named memory and forecast layers do not establish learned memory or prospective prediction.

## S8.3 Measures and comparison scope

For route $r$, condition $c$ and anchor $k$, let $Y_{rcktd}$ be the P-layer output at frame $t$ and dimension $d$. With $D = 202$ and $T = 423$, the path measure is

$$R_{rc} = \sqrt{\frac{\sum_{k=0}^{K-1} \sum_{t,d} (Y_{rcktd} - Y_{rc0td})^2}{KTD}}.$$

Endpoint RMS uses only $k = K-1$ and denominator $TD$. Global pooling weights dimensions equally, hence mechanisms with more P dimensions receive more total weight. Per-mechanism RMS pools only its own declared P slice. The paired flute–piano contrast compares their matched tensors directly; it is not the difference between their radial distances from a reference. Signed endpoint means are retained separately.

Frames are compared without temporal realignment. Timing changes can therefore increase RMS through a shift of an otherwise similar trajectory. Original-level and common-RMS conditions remain separate. Rank correlations are descriptive; rank stability does not establish trajectory convergence. Midpoint interpolation error uses the complete saved P-layer time-course tensor (97 anchors, 423 frames and 202 dimensions) and is normalized by the corresponding full-path RMS.




**Figure S8. Reference-dependent responses and their computational audit. a,** Equal-RMS global path and endpoint displacement across 202 P dimensions. S denotes synthetic transcription and B the bird source; F and P denote the fixed sample-derived flute and piano endpoints. The same endpoints are reused across references. **b,** Per-mechanism path RMS for the post hoc union of the five largest responses on each of the five primary paths (13 names); a common raw-unit colour scale is used. Full tables retain all 84 names. **c,** Norms of total change, explicit pleasantness/inverse-roughness terms and remaining terms for SLEE and CDMR. RMS norms are not additive and the remainder is not independent of the direct terms. The displayed SDNPS equation is an implemented sensory proxy. **d,** Matched-level 49-to-97-anchor midpoint error and mechanism rank correlation. Very stable rankings coexist with non-negligible P-layer tensor interpolation errors. All panels derive from the frozen assay's archived summaries; no new engine execution is performed for this figure.

## S8.4 Findings and their scope

From the transcription, piano produces larger path RMS than flute (0.136146 versus 0.098160); from the bird, flute produces larger displacement (0.064883 versus 0.051725). Endpoint RMS has the same reversal (main Table 1). The paired direct-path flute–piano contrast has path RMS 0.060788 and endpoint RMS 0.114629. Reference-dependent ordering does not itself require a cognitive or nonlinear representation: the one-dimensional coordinates B=0, P=1, F=2 and S=3 already reverse the radial ordering. This is a mathematical counterexample, not an MI measurement.

Isolated spectral replacement gives path RMS 0.050211, timing 0.030898 and pitch 0.018750. Unequal acoustic magnitudes and renderer changes prevent interpreting these as universal perceptual weights. The factorial change is 0.09933 and its additive-model residual 0.05590 (ratio 0.563). This ratio is not explained variance or a uniquely cognitive interaction. Renderer, front-end and clipping nonlinearities can contribute. Gain-only responses of 0.01382–0.02762 are comparable to the isolated pitch response.

## S8.5 Executable formulas, labels and revision decisions

| Output / request | Executed quantity or recorded discrepancy | Interpretation and required action |
|---|---|---|
| SDNPS P2 | clamp[1 − (roughness + Sethares dissonance)/2, 0, 1]. P1 multiplies inverse inharmonicity by tristimulus balance before clipping. | Algebraic sensory proxies. A neural-pitch or brainstem interpretation requires appropriate observations; the label is not a measurement. |
| SLEE P outputs | Explicit pleasantness and inverse-roughness terms have combined coefficients 0.50, 0.55 and 0.50. | Direct acoustic sensitivity is a candidate explanation. Input contribution and learned-expectation interpretation are separate claims. |
| CDMR P outputs | Corresponding combined coefficients are 0.45, 0.45 and 0.60. | A large response does not establish a physiological mismatch signal. The remaining terms can themselves depend on the same acoustic features. |
| SLEE request | Index 23 is labelled `pitch_change`, but the feature map selects `distribution_flatness`. Horizon index 3 gives 23.2 ms; comments state 100 ms. | Preserve the frozen equation for this assay. Decide whether the intended specification or implementation should change in a separately versioned model. |
| CDMR request | Index 33 is named `x_l4l5`, while the actual coordinate is `chroma_Ab`. | Evaluate the interpretation against the input actually consumed, not the comment's generic pattern-coupling description. |
| STANM request | The source audit records index 10 (`loudness`) where some comments describe spectral flux. No independently manipulated attention cue enters the assay. | A fixed acoustic response does not test cue-dependent attention. Literal mismatches include synonyms as well as substantive discrepancies. |
| Upstream scope | The audited scope-filter path forwards full declared-upstream tensors despite comments describing dimension filtering. Fallbacks and consumed values depend on execution. | Report executable behaviour. Availability of an input does not establish its consumption or nonzero effect. |

