

Antidote: A Governed Framework for Inspectable Person-and-Moment Generative Audio Journeys

System Design and Prospective Feasibility Protocol

Alan Szmyt

Ego Hygiene

2026-09-07

Abstract

Background: Prior personalized and closed-loop music systems connect current and desired affect, semantic planning, controllable generation, and feedback, but those components do not by themselves make the full path from authorized context to later response auditable. **Objective:** We ask how a local research instrument can represent, generate, and evaluate an audio journey for one person in one moment while keeping context, intent, planned controls, realized acoustics, exposure, and response non-interchangeable. **Design:** Antidote combines a bounded comparator review with a formal record model, a proposed two-rate generate-ahead architecture, an explicit human-authority model, and a frozen prospective N-of-1 feasibility protocol. Rather than claiming novelty for the individual components, the manuscript treats end-to-end inspectability as the design contribution. **Current evidence:** The repository contains an executable synthetic session using a rule-guided planner and deterministic mock worker. No qualifying real-model package or formal human study exists, and collection authority remains false. **Conclusion:** The manuscript answers the design question by specifying an inspectable, provenance-linked chain and the gates needed to test it. Technical feasibility, perceived continuity, useful within-person advisory relationships, safety, efficacy, and neurological mechanism remain unanswered.

Keywords: adaptive music, affect, generative audio, human-computer interaction, interpretability, personalization

Source revision: 34f765346900

Contents

1 Introduction 7

1.1 Research gap 7

1.2 Staged research questions 9

1.3 Bounded contributions 9

2 Related work 10

2.1 Musical emotion, meaning, and response constructs 10

2.2 Current-to-target and longitudinal personalization 10

2.3 Closed-loop and psychophysiological music systems 11

2.4 Controllable generation and real-time architecture 11

2.5 Auditory beats: behavioral findings versus entrainment 12

2.6 Music as setting and phenomenological variable 12

2.7 Capability-level comparison 12

2.8 Bounded contribution 13

3 System design 13

3.1 Person–moment conditional model 14

3.2 Inspectable record chain 14

3.3 Semantic intent, mixins, and adapter compilation 15

3.4 Two-rate architecture and authority 16

3.5 State belief, receding horizon, and generate-ahead buffer 19

3.6 Semantic, musical, and waveform continuity 20

3.7 Exposure, response, and advisory updating 21

3.8 Provenance and failure boundaries 22

4 Methods 24

4.1 Planned measurement layers 25

4.2 Evidence classification 25

4.3 Research questions and feasibility stages 25

4.4 Units, conditions, and session structure 25

4.5 Technical verification 26

4.6 Within-person feasibility measures 27

4.7 Assignment, carryover, and missingness 28

4.8 Analysis and sensitivity plan 28

4.9	Protocol freeze, safety, and deviations	30
5	Results	31
5.1	Technical verification reporting	32
5.2	Control adherence and continuity	32
5.3	Exposure and response completeness	34
5.4	Subjective response and aftereffect	36
5.5	Null, negative, missing, and adverse observations	36
5.6	Optional physiological observations	38
5.7	Promotion and replacement rule	38
6	Discussion	38
6.1	Contribution at the design evidence level	38
6.2	Moment-specific and negotiated personalization	40
6.3	Alternative explanations and architectures	40
6.4	Bounded future research program	41
7	Limitations	42
7.1	Construct and measurement validity	43
7.2	Internal and statistical validity	43
7.3	External and ecological validity	44
7.4	System, model, and continuity failures	44
8	Ethics statement	44
8.1	Emotional safety and participant authority	46
8.2	Consent, privacy, and data governance	46
8.3	Licensing, bias, accessibility, and hearing safety	47
8.4	Independent review and collection gate	47
9	Data and code availability	48
9.1	Public artifacts and reproducibility boundary	48
9.2	Excluded and access-controlled material	48
10	Acknowledgements	49
11	Contributor statement	49
12	Conclusion	50

A	Notation and symbol glossary	54
B	Conceptual and operational equation classification	56
C	Working protocol checklist	58
D	Claim and evidence classification index	60
E	Consent and context-projection record	60
F	Provenance and reproducibility checklist	61
G	Further reading and research shelf	63
H	Migration provenance	63

List of Figures

1	Person-and-moment inspectability chain separating an approved semantic journey, measured acoustic realization, actual exposure, response, and any later advisory proposal. .	8
2	Consent-scoped transformation from private context to the bounded projection available to planning and generation.	15
3	Typed semantic mixins combine into a reviewable conditioning state under explicit human authority.	17
4	Proposed two-rate Antidote architecture. Human-correctable deliberation authorizes one immutable generation specification; variable-latency generation fills a verified future-audio buffer that decouples model work from deterministic playback deadlines. Labels distinguish the executable synthetic slice from proposed components.	18
5	Receding-horizon planning and buffer deadlines under an uncertain, person-correctable state belief.	20
6	Independent semantic-conditioning and waveform-continuity checks across adjacent generated segments.	21
7	Conceptual map of the person-moment representation, journey decision, continuity, and cautious update equations.	23
8	Provenance-linked session records pass through consent, redaction, privacy review, and explicit export approval.	24
9	Prospective feasibility-session sequence with approval, exposure, response, aftereffect, and stopping boundaries.	26

List of Tables

1	Evidence classes used by Antidote and the promotion boundary between them.	8
2	Capability-level comparison of nine reviewed systems and the proposed Antidote profile. Tokens report only what each artifact describes, not evidence strength or benefit.	13
3	Prospective measurement layers, instruments, timing, provenance, and permitted interpretation.	29
4	Governed T0 and T1 result slots with their required source records, protocol, analyses, and current unavailable states; no qualifying values are registered.	33
5	Governed H1 flow and response-completeness slots with required source records and analyses; collection is not authorized and no values are registered.	35
6	Governed slots for null, opposite-direction, mismatched, burdensome, harmful, missing, stopped, interrupted, and adverse observations; no H1 values exist.	37
7	Governed risk register pairing current controls with residual uncertainty and the gate that blocks stronger use. Mixed denotes only bounded synthetic implementation; specified controls remain unverified, and formal collection remains blocked.	45
8	Current protocol, reporting-contract, stage, source-package, evidence-class, and claim-promotion audit; no qualifying T0, T1, or H1 result package is registered.	62

1 Introduction

Personalizing audio is not only a question of which music suits a person. It is also a question of what that person wants to explore in a particular moment, how the intended transition should unfold, what audio is actually produced and heard, and how the person’s response should—or should not—inform a later proposal. Prior systems already connect parts of this problem. They select music from current and desired affect while learning person-specific associations [1]; generate music along affective trajectories inside physiological feedback loops [2, 3]; expose interpretable controls for real-time affective generation [4]; and mediate estimated state through a semantic plan before controllable generation [5]. These precedents make a broad claim to “personalized,” “closed-loop,” or “interpretable” music generation untenable.