Table S8a. Selected implemented quantities and source-audit findings relevant to prominent responses. Stable names identify the frozen outputs. The complete registry, dimensions, upstream declarations and demand audit accompany the data record. No formula was repaired within the reported experiment.

Across transcription-to-instrument paths, the direct pleasantness/inverse-roughness change has cosine alignment 0.978–0.988 with total SLEE/CDMR change; residual-to-total RMS ranges from 0.302 to 0.355. Neither quantity estimates independent causal contribution or explained variance. Full algebraic reconstruction, interventions on specified terms and comparisons on an appropriate observation remain distinct tests.

Documentation discrepancies do not all imply an erroneous algorithm. An intended definition must first be compared with both prose and executed inputs. Changing a feature index or equation requires a new engine identity and reproduction of affected results, with the frozen assay retained as a separate record. The current revision reports these findings without changing the engine.

## S8.6 Numerical resolution and additional checks

The matched-level 49-to-97 midpoint errors are 19.556% (B–S), 7.264% (S–F), 4.387% (S–P), 8.813% (B–F) and 9.358% (B–P). Mechanism-rank correlations span 0.99717–0.99998. Ranking is stable under this refinement; local curves and endpoints are not thereby converged. Finite endpoint probes do not establish mathematical discontinuity or a human perceptual threshold.

The revision-stage review independently recomputed global RMS from 202 dimension summaries and from P-dimension-weighted mechanism summaries, within $2 \times 10^{-8}$. All 1,344 mechanism curves (84 names, eight routes and two level conditions) matched their CSV path-RMS summaries. These files share the original execution; their agreement is not an independent engine replication.

Additional, explicitly post hoc checks gave equal-mechanism path RMS of 0.096265, 0.133490, 0.063289 and 0.051693 for S–F, S–P, B–F and B–P. Removing each mechanism in turn preserved the original dimension-weighted path ordering in both level conditions (168/168 checks). This rules out dependence on any single included mechanism for this ordering. It does not establish robustness to all rescalings, multiple exclusions or new sources; these checks are not additional independent experiments.

## S8.7 Execution bridge and data identity

| Property | Swan Lake / architecture record | Birdsong assay |
|---|---|---|
| Input and state | 182.597-s saved execution; selected 30–50-s trace and a targeted temporal-request intervention. | 2.45-s full phrase; fresh model construction for each anchor; no inter-anchor learning. |
| Source identity | Canonical configuration retains 854 source-pin entries; the earlier receipt reports its own verification scope. | The assay pin contains 855 entries. Of 853 overlapping entries, all hashes match; this is a source-pin comparison, not adapter equivalence. |
| Registry and ordering | 84 public names, 889 dimensions; depth order with discovery order at equal depth. Consumption is observed in the saved trace. | Same counts; first-function-wins collisions explicitly retained. Adapter uses maximum of declared and role depth with the canonical executor. |
| Preparation | Separate historical adapter and reconstruction records; supplied full-record context. Normalization must be identified for each execution. | Strict float32 decoding, mono mean, fixed mel front end and clip maximum; no silent decoder fallback. |
| Readout branches | 131 weighted mappings and eight selected recursive classes are distinguished in the architecture; source availability does not establish activation in a trace. | 131 unstretched weighted aggregates saved; the eight separate recursive classes are not run. RAM and neurochemical projections are derived outputs. |
| Dependencies | Successful reads were recorded for one execution; missing membership/get fallbacks are not exhaustively observed. | Declared upstream tensors and executor behaviour retained. A complete equality of consumed/fallback paths across these inputs is not asserted. |

Table S8b. Scope of the execution bridge. The broader 853-entry pin comparison was performed during this revision; the preceding review also matched the 80 entries shared with the figure-specific source manifest. Counts of matching sources must not be substituted for configuration identity.

The original execution receipts report nine full-repeat checks and four RAM/neuromodulator reconstructions. The review verified all 855 frozen source-file hashes, but did not rerun the engine or rehash all 1,146 stimuli. The source audit preserves the earlier decoder-recovery history: reconstructed original stimuli matched their archived hashes, and strict decoding restored exact output reproduction. These execution receipts and the revision-stage arithmetic audit are separate records.

The accompanying `birdsong-assay/` directory contains the supplied 10-page manuscript; `analysis.json`, `curves.npz`, four response CSVs and grounding summaries; source metadata; design/amendments; source audit; original adapter/analysis code; and experiment receipts. The response CSVs have 1,344 mechanism, 14,224 dimension, 416 regional and 64 neuromodulator rows. All output dimensions are retained in the dimension table; P-layer selection is recorded separately. The archive README specifies which scripts require the original full experiment tree.

The experiment manuscript SHA-256 begins `825a22f9de29fd44`; the complete hash, engine-pin hash, adapter hash and every packaged record appear in the revision manifest. `qa/execution-source-bridge.json` records the broader pin comparison. The dated review's `revision-summary-audit.json` contains the numerical checks and identifies them as summary-level analyses. S8's figure generator records its inputs and selected mechanism names.

The full frozen engine, complete stimuli and frame-level outputs are controlled-access records, available on request to the author. The experiment is an unpublished research manuscript; no new DOI or peer-review status is asserted. The earlier Zenodo archive is not presented as containing this assay. Independent reconstruction requires its matching source, adapter, inputs and environment.

## S8.8 What the next experiments must distinguish

Matched comparisons of $C^3$, $R^3$, log-mel and envelope distances on these same paths would locate the reference ordering within the pipeline. Term-level reconstruction and interventions would distinguish direct algebraic contributions from downstream reweighting. Independent phrases, registers and envelope/bandwidth controls would test stimulus generality. A small, theory-selected output set could then be compared with listener dissimilarity or discrimination under source- and participant-separated evaluation. Claims about learning or context require an identical probe under manipulated histories and an explicitly state-dependent model. These are proposed tests, not completed validation.

## References


Allen, E. J., Mesik, J., Kay, K. N. & Oxenham, A. J. Distinct Representations of Tonotopy and Pitch in Human Auditory Cortex. The Journal of Neuroscience 42, 416-434 (2022). Source

Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M. & Brattico, E. Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. NeuroImage 59, 3677-3689 (2012). Source

Auksztulewicz, R. & Friston, K. Repetition suppression and its contextual determinants in predictive coding. Cortex 80, 125-140 (2016). Source

Bangert, M. et al. Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. NeuroImage 30, 917-926 (2006). Source

Barchet, A. V., Rimmele, J. M., and Pelofi, C. (2024). TenseMusic: An automatic prediction model for musical tension. PLOS ONE, 19(1), e0296385. Source

Barrett, F. S., Grimm, K. J., Robins, R. W., Wildschut, T., Sedikides, C. & Janata, P. Music-evoked nostalgia: Affect, memory, and personality. Emotion 10, 390-403 (2010). Source

Basiński, K., Celma-Miralles, A., Quiroga-Martinez, D. R. & Vuust, P. Inharmonicity enhances brain signals of attentional capture and auditory stream segregation. Communications Biology 8, 1584 (2025). Source

Bellmann, O. T. & Asano, R. Neural correlates of musical timbre: an ALE meta-analysis of neuroimaging data. Frontiers in Neuroscience 18, 1373232 (2024). Source

Bidelman, G. M. The Role of the Auditory Brainstem in Processing Musically Relevant Pitch. Frontiers in Psychology 4 (2013). Source

Bidelman, G. M., and Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. Journal of Neuroscience, 29(42), 13165–13171. Source

Bigand, F., Bianco, R., Abalde, S. F., Nguyen, T. & Novembre, G. EEG of the Dancing Brain: Decoding Sensory, Motor, and Social Processes during Dyadic Dance. The Journal of Neuroscience 45, e2372242025 (2025). Source

Bolland, E., De Burca, A., Wang, S. H., Khalil, A. & McLoughlin, G. Efficacy of auditory gamma stimulation for cognitive decline: a systematic review of individual and group differences across cognitively impaired and healthy populations. npj Aging 12, 8 (2025). Source

Bonetti, L. et al. Spatiotemporal brain hierarchies of auditory memory recognition and predictive coding. Nature Communications 15, 4313 (2024). Source

Bravo, F., Cross, I., Stamatakis, E. A. & Rohrmeier, M. Sensory cortical response to uncertainty and low salience during recognition of affective cues in musical intervals. PLOS ONE 12, e0175991 (2017). Source