They also expose a more specific research problem. A requested transition is not the same record as an estimated present state; an accepted semantic plan is not the audio a model realizes; a generated artifact is not evidence that it was played; and the affect a listener perceives in music is not necessarily the affect that listener feels. Emotional intensity, immediate usefulness or harm, and a later aftereffect are distinct again. If these transformations are collapsed into a prompt, a model output, or a single success score, it becomes difficult to reconstruct which consented context informed a proposal, where a mismatch arose, what the person actually encountered, or what evidence could justify changing a later proposal. We therefore treat inspectability of the complete transformation chain as the research object; we do not treat inspectability as a guarantee of correctness, meaningful consent, privacy, safety, or benefit.

Antidote originated in one researcher’s self-report of a strongly cathartic intentional generative-audio session. That lived observation motivates asking whether more deliberate person-and-moment audio journeys can be represented and studied. It is not a study result, an efficacy signal, a neurological mechanism, or evidence that the experience will generalize. Throughout this manuscript it remains a hypothesis-generating observation.

We organize the resulting design space around four linked but non-interchangeable terms. The *person* contributes explicit authority, semantic language, exclusions, and potentially relevant prior records. The *moment* is a consent-scoped and correctable representation of present context, not an objective diagnosis or inferred need. The *journey* is an accepted, time-indexed progression toward an intended experience, not proof that a state transition occurred. The *audio* is the realized artifact, whose identity and measured properties remain separate from its requested meaning and from the person’s response. Figure 1 introduces this person–moment–journey–audio chain. Its return path is a hypothesis for later within-person inquiry, not an observed adaptive mechanism.

Table 1 defines the evidence vocabulary used across the complete manuscript and the boundary each class must cross before it can support a stronger statement.

1.1 Research gap

The relevant literature does not leave an empty technical category for Antidote to fill. Longitudinal person-specific modeling, current-to-target selection, generated affective trajectories, smooth real-time feedback, interpretable musical controls, and semantic intervention planning all have documented precedent [1–5]. Furthermore, a recent preliminary brain–computer music study reported nonsignificant target-emotion and time effects while person and trial variation dominated its observed variance [6]. Technical loop closure therefore cannot be assumed to establish reliable affective change.

Instead, our bounded review identifies an operational gap: it did not find one evaluated system that preserves, as a complete governed chain, (1) the context a person permitted for a particular purpose,

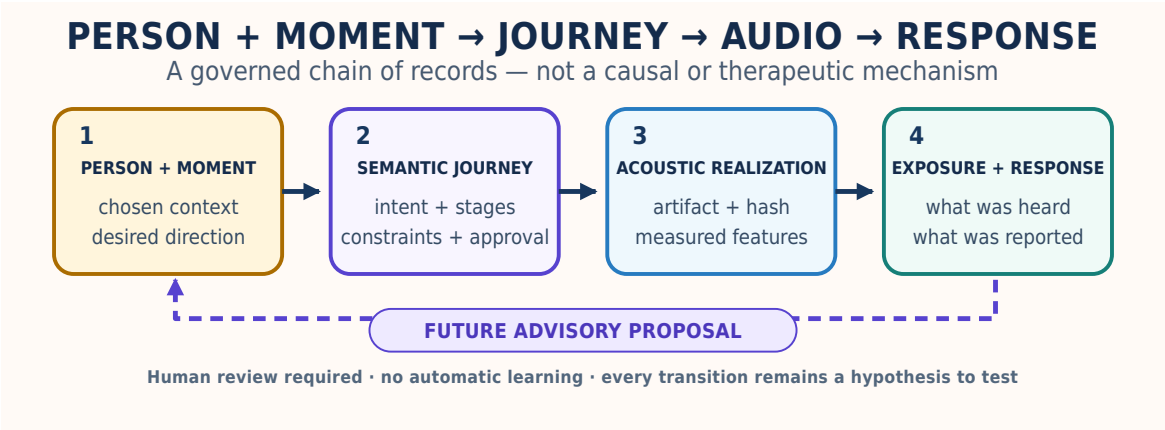


Figure 1: Person-and-moment inspectability chain separating an approved semantic journey, measured acoustic realization, actual exposure, response, and any later advisory proposal.

Figure description: Four numbered cards run left to right. Person and moment holds chosen context and desired direction; semantic journey holds intent, stages, constraints, and approval; acoustic realization holds the content-hashed artifact and measured features; exposure and response distinguishes what was heard from what was reported. A separate dashed return path is labeled future advisory proposal, human review required, and no automatic learning. Every forward transition is a proposed relationship to test rather than a causal or therapeutic mechanism.

Class	Meaning	Acceptable support	Promotion boundary
Source	What an external primary or authoritative source actually establishes	Primary or authoritative work plus a reviewed source record.	Supports only what the source establishes; never promotes itself to an Antidote claim.
Hypothesis	What Antidote proposes could be true	Scoped rationale, uncertainty, and a testable research question.	Remains tentative until qualifying observations and prespecified analysis exist.
Observation	What occurred in a recorded session or experiment	A qualifying artifact, protocol, measurement, or person report.	Does not become an interpretation, outcome, or generalized claim automatically.
Interpretation	What a person or analysis suggests the observation may mean	Explicit reasoning linked to evidence and alternatives.	Must retain uncertainty, conflict, and plausible alternatives.
Claim	What the manuscript or product states as supported	Ledger ID, named evidence, permitted wording, and review disposition.	Must stay within its ledger boundary; otherwise downgrade, revise, or reject.

Each promotion is a separately reviewed decision. A generated explanation is not a source, and a synthetic fixture or compelling experience is not automatically a human-outcome result.

Table 1: Evidence classes used by Antidote and the promotion boundary between them.

Table description: Rows define what each evidence class means, what artifact may support it, and which promotion step requires review. The table emphasizes that generated explanations are not sources, observations are not automatically claims, and no layer silently promotes itself to the next.

(2) user-extensible semantic intent and exclusions, (3) an accepted time-indexed journey plan, (4) the generated artifact and measured acoustic realization, (5) actual exposure, (6) perceived expression, felt response, immediate usefulness or harm, and later aftereffect as separate observations, and (7) any proposed within-person update with its provenance and explicit human approval. “Did not find” is a bounded-review result, not proof that no such system exists. The proposed combination may constitute a useful research-system contribution; its global novelty, technical feasibility, learnability, and benefit remain unresolved.