Briley, P. M., Breakey, C. & Krumbholz, K. Evidence for Pitch Chroma Mapping in Human Auditory Cortex. Cerebral Cortex 23, 2601-2610 (2013). Source

Bruzzone, S. E. P., Lumaca, M., Brattico, E., Vuust, P., Kringelbach, M. L. & Bonetti, L. Dissociated brain functional connectivity of fast versus slow frequencies underlying individual differences in fluid intelligence: a DTI and MEG study. Scientific Reports 12, 4746 (2022). Source

Burchardt, L. S., Varkevisser, J. M. & Spierings, M. J. Zebra finch tutees not only share the melody but also the rhythm of their tutor's song. Scientific Reports 15, 35573 (2025). Source

Chabin, T. et al. Cortical Patterns of Pleasurable Musical Chills Revealed by High-Density EEG. Frontiers in Neuroscience 14, 565815 (2020). Source

Cheung, V. K. M., Harrison, P. M. C., Meyer, L., Pearce, M. T., Haynes, J.-D., and Koelsch, S. (2019). Uncertainty and surprise jointly predict musical pleasure and amygdala, hippocampus, and auditory cortex activity. Current Biology, 29(23), 4084–4092.e4. Source

Cousineau, M., Bidelman, G. M., Peretz, I. & Lehmann, A. On the Relevance of Natural Stimuli for the Study of Brainstem Correlates: The Example of Consonance Perception. PLOS ONE 10, e0145439 (2015). Source

Crespo-Bojorque, P., Monte-Ordoño, J. & Toro, J. M. Early neural responses underlie advantages for consonance over dissonance. Neuropsychologia 117, 188-198 (2018). Source

Daly, I. et al. Affective brain–computer music interfacing. Journal of Neural Engineering 13, 046022 (2016). Source

Deng, X., Jiang, N., Huang, Z. & Wang, Q. Cortical Activation and Functional Connectivity Response to Different Interactive Modes of Virtual Reality (VR)-Induced Analgesia: A Prospective Functional Near-Infrared Spectroscopy (fNIRS) Study. Journal of Pain Research 18, 1095-1108 (2025). Source

Doelling, K. B. & Poeppel, D. Cortical entrainment to music and its modulation by expertise. Proceedings of the National Academy of Sciences 112 (2015). Source

Eerola, T. (2025). Prevalence of transparency and reproducibility-related research practices in music psychology (2017–2022). Musicae Scientiae, 29(2), 385–399. Source

Eerola, T., and Lahdelma, I. (2021). The anatomy of consonance/dissonance: Evaluating acoustic and cultural predictors across multiple datasets with chords. Music & Science, 4, 1–19. Source

Eerola, T., and Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. Psychology of Music, 39(1), 18–49. Source

Egermann, H., Pearce, M. T., Wiggins, G. A., and McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. Cognitive, Affective, & Behavioral Neuroscience, 13(3), 533–553. Source

Ehrlich, S. K., Agres, K. R., Guan, C. & Cheng, G. A closed-loop, music-based brain-computer interface for emotion mediation. PLOS ONE 14, e0213516 (2019). Source

Fang, R., Ye, S., Huangfu, J. & Calimag, D. P. Music therapy is a potential intervention for cognition of Alzheimer's Disease: a mini-review. Translational Neurodegeneration 6, 2 (2017). Source

Fernández-Rubio, G., Brattico, E., Kotz, S. A., Kringelbach, M. L., Vuust, P. & Bonetti, L. Magnetoencephalography recordings reveal the spatiotemporal dynamics of recognition memory for complex versus simple auditory sequences. Communications Biology 5, 1272 (2022). Source

Fishman, Y. I. et al. Consonance and Dissonance of Musical Chords: Neural Correlates in Auditory Cortex of Monkeys and Humans. Journal of Neurophysiology 86, 2761-2788 (2001). Source

Gold, B. P., Mas-Herrero, E., Zeighami, Y., Benovoy, M., Dagher, A. & Zatorre, R. J. Musical reward prediction errors engage the nucleus accumbens and motivate learning. Proceedings of the National Academy of Sciences 116, 3310-3315 (2019a). Source

Gold, B. P., Pearce, M. T., Mas-Herrero, E., Dagher, A., and Zatorre, R. J. (2019b). Predictability and Uncertainty in the Pleasure of Music: A Reward for Learning? Journal of Neuroscience, 39(47), 9397–9409. Source

Gold, B. P., Pearce, M. T., McIntosh, A. R., Chang, C., Dagher, A. & Zatorre, R. J. Auditory and reward structures reflect the pleasure of musical expectancies during naturalistic listening. Frontiers in Neuroscience 17, 1209398 (2023). Source

Grahn, J. A. & Brett, M. Rhythm and Beat Perception in Motor Areas of the Brain. Journal of Cognitive Neuroscience 19, 893-906 (2007). Source

Haiduk, F., Zatorre, R. J., Benjamin, L., Morillon, B. & Albouy, P. Spectrotemporal cues and attention jointly modulate fMRI network topology for sentence and melody perception. Scientific Reports 14, 5501 (2024). Source

Halpern, A. R., Zatorre, R. J., Bouffard, M. & Johnson, J. A. Behavioral and neural correlates of perceived and imagined musical timbre. Neuropsychologia 42, 1281-1292 (2004). Source

Harrison, E. C. et al. Neural mechanisms underlying synchronization of movement to musical cues in Parkinson disease and aging. Frontiers in Neuroscience 19, 1550802 (2025). Source

Harrison, P. M. C., and MacConnachie, J. M. C. (2024). Consonance in the carillon. Journal of the Acoustical Society of America, 156(2), 1111–1122. Source

Hausfeld, L., Disbergen, N. R., Valente, G., Zatorre, R. J. & Formisano, E. Modulating Cortical Instrument Representations During Auditory Stream Segregation and Integration With Polyphonic Music. Frontiers in Neuroscience 15, 635937 (2021). Source

Hoddinott, J. D. & Grahn, J. A. Neural representations of beat and rhythm in motor and association regions. Cerebral Cortex 34, bhae406 (2024). Source

Jacobsen, J. H., Stelzer, J., Fritz, T. H., Chételat, G., La Joie, R. & Turner, R. Why musical memory can be preserved in advanced Alzheimer's disease. Brain 138, 2438-2450 (2015). Source

Janata, P. The Neural Architecture of Music-Evoked Autobiographical Memories. Cerebral Cortex 19, 2579-2594 (2009). Source

Kheirkhah, M. et al. Mindfulness, music, visual occlusion in ketamine therapy for depression: do they change outcomes? A qualitative and quantitative analysis of a randomized controlled trial. Frontiers in Psychiatry 16, 1642025 (2025). Source

Kim, C. H., Jin, S. H., Kim, J. S., Kim, Y., Yi, S. W. & Chung, C. K. Dissociation of Connectivity for Syntactic Irregularity and Perceptual Ambiguity in Musical Chord Stimuli. Frontiers in Neuroscience 15, 693629 (2021). Source

Kim, S. G., Mueller, K., Lepsien, J., Mildner, T. & Fritz, T. H. Brain networks underlying aesthetic appreciation as modulated by interaction of the spectral and temporal organisations of music. Scientific Reports 9, 19446 (2019). Source

Kleber, B., Zeitouni, A. G., Friberg, A. & Zatorre, R. J. Experience-Dependent Modulation of Feedback Integration during Singing: Role of the Right Anterior Insula. The Journal of Neuroscience 33, 6070-6080 (2013). Source

Koelsch, S. Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. Psychophysiology 46, 179-190 (2009). Source

Koelsch, S., Fritz, T., v. Cramon, D. Y., Müller, K. & Friederici, A. D. Investigating emotion with music: An fMRI study. Human Brain Mapping 27, 239-250 (2006). Source

Koelsch, S., Gunter, T., Friederici, A. D. & Schröger, E. Brain Indices of Music Processing: "Nonmusicians" are Musical. Journal of Cognitive Neuroscience 12, 520-541 (2000). Source

Koelsch, S., Schröger, E. & Tervaniemi, M. Superior pre-attentive auditory processing in musicians. NeuroReport 10, 1309-1313 (1999). Source

Kohler, N. et al. Distinct and content-specific neural representations of self- and other-produced actions in joint piano performance. Frontiers in Human Neuroscience 19, 1543131 (2025). Source

Kokal, I., Engel, A., Kirschner, S. & Keysers, C. Synchronized Drumming Enhances Activity in the Caudate and Facilitates Prosocial Commitment - If the Rhythm Comes Easily. PLOS ONE 6, e27272 (2011). Source

Kraemer, D. J. M., Macrae, C. N., Green, A. E. & Kelley, W. M. Sound of silence activates auditory cortex. Nature 434, 158-158 (2005). Source