This framing shifts the central question away from whether a system can infer an emotion and generate music. It asks whether a research instrument can make the entire person-and-moment audio hypothesis inspectable enough to specify, challenge, reproduce, and evaluate without promoting one record type into another.

1.2 Staged research questions

The manuscript separates a design question from two later feasibility questions:

RQ1—design. How can a local research instrument represent, generate, and evaluate a person-and-moment audio journey while keeping consented context, semantic intent, planned controls, realized acoustics, exposure, response, and longitudinal updates inspectable and non-interchangeable?

RQ2—technical feasibility. Can the implemented instrument preserve its declared authority, provenance, control-adherence, continuity, interruption, and recovery contracts across a complete local session?

RQ3—within-person feasibility. Across repeated sessions, do provenance-linked person-and-moment records support useful advisory relationships among semantic intent, acoustic realization, and self-reported response without increasing burden, surprise, mismatch, or harm?

The design/protocol manuscript is structured to address RQ1 by formalizing the research instrument and its evaluation boundary. RQ2 requires an end-to-end implemented session and auditable technical checks. RQ3 additionally requires a frozen protocol, resolved ethics and privacy gates, qualifying repeated-session observations, and uncertainty-aware analysis. Neither later question is answered by the originating experience, the literature review, a repository fixture, or the proposed architecture.

1.3 Bounded contributions

Within those limits, the complete design/protocol manuscript is scoped to five contributions at explicitly different evidence states:

1. it *bounds* the problem against prior closed-loop, current-to-target, controllable, physiologically adaptive, and longitudinal music systems rather than presenting those ingredients as new;
2. it *formalizes* an end-to-end inspectability contract that keeps consented context, semantic intent, planning, acoustic realization, exposure, response layers, provenance, and advisory updating distinct;
3. it *proposes* a two-rate adaptive-audio architecture that separates uncertain, person-correctable deliberation from deterministic generate-ahead rendering, without claiming a validated controller or seamless listening experience;
4. it *requires* a local-first authority model in which context use, plan acceptance, generation, playback, retention, export, and personal-model updates remain interruptible and human-correctable, while distinguishing the partially implemented mock path from proposed runtime behavior; and

5. it *specifies prospectively* an evidence-disciplined feasibility boundary that retains null, negative, missing, interrupted, and adverse observations and never substitutes intended control, acoustic adherence, perceived expression, felt response, usefulness or harm, or aftereffect for one another.

The remainder of the manuscript locates these contributions in prior work, formalizes the system and its mathematical objects, specifies the prospective feasibility method and results boundary, discusses only evidence-proportional interpretations, states limitations and ethical constraints, and records availability and reproducibility conditions before concluding with the gates required to revisit RQ2 and RQ3.

2 Related work

The living source atlas is intentionally broader than both the bibliography shelf and the sources cited here. This section uses only promoted records and organizes them by the question for which each source is informative. A study may establish that an architecture existed while providing weak evidence for a listener outcome; a theoretical framework may clarify measurement without validating an implementation; and a normative or engineering precedent may support a design transfer without becoming empirical affective evidence.

2.1 Musical emotion, meaning, and response constructs

Musical response is not a direct readout of an acoustic feature. Juslin and Västfjäll describe multiple candidate induction mechanisms—including brain-stem reflexes, evaluative conditioning, emotional contagion, imagery, episodic memory, and musical expectancy in addition to appraisal—whose timing, volitional influence, cultural dependence, and relation to musical structure differ [7]. The framework is theoretical, but it makes a single acoustic-to-emotion rule an inadequate research model. It also motivates retaining a person’s context and interpretation rather than treating a classifier output as the mechanism that operated.

Measurement work reinforces that distinction. Across four studies, Zentner et al. derived and replicated the nine-factor Geneva Emotional Music Scale and reported differences by genre and by whether emotion was perceived in the music or felt by the listener [8]. This supports a music-specific response vocabulary while also showing why no group-derived taxonomy should replace personally meaningful language. AffectMachine- Classical offers a complementary technical boundary: its 26-participant study evaluated whether generated excerpts communicated intended affect, not whether listeners felt that affect or benefited from it [4]. Expression alignment, felt response, usefulness, and harm are therefore separate outcome questions.

An exclusively positive account of musical intensity would be incomplete. Silverman et al.’s theoretical model treats music-induced harm as potentially affective, behavioral, cognitive, identity-related, interpersonal, physical, or spiritual, arising through interactions among context, deliverer, music, and recipient [9]. The model does not estimate risk or validate Antidote’s safeguards. It does require the research record to retain mismatch, unwelcome intensity, and harm for the person’s interpretation rather than recoding them as failed optimization or successful catharsis.

2.2 Current-to-target and longitudinal personalization

Person- and moment-conditioned music predates Antidote. The personalized affective music player combined current and desired affect, physiological observations, and person-specific associations to select existing music over multiple weeks [1]. Its three-participant field evaluation is too small to establish robust generalization, but it is direct precedent for current-to-goal selection and longitudinal personal

modeling.

The LUCID trial extended this design neighborhood to an explicit calm target, the iso principle, preference information, and reinforcement-learning-based sequencing of pre-existing tracks and nature sounds [10]. The open-label single-session study randomized 163 participants across music-plus-beats, music, beats, and pink-noise conditions; some differences appeared in the moderate trait-anxiety subgroup and were less differentiated in the high-anxiety subgroup. Commercial involvement, subgroup dependence, and the inseparability of several design components prevent the study from validating a universal journey or mechanism. BEAMERS likewise documents strong person-dependent variation in response to the same song while using EEG in a personalized recommendation setting [11]; as a preprint, it is a comparator rather than an independent basis for an outcome claim.

Together, these systems reject novelty claims based on personalization, current-to-target sequencing, physiology, or longitudinal learning alone. They also leave a methodological reason to model the unit of inquiry as a person–moment–journey record rather than a fixed “best music” profile.

2.3 Closed-loop and psychophysiological music systems

Physiological feedback has been connected to music at several levels of complexity. iHeartLift used heart-rate-related feedback to manipulate music tempo, demonstrating a narrow real-time loop rather than a semantic generator [12]. Daly et al.’s affective brain–computer music interface combined EEG and other physiology, a desired affective trajectory, algorithmic generation, and a person-specific model across sessions [2]. Eight participants completed its longitudinal protocol; attrition, constrained generation, and signal limitations qualify the result. Ehrlich et al. reported continuous generated musical feedback with smooth parameter control and user calibration, using an 11-listener perceptual study and a five-participant BCI pilot [3]. The preliminary outcomes were heterogeneous and did not establish reliable regulation.