Lahav, A., Saltzman, E. & Schlaug, G. Action Representation of Sound: Audiomotor Recognition Network While Listening to Newly Acquired Actions. The Journal of Neuroscience 27, 308-314 (2007). Source

Lahdelma, I., Eerola, T., Ahmad, N., Clayton, M., Armitage, J., Bhattacharyya, B., and Munsamy, N. (2024). Musical expertise better predictor of tension in harmonic intervals than psychoacoustics across North and South Indian listeners. Dataset; manuscript listed as in preparation in the source record. Source

Leeuwis, N., Pistone, D., Flick, N. & van Bommel, T. A Sound Prediction: EEG-Based Neural Synchrony Predicts Online Music Streams. Frontiers in Psychology 12, 672980 (2021). Source

Leipold, S., Klein, C. & Jäncke, L. Musical Expertise Shapes Functional and Structural Brain Networks Independent of Absolute Pitch Ability. The Journal of Neuroscience 41, 2496-2511 (2021). Source

Liang, J. et al. The Brain Mechanisms of Music Stimulation, Motor Observation, and Motor Imagination in Virtual Reality Techniques: A Functional Near-Infrared Spectroscopy Study. eneuro 12, ENEURO.0557-24.2025 (2025). Source

Loui, P., Patterson, S., Sachs, M. E., Leung, Y., Zeng, T. & Przysinda, E. White Matter Correlates of Musical Anhedonia: Implications for Evolution of Music. Frontiers in Psychology 8, 1664 (2017). Source

Maess, B., Koelsch, S., Gunter, T. C. & Friederici, A. D. Musical syntax is processed in Broca's area: an MEG study. Nature Neuroscience 4, 540-545 (2001). Source

Marjieh, R., Harrison, P. M. C., Lee, H., Deligiannaki, F., and Jacoby, N. (2024). Timbral effects on consonance disentangle psychoacoustic mechanisms and suggest perceptual origins for musical scales. Nature Communications, 15, 1482. Source

Martínez-Molina, N., Mas-Herrero, E., Rodríguez-Fornells, A., Zatorre, R. J. & Marco-Pallarés, J. Neural correlates of specific musical anhedonia. Proceedings of the National Academy of Sciences 113 (2016). Source

McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. Psychological Review 102, 419-457 (1995). Source

Mencke, I., Omigie, D., Wald-Fuhrmann, M. & Brattico, E. Atonal Music: Can Uncertainty Lead to Pleasure?. Frontiers in Neuroscience 12, 979 (2019). Source

Millidge, B., Seth, A. K. & Buckley, C. L. Predictive coding: a theoretical and experimental review. arXiv:2107.12979v4 (2022). Source

Mischler, G., Li, Y. A., Bickel, S., Mehta, A. D. & Mesgarani, N. The impact of musical expertise on disentangled and contextual neural encoding of music revealed by generative music models. Nature Communications 16, 8874 (2025). Source

Mitterschiffthaler, M. T., Fu, C. H., Dalton, J. A., Andrew, C. M. & Williams, S. C. A functional MRI study of happy and sad affective states induced by classical music. Human Brain Mapping 28, 1150-1162 (2007). Source

Møller, C. et al. Audiovisual structural connectivity in musicians and non-musicians: a cortical thickness and diffusion tensor imaging study. Scientific Reports 11, 4324 (2021). Source

Noboa, M. d. L., Kertész, C. & Honbolygó, F. Neural entrainment to the beat and working memory predict sensorimotor synchronization skills. Scientific Reports 15, 10466 (2025). Source

Norman-Haignere, S. V. et al. Multiscale temporal integration organizes hierarchical computation in human auditory cortex. Nature Human Behaviour 6, 455-469 (2022). Source

Norman-Haignere, S., Kanwisher, N. & McDermott, J. H. Cortical Pitch Regions in Humans Respond Primarily to Resolved Harmonics and Are Located in Specific Tonotopic Regions of Anterior Auditory Cortex. The Journal of Neuroscience 33, 19451-19469 (2013). Source

Nozaradan, S., Peretz, I. & Mouraux, A. Selective Neuronal Entrainment to the Beat and Meter Embedded in a Musical Rhythm. The Journal of Neuroscience 32, 17572-17581 (2012). Source

Nozaradan, S., Peretz, I., Missal, M. & Mouraux, A. Tagging the Neuronal Entrainment to Beat and Meter. Journal of Neuroscience 31, 10234-10240 (2011). Source