MindMelody is the closest recent semantic comparator. It connects real-time EEG processing to global and local affect trajectories, a retrieval-augmented semantic intervention plan, controllable MusicGen-based generation, and physiological and subjective feedback [5]. Its reviewed v2 preprint reports technical evaluations and a randomized within-subject non-clinical pilot, but the pilot participant count was not found in the visible methods. The system is strong prior art for an affect estimate to semantic plan to generated music to response loop; its pilot is not Antidote evidence and does not settle clinical efficacy.

Conflicting findings must remain in the same comparison. A 22-participant preliminary minimalist brain–computer musical interface reported nonsignificant target-emotion and time effects, with participant and trial variation dominating the observed variance [6]. At field level, a systematic review found heterogeneous music-neurofeedback methods and no consensus success measure [13]. A separate scoping review of commercial music-based digital therapeutics found product presence but insufficient direct evidence for real-world effectiveness or superiority over generic music [14]. Technical loop closure, product availability, and beneficial response are therefore three different evidence classes.

2.4 Controllable generation and real-time architecture

Interpretable generation controls also have established precedent. AffectMachine-Classical maps continuous valence and arousal controls to a rule-based and probabilistic generator that can change musical parameters bar by bar [4]; MindMelody exposes semantic plan fields including tempo, texture density, section dynamics, and a target affect trajectory [5]. Mustango adds explicit chord, beat, tempo, and key conditioning to text-to-music generation and evaluates condition adherence [15]. JASCO demonstrates global text conditioning combined with local time-varying symbolic and audio conditions [16]. These

systems support a translation from human-readable intent into inspectable generator controls. They do not make prompt mixtures perceptually linear, guarantee requested conditions in the realized artifact, or establish felt-response control.

Antidote’s timing and adaptation proposal therefore transfers engineering ideas rather than outcome evidence. Adaptive game-music research supplies precedents for transition-aware musical structure [17]; just-in-time adaptive intervention design separates decision points, tailoring variables, and intervention options [18]; model-predictive control supplies receding-horizon re-estimation [19]; and partially observable planning supplies belief-state reasoning under uncertainty [20]. Interactive machine learning further motivates visible correction and override instead of treating an estimate as authority [21]. These transfers inform a two-rate, generate-ahead architecture. They do not model a person as a controllable plant, validate an affect transition model, define a therapeutic reward, or establish a stable real-time generator.

2.5 Auditory beats: behavioral findings versus entrainment

Auditory-beat outcomes and brainwave entrainment require separate claims. The LUCID factorial comparison permits reporting subgroup-dependent behavioral differences among music, beat, combined, and noise conditions [10]; it does not isolate a universal beat-frequency mechanism. In contrast, Ingendoh et al.’s systematic review asked specifically whether binaural-beat stimulation entrains EEG activity. Among 14 included studies, five supported the entrainment hypothesis, eight contradicted it, and one was mixed, with substantial heterogeneity in stimulation, design, EEG measures, and analysis [22]. Binaural beats can thus be treated as testable psychoacoustic controls, but the reviewed evidence does not justify a rule that a frequency produces a known neural or therapeutic state.

2.6 Music as setting and phenomenological variable

Psychedelic research provides adjacent evidence that music can shape an experience without being reducible to a treatment ingredient. In interviews with 19 patients receiving psilocybin with psychological support for treatment-resistant depression, Kaelen et al. identified welcome influences such as emotion, imagery, guidance, openness, calm, and safety, alongside unwelcome emotion, resistance, and misguidance [23]. Music-experience ratings were associated with acute experience and later symptom change, but the small exploratory study had no no-music comparison and cannot isolate music, playlist, preparation, support, drug, or expectancy effects.

Rowe et al.’s scoping review included 19 papers from 10 studies and 330 total participants [24]. It found preliminary relationships with intensified emotion, meaning and imagery processes, and neural-dynamics measures, while emphasizing small samples, methodological inconsistency, and poorly standardized music conditions. The eligible evidence concerned LSD, psilocybin, and ayahuasca rather than establishing a ketamine–music comparison. This literature supports treating music as an active, temporally structured, sometimes unwelcome setting variable. It does not validate generated therapeutic music, an optimized soundtrack, or Antidote as a psychedelic or ketamine intervention.

2.7 Capability-level comparison

Table 2 projects the governed comparator matrix into capability families. A reported mark means only that the reviewed artifact described the capability; it is not an evidence grade or a positive outcome. Partial and unreported cells remain explicit, and the Antidote row is labeled as a proposed reference profile rather than an evaluated system.

System	Sem.	Journey	Audio	Loop	Longit.	Prov.
MindMelody [5]	R	R	R	R	NR	NR
Daly aBCMI [2]	P	R	R	R	R	P
Ehrlich BCI [3]	P	R	R	R	P	P
AffectMachine [4]	R	R	R	Prop.	NR	NR
Affective player [1]	P	P	R	P	R	NR
BEAMERS [11]	P	P	R	NR	P	NR
LUCID [10]	P	R	R	NR	NR	P
iHeartLift [12]	NR	P	R	R	NR	NR
Minimalist BCMI [6]	P	P	R	R	NR	P
Antidote profile	Prop.	Prop.	Prop.	Prop.	Prop.	Prop.

R = reported; P = partial; NR = not reported in the reviewed artifact; NA = not applicable; Prop. = proposed, not evaluated. A token identifies reported capability only, not evidence strength, implementation quality, safety, or benefit.

Table 2: Capability-level comparison of nine reviewed systems and the proposed Antidote profile. Tokens report only what each artifact describes, not evidence strength or benefit.

Table description: MindMelody, Daly aBCMI, Ehrlich BCI, AffectMachine, the personalized affective music player, BEAMERS, LUCID, iHeartLift, and the Minimalist BCMI are projected from the governed comparator JSON. Columns cover semantic representation, journey representation, audio strategy, within-session adaptation, longitudinal learning, and provenance. Reported audio strategy may be generation, sonification, selection, sequencing, or playback as defined in the source cell. The final Antidote row says proposed in every capability family and is not evidence of implementation, novelty, safety, or benefit.

2.8 Bounded contribution

Antidote therefore does not claim those components individually as new. It formalizes a governed combination that preserves consented context, user-extensible semantic intent, time-indexed planning, measured acoustic realization, actual exposure, perceived and felt response, usefulness or harm, aftereffect, provenance, and human-approved within-person updating as distinct records. The bounded review did not identify that complete chain in one evaluated system; this is a contribution rationale, not a global novelty claim or proof of learnability or benefit.

The gap handed to System Design is consequently narrower than “adaptive music.” Prior work establishes the ingredients and several close loops; the remaining task is to define how one research instrument preserves authority and evidence status at each transformation, including failures and disagreements. The following section formalizes that record chain without treating design coherence as evidence of feasibility or benefit.

3 System design

Antidote is designed as an inspectable research instrument, not as an opaque audio recommender or an autonomous affect controller. Its unit of inquiry is a person encountering a particular acoustic realization in a particular moment after approving the context and journey that informed it. The architecture therefore preserves each transformation—source record, working projection, semantic intent, journey plan, model conditioning, generated artifact, measured acoustics, exposure, response, and any later update—as a separately identified object. This section addresses RQ1 by defining that chain and the authority, timing, uncertainty, and failure boundaries around it.

The equations below are a design language. Each is visibly classified as *definitional*, *conceptual*, a *proposed operationalization*, an *estimated model*, or a *future empirical model*. A definition names records without asserting a causal law. A conceptual equation transfers a useful structure from another field. A proposed operationalization states a quantity that could be implemented and tested. An estimated model produces an uncertain system view rather than ground truth. A future empirical model is neither

fitted nor used by the present prototype. Appendix A centralizes the notation, and Appendix B records the status and prohibited inference of every equation.

3.1 Person–moment conditional model

Let i index an individual and t a decision point within or between sessions. Antidote does not represent the person as a fixed feature vector. Instead, it defines the information available for one bounded moment as the heterogeneous record

$$z_{i,t} = \left(o_{i,t}^{\text{usr}}, o_{i,t}^{\text{opt}}, c_{i,t}^{\text{work}}, g_{i,t}, k_{i,t}^{\text{usr}}, \mathcal{H}_{i,<t} \right). \quad (1)$$

ANT-EQ-001—Definitional; implementation: represented in the executable mock slice.

Here o^{usr} contains explicit check-in observations, o^{opt} contains any separately consented optional observations, c^{work} is the reviewed context projection, g is the desired journey, k^{usr} contains direct semantic-control settings, and $\mathcal{H}$ denotes eligible prior evidence. Equation 1 is a tuple, not a sum: its members need not share units and none is silently converted into “the person’s state.” Missing optional observations remain missing. The same individual may therefore produce different moment records at two times, and the same audio may be associated with different responses without either response being treated as anomalous.

The desired journey $g_{i,t}$ is a direction and time-indexed proposal—for example, staying with, softening, exploring, releasing, focusing, or closing—rather than a promised terminal state. The current schemas permit compact valence, arousal, intensity, and confidence fields, but preserve free-text description because a low-dimensional affect coordinate cannot exhaust the person’s meaning. Likewise, prior history is eligible evidence only when its retention and learning permissions remain active and the person accepts its relevance to the present moment.

3.2 Inspectable record chain

The context boundary separates retained source records from replaceable interpretations. Following the repository’s personal-model contract, the context state is defined as

$$S_{i,t}^{\text{ctx}} = (L_{i,t}, P_{i,t}, V_{i,t}), \quad (2)$$

ANT-EQ-002—Definitional; implementation: append-only events and derived projections exist for the executable mock slice.

where L is the consented source-event record, P contains derived projections, and V contains their versions and provenance. This partition is compatible with local-first and event-derived views, but neither local storage nor an append-only history proves privacy, truth, or correct interpretation [25, 26]. A current consent grant $A_{i,t}$ then constrains the view used for a declared purpose:

$$c_{i,t}^{\text{work}} = \pi \left(S_{i,t}^{\text{ctx}}; A_{i,t} \right). \quad (3)$$

ANT-EQ-003—Proposed operationalization; implementation: manual session-scoped projection is enforced in the executable mock slice.

The projection operator π may select, minimize, summarize, or omit records, but it must preserve source-event identifiers, derivation version, purpose, review state, expiry, and any person-authored edit or exclusion. Revoked, expired, excluded, or out-of-purpose fields do not enter the working view. The source events remain distinguishable from the projection so a new derivation can be audited without rewriting what was originally authorized. Consent to inspect, project, generate, retain, learn, export, and publish remains separate; manual semantic entry remains usable without historical context.

The record chain continues from the accepted working projection to an immutable moment record, a versioned journey-plan revision, an approved generation specification, a generation attempt, a content-hashed artifact and feature report, an actual exposure, and separated response observations. No skipped link is inferred from a later one: generating an artifact does not establish playback, and playback does not establish a response.

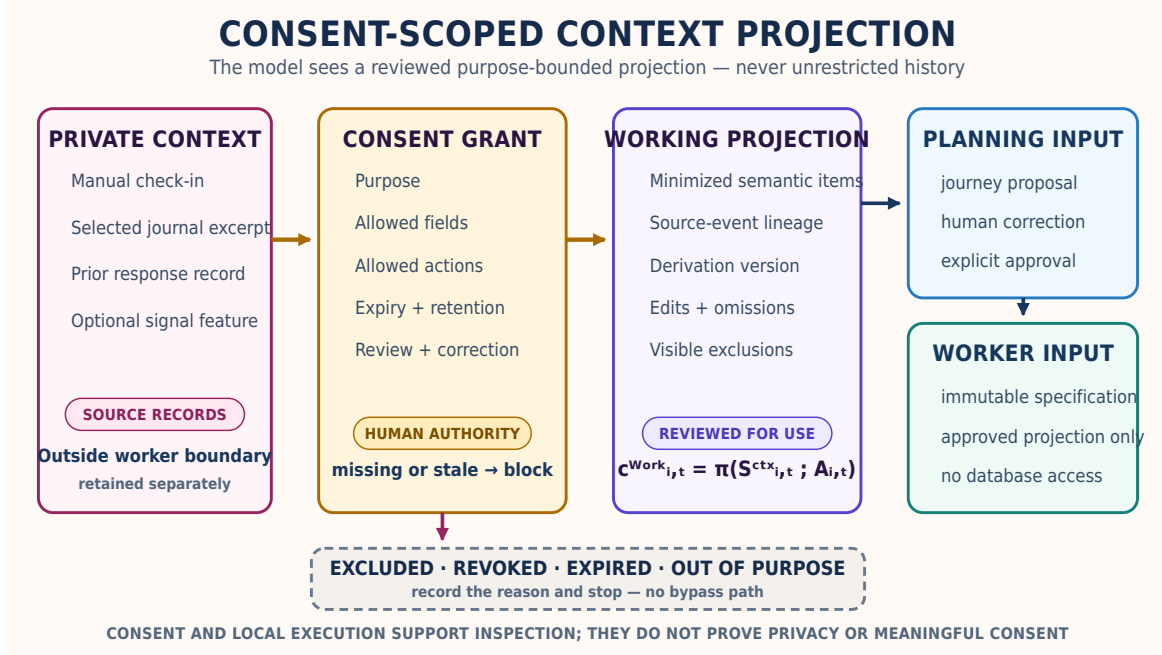


Figure 2: Consent-scope transformation from private context to the bounded projection available to planning and generation.