Okada, K. i., Takeya, R. & Tanaka, M. Neural signals regulating motor synchronization in the primate deep cerebellar nuclei. Nature Communications 13, 2504 (2022). Source

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A. & Ross, B. Timbre-specific enhancement of auditory cortical representations in musicians. Neuroreport 12, 169-174 (2001). Source

Paraskevopoulos, E., Chalas, N., Anagnostopoulou, A. & Bamidis, P. D. Interaction within and between cortical networks subserving multisensory learning and its reorganization due to musical expertise. Scientific Reports 12, 7891 (2022). Source

Partanen, E., Mårtensson, G., Hugoson, P., Huotilainen, M., Fellman, V. & Ådén, U. Auditory Processing of the Brain Is Enhanced by Parental Singing for Preterm Infants. Frontiers in Neuroscience 16, 772008 (2022). Source

Patel, A. D. & Iversen, J. R. The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. Frontiers in Systems Neuroscience 8 (2014). Source

Patterson, R. D., Uppenkamp, S., Johnsrude, I. S. & Griffiths, T. D. The Processing of Temporal Pitch and Melody Information in Auditory Cortex. Neuron 36, 767-776 (2002). Source

Penagos, H., Melcher, J. R. & Oxenham, A. J. A Neural Representation of Pitch Salience in Nonprimary Human Auditory Cortex Revealed with Functional Magnetic Resonance Imaging. The Journal of Neuroscience 24, 6810-6815 (2004). Source

Porfyri, I., Paraskevopoulos, E., Anagnostopoulou, A., Styliadis, C. & Bamidis, P. D. Multisensory vs. unisensory learning: how they shape effective connectivity networks subserving unimodal and multimodal integration. Frontiers in Neuroscience 19, 1641862 (2025). Source

Potes, C., Gunduz, A., Brunner, P. & Schalk, G. Dynamics of electrocorticographic (ECoG) activity in human temporal and frontal cortical areas during music listening. NeuroImage 61, 841-848 (2012). Source

Putkinen, V., Seppälä, K., Harju, H., Hirvonen, J., Karlsson, H. K. & Nummenmaa, L. Pleasurable music activates cerebral μ-opioid receptors: a combined PET-fMRI study. European Journal of Nuclear Medicine and Molecular Imaging 52, 3540-3549 (2025). Source

Qiu, R. et al. The impact of musical intervention during fetal and infant stages on social behavior and neurodevelopment in mice. Translational Psychiatry 15, 408 (2025). Source

Rathcke, T., Smit, E., Zheng, Y. & Canzi, M. Perception of temporal structure in speech is influenced by body movement and individual beat perception ability. Attention, Perception, & Psychophysics 86, 1746-1762 (2024). Source

Ross, J. M. & Balasubramaniam, R. Time Perception for Musical Rhythms: Sensorimotor Perspectives on Entrainment, Simulation, and Prediction. Frontiers in Integrative Neuroscience 16, 916220 (2022). Source

Rupp, A., Englitz, B., Balaguer-Ballester, E. & Andermann, M. Editorial: Early neural processing of musical melodies. Frontiers in Human Neuroscience 16, 1109500 (2022). Source

Sachs, M. E., Kozak, M. S., Ochsner, K. N. & Baldassano, C. Emotions in the Brain Are Dynamic and Contextually Dependent: Using Music to Measure Affective Transitions. eneuro 12, ENEURO.0184-24.2025 (2025). Source

Sakakibara, Y. et al. A Nostalgia Brain-Music Interface for enhancing nostalgia, well-being, and memory vividness in younger and older individuals. Scientific Reports 15, 32337 (2025). Source

Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., and Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. Nature Neuroscience, 14(2), 257–262. Source

Salimpoor, V. N., van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A. & Zatorre, R. J. Interactions Between the Nucleus Accumbens and Auditory Cortices Predict Music Reward Value. Science 340, 216-219 (2013). Source

Samiee, S., Vuvan, D., Florin, E., Albouy, P., Peretz, I. & Baillet, S. Cross-Frequency Brain Network Dynamics Support Pitch Change Detection. The Journal of Neuroscience 42, 3823-3835 (2022). Source

Sansare, A., Weinrich, M., Bernard, J. A. & Lei, Y. Enhancing Balance Control in Aging Through Cerebellar Theta-Burst Stimulation. The Cerebellum 24, 161 (2025). Source

Sarasso, P. et al. Aesthetic appreciation of musical intervals enhances behavioural and neurophysiological indexes of attentional engagement and motor inhibition. Scientific Reports 9, 18550 (2019). Source