Figure description: A private context store sits outside the generation boundary. A consent grant selects purpose, fields, retention, and expiration, producing a reviewable working projection. Only that minimized projection can inform planning and the immutable generation specification. Excluded, revoked, or expired fields terminate at a blocked path rather than reaching the worker.

3.3 Semantic intent, mixins, and adapter compilation

Semantic intent is the model-independent statement of what the person wants a journey to explore. It combines a desired transition with words, metaphors, memories, inclusions, exclusions, priorities, and bounded acoustic preferences. A semantic mixin m_j is a typed and versioned control object with a human-readable purpose, neutral position, bounds, conflicts, and adapter mapping. It is not a hidden prompt fragment and does not claim a linear mapping to acoustics or felt emotion.

For mixin j , the effective normalized control is determined by an explicit authority order:

$$k_{j,t}^{\text{eff}} = \begin{cases} k_{j,t}^{\text{usr}}, & \text{when the person sets a current value,} \\ k_{j,t}^{\text{acc}}, & \text{when the person accepts a suggestion,} \\ k_j^0, & \text{otherwise,} \end{cases} \quad k_{j,t}^{\text{eff}} \in [0, 1]. \quad (4)$$

ANT-EQ-004—Proposed operationalization; implementation: the generalized mixin control surface is not implemented.

An inferred value has no branch in Equation 4: it gains authority only through explicit acceptance. A current person-authored value therefore outranks an accepted suggestion, which outranks the neutral default. An active exclusion or safety constraint may reject a value or block a transition, but it may

not silently substitute a different intent. Stop, revocation, and exclusion act immediately and are never smoothed.

For ordinary non-safety controls, a visible rate parameter may reduce jitter:

$$\alpha_{j,t} = (1 - \rho_j)\alpha_{j,t-1} + \rho_j k_{j,t}^{\text{eff}}, \quad 0 < \rho_j \leq 1. \quad (5)$$

ANT-EQ-005—Proposed operationalization; implementation: control smoothing is not implemented or perceptually validated.

Both k and α are dimensionless. The rate ρ_j , neutral zone, and maximum rate of change are mix-in-specific, visible policy rather than model invention. Smoothing can reduce control jitter; it cannot prove a smooth musical transition or a helpful experience.

The resulting semantic condition for journey stage n is

$$q_{i,n} = \text{Resolve} \left(q_{i,n}^{\text{base}} \oplus \bigoplus_{j=1}^J \alpha_{j,n} m_j; \Xi_{i,t}, \prec_{i,t} \right), \quad (6)$$

ANT-EQ-006—Definitional; implementation: the typed generalized composition remains formalized only.

where $\oplus$ denotes typed composition, Ξ contains active exclusions and constraints, and $\prec$ records the visible precedence order used to resolve conflicts. The operator cannot be interpreted as numerical addition or raw text concatenation. Human-scale controls may be grounded in established conditioning dimensions such as tempo, key, chord, and beat, while remaining separate from the model’s native prompt representation [15, 16].

The accepted journey plan $J_{i,t}$ remains model independent. A replaceable adapter a compiles one stage against the adapter capability set $\mathcal{K}_a$:

$$\begin{aligned} \gamma_{i,n}^{(a)} &= \text{Compile}_a(J_{i,t}, q_{i,n}, \mathcal{K}_a), \\ \varphi_{i,n}^{(a)} &= \text{Prompt}_a(\gamma_{i,n}^{(a)}). \end{aligned} \quad (7)$$

ANT-EQ-007—Proposed operationalization; implementation: only the immutable mock generation boundary is executable.

The conditioning state γ records model-native controls, conflicts, downgrades, exclusions, uncertainty, and provenance. The prompt projection φ is merely one adapter-specific projection of γ ; it is not the journey plan or the canonical semantic record. An unsupported control must produce a visible downgrade or failure before generation. The person may accept the revised specification, change the plan, choose another compatible adapter, or stop. This separation permits model comparison without rewriting Antidote’s domain language and recognizes that requested conditions are not guaranteed to be realized by current generators [15, 16].

3.4 Two-rate architecture and authority

The proposed architecture uses two clocks because model deliberation and audio rendering have different timing and authority requirements. A slower *deliberative loop* operates at meaningful decision points. It may inspect authorized observations, maintain alternative state hypotheses, combine person-authored intent, propose a short journey horizon, and request future generation. Every material change remains visible and correctable. A faster *deterministic audio loop* schedules already verified material, applies bounded gain and transition curves, tracks playback, and responds immediately to stop or underrun conditions. It does not infer psychology or rewrite semantic intent. Control-rate

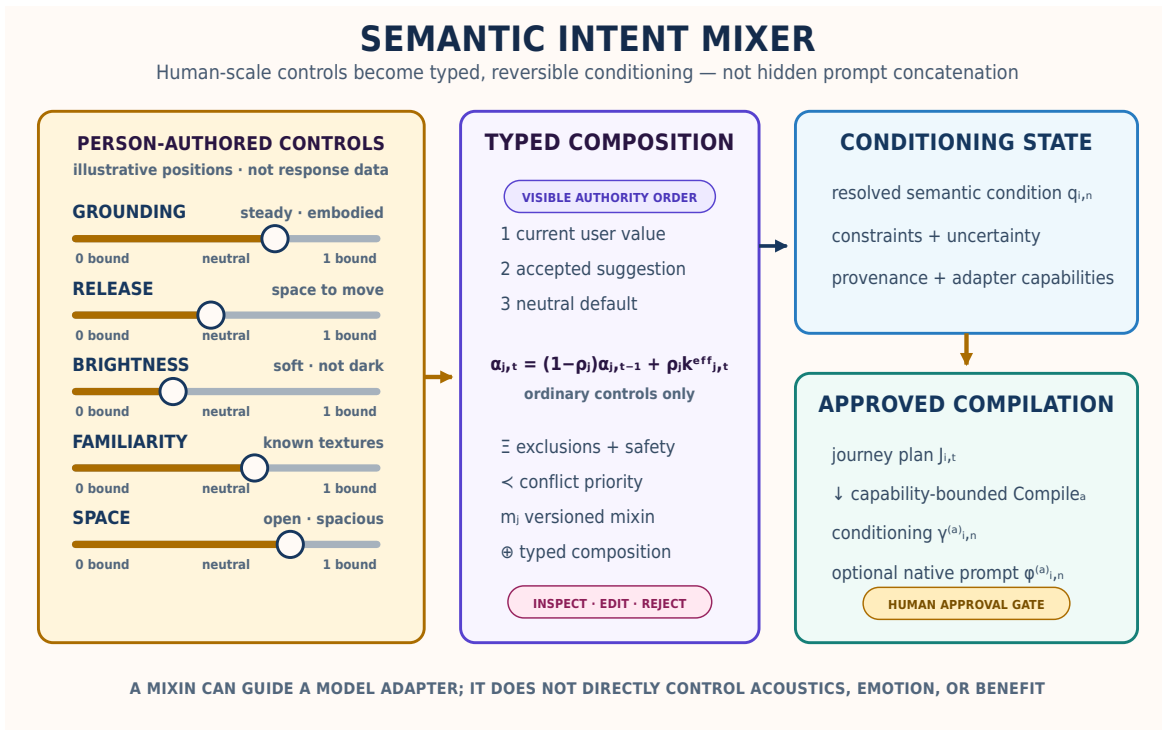


Figure 3: Typed semantic mixins combine into a reviewable conditioning state under explicit human authority.

Figure description: Several semantic mixin cards expose a human-readable purpose, current value, neutral position, bounds, exclusions, priority, and adapter mapping. The cards combine into a versioned conditioning state that the person may inspect, correct, accept, or reject before generation. No arrow claims that a mixin directly causes an emotion.

and render-rate separation has normative and neighboring-system precedents, but does not establish Antidote’s conformance or response quality [3, 17, 27].

Authority crosses these loops asymmetrically. The person chooses context, desired journey, mixins, plan acceptance, generation, playback, retention, export, and whether evidence may inform a later proposal. The Rust domain core enforces those transitions and owns the canonical session record. A model worker receives only an approved generation specification and cannot read the database, change the personal model, initiate playback, or publish an interpretation. The renderer may halt or fall back when a deadline or safety condition fails, but cannot silently choose a new emotional goal.

At the source revision identified in this manuscript’s build provenance, an executable synthetic slice already composes a React/Tauri interface, Rust session authority, session-scoped context review, deterministic Level-1 journey planning, SQLite events, a supervised mock worker, verified synthetic WAV production, deliberate playback state, response capture, and safety halts. This is repository evidence, not human-outcome evidence. The generalized Semantic Intent Mixer, probabilistic state estimation, receding-horizon controller, generate-ahead queue, continuity engine, real music model, measured control adherence, privacy-reviewed export, optional sensors, and autonomous adaptation are not implemented. The governed final-figure specification requires both groups to be labeled so the proposed system cannot borrow credibility from the mock path.

“Closed loop” therefore names a governed possibility for later evidence to inform a new proposal. In v0, planning is rule guided, responses are recorded, and a personal-model change requires a separate evidence-linked proposal and human acceptance. The worker and controller have no autonomous learning authority.

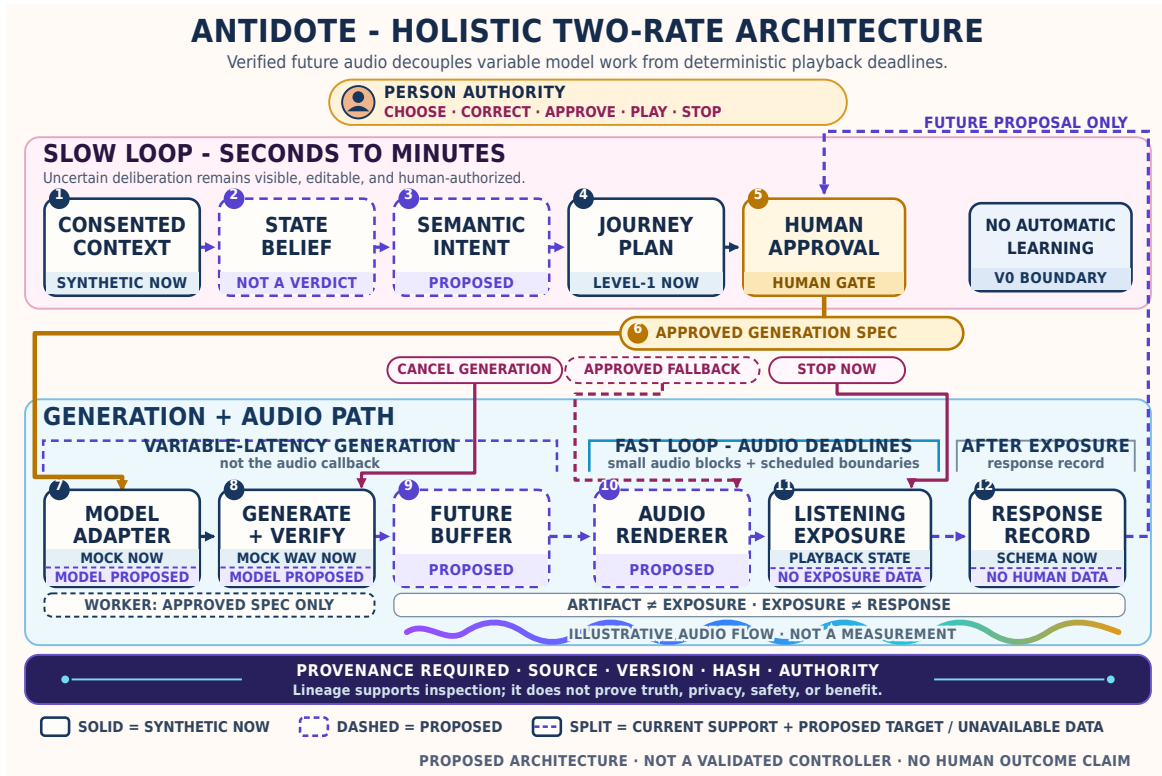


Figure 4: Proposed two-rate Antidote architecture. Human-correctable deliberation authorizes one immutable generation specification; variable-latency generation fills a verified future-audio buffer that decouples model work from deterministic playback deadlines. Labels distinguish the executable synthetic slice from proposed components.

Figure description: The numbered path begins with consented context, a correctable state belief, semantic intent, a journey plan, and explicit human approval. Approval produces one immutable generation specification that crosses into a capability-scoped model adapter. Variable-latency generation and verification prepare future audio before the proposed deterministic renderer consumes small audio blocks and segment boundaries for a deliberate playback path. Exposure remains a separate record of what was actually heard, and a response remains a separate observation that can return only as a dashed future proposal requiring human review; no automatic learning path exists. Solid cards identify the executable synthetic slice, dashed cards identify proposed capabilities, and split cards distinguish current mock or schema support from target behavior. The listening/exposure and response cards are split because playback state and response schemas exist while no qualifying human exposure or response is observed. Separate cancel, approved-fallback, and immediate-stop paths bypass semantic deliberation. The provenance spine requires source, version, hash, and authority lineage while explicitly making no claim of truth, privacy, safety, benefit, hard-real-time generation, or human outcome.