Sarasso, P., Perna, P., Barbieri, P., Neppi-Modona, M., Sacco, K. & Ronga, I. Memorisation and implicit perceptual learning are enhanced for preferred musical intervals and chords. Psychonomic Bulletin & Review 28, 1623-1637 (2021). Source

Schoeller, F., Jain, A., Haar Horowitz, A., Yan, G., Hu, X., Maes, P., and Salomon, R. (2023). ChillsDB: A gold standard for aesthetic chills stimuli. Scientific Data, 10, 307. Source

Scholkmann, F. et al. Creative music therapy in preterm infants effects cerebrovascular oxygenation and perfusion. Scientific Reports 14, 28249 (2024). Source

Schwartz, D. A., Howe, C. Q., and Purves, D. (2003). The statistical structure of human speech sounds predicts musical universals. Journal of Neuroscience, 23(18), 7160–7168. Source

Sethares, W. A. (1993). Local consonance and the relationship between timbre and scale. Journal of the Acoustical Society of America, 94(3), 1218–1228. Source

Spence, C. Crossmodal correspondences: A tutorial review. Attention, Perception, & Psychophysics 73, 971-995 (2011). Source

Squire, L. R. & Alvarez, P. Retrograde amnesia and memory consolidation: a neurobiological perspective. Current Opinion in Neurobiology 5, 169-177 (1995). Source

Tabas, A., Andermann, M., Schuberth, V., Riedel, H., Balaguer-Ballester, E. & Rupp, A. Modeling and MEG evidence of early consonance processing in auditory cortex. PLOS Computational Biology 15, e1006820 (2019). Source

Thaut, M. H., McIntosh, G. C. & Hoemberg, V. Neurobiological foundations of neurologic music therapy: rhythmic entrainment and the motor system. Frontiers in Psychology 5 (2015). Source

Thaut, M. H., Miller, R. A. & Schauer, L. M. Multiple synchronization strategies in rhythmic sensorimotor tasks: phase vs period correction. Biological Cybernetics 79, 241-250 (1998). Source

Trainor, L. J. & Unrau, A. Development of Pitch and Music Perception. Springer Handbook of Auditory Research, 223-254 (2012). Source

van der Walle, H. A., Wu, W., Margulis, E. H., and Jakubowski, K. (2025). MUSIFEAST-17: MUsic Stimuli for imagination, familiarity, emotion, and Aesthetic STudies across 17 genres. Behavior Research Methods, 57, 204. Source

Vuust, P., Brattico, E., Seppänen, M., Näätänen, R. & Tervaniemi, M. The sound of music: Differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. Neuropsychologia 50, 1432-1443 (2012). Source

Wikman, P. et al. Selective Attention Shapes Neural Representations of Complex Auditory Scenes: The Roles of Object Identity and Scene Composition. The Journal of Neuroscience 45, e0506252025 (2025). Source

Yamashita, K. et al. A pilot study on simultaneous stimulation of the primary motor cortex and supplementary motor area using gait-synchronized rhythmic brain stimulation to improve gait variability in post-stroke hemiparetic patients. Frontiers in Human Neuroscience 19, 1618758 (2025). Source

Yokota, Y. et al. Auditory stimulation at individual gamma frequency enhances cognitive performance. Scientific Reports 15, 38697 (2025). Source

Yuan, Y., Gayet, S., Wisman, D. C., van der Stigchel, S. & van der Stoep, N. Decoding Auditory Working Memory Load From EEG Alpha Oscillations. Psychophysiology 62, e70210 (2025). Source

Zamorano, A., Zatorre, R., Vuust, P., Friberg, A., Birbaumer, N. & Kleber, B. Singing training predicts increased insula connectivity with speech and respiratory sensorimotor areas at rest. Brain Research 1813, 148418 (2023). Source

Zatorre, R. J. & Halpern, A. R. Mental Concerts: Musical Imagery and Auditory Cortex. Neuron 47, 9-12 (2005). Source

Zatorre, R. J. Hemispheric asymmetries for music and speech: Spectrotemporal modulations and top-down influences. Frontiers in Neuroscience 16, 1075511 (2022). Source

Zhang, K., Zhang, H., Li, S., Yang, C., and Sun, L. (2018). The PMEmo dataset for music emotion recognition. Proceedings of the 2018 ACM International Conference on Multimedia Retrieval, 135–142. Source

R3 source and archive index; S8 assay records in the accompanying package