3.5 State belief, receding horizon, and generate-ahead buffer

Real-time evaluation begins with observations rather than a verdict about what the person needs. Let $o_{i,1:t}$ contain only provenance-linked explicit reports, control changes, playback actions, and optional signals authorized for the current purpose. A candidate state estimator may retain alternatives as a belief distribution

$$b_{i,t}(x) = \Pr(X_{i,t} = x \mid o_{i,1:t}, u_{i,1:t-1}). \quad (8)$$

ANT-EQ-008—Estimated model; implementation: no state estimator or sensor path exists in v0.

Equation 8 transfers belief-state notation from planning under partial observation; Antidote does not claim to implement a POMDP or to observe a hidden true emotion [20]. The state space, observation model, calibration, and confidence representation remain study questions. Person correction changes the system record and outranks the model view. Conflicting, missing, stale, or out-of-distribution observations should widen uncertainty or suspend a proposal rather than create false precision. Reviews of passive monitoring and electrodermal affect models reinforce this validation burden: missingness, feature variation, small samples, limited generalizability, and weaker valence inference constrain any future sensor use [28, 29].

At a decision point, a conceptual receding-horizon planner may compare a sequence $U = (u_t, \dots, u_{t+H_p})$ of semantic and musical control proposals:

$$\begin{aligned} \mathcal{U}_{i,t}^{\text{adm}} &= \mathcal{U}(A_{i,t}, \Xi_{i,t}, \mathcal{K}_a), \\ U_{i,t:t+H_p}^* &\in \arg \min_{U \in \mathcal{U}_{i,t}^{\text{adm}}} \sum_{\tau=t}^{t+H_p} \left[\bar{L}_g(\hat{x}_{i,\tau|t}, g_{i,\tau}, u_{i,\tau}) + \lambda_\Delta \bar{D}_\Delta(u_{i,\tau}, u_{i,\tau-1}) + \lambda_R \bar{R}(u_{i,\tau}) + \lambda_B \bar{C}_{\text{burden}}(u_{i,\tau}) \right]. \end{aligned} \quad (9)$$

ANT-EQ-009—Conceptual; implementation: v0 uses a deterministic rule-guided planner rather than this optimizer.

The admissible set is constrained by active consent A , exclusions and safety rules Ξ , and adapter capabilities $\mathcal{K}_a$. The four barred terms are normalized, dimensionless design costs: represented journey mismatch, control change, uncertainty or risk, and interaction burden. Their weights are therefore dimensionless policy choices, not learned clinical utilities. $\hat{x}_{i,\tau|t}$ is a planning estimate, not a predicted lived state. Only the first proposed action is shown; nothing is committed until the person approves it, after which the system observes and may replan. This is an MPC-inspired transfer under JITAI-like decision points, not a stability, optimality, intervention, or affect-dynamics claim [18, 19].

Generation must also finish before its audio is needed. If $B_{i,t}$ is the verified playable buffer in seconds, the proposed readiness condition is

$$B_{i,t} \geq \ell_{\text{observe}} + \ell_{\text{plan}} + \ell_{\text{generate}} + \ell_{\text{verify}} + \ell_{\text{schedule}} + M_{i,t}. \quad (10)$$

ANT-EQ-010—Proposed operationalization; implementation: generation latency and a generate-ahead buffer are not yet benchmarked.

Every term in Equation 10 is measured in seconds; $M_{i,t}$ is an uncertainty margin derived from the declared hardware, adapter, and latency distribution. If the condition fails, the renderer uses a previously approved continuation, loop, ambient bed, or graceful ending. It does not insert unverified bytes or block the audio callback. Emerging streaming-generation work makes changing-input, chunk-wise generation a plausible future adapter, but does not establish local performance or production readiness [30].

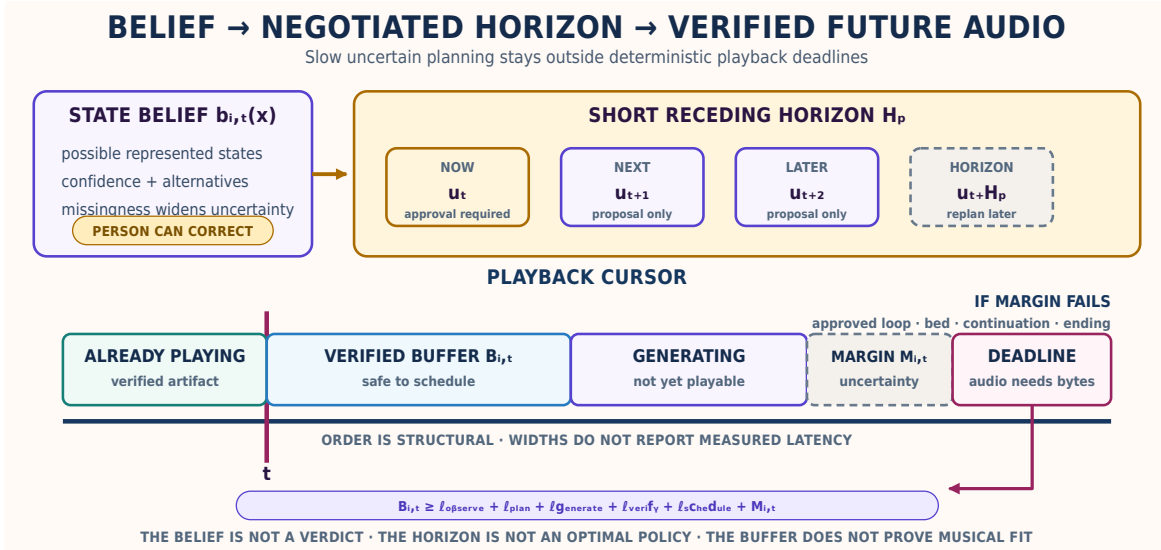


Figure 5: Receding-horizon planning and buffer deadlines under an uncertain, person-correctable state belief.

Figure description: A time axis shows the current playback cursor, verified buffered audio, generation in progress, and a future planning horizon. Above it, an uncertain state belief is corrected by the person before the next plan is accepted.

Deadline and safety margins are marked independently from claims about subjective response or model quality.

3.6 Semantic, musical, and waveform continuity

A continuously changing prompt does not by itself produce a continuous listening experience. Antidote therefore separates three questions: whether neighboring semantic conditions change within their declared bounds, whether the realized audio is musically and acoustically compatible at a boundary, and whether the renderer joins verified samples without a signal discontinuity.

For adjacent stages n and $n + 1$, semantic-plan distance is

$$D_n^{\text{sem}} = d_Q(q_{i,n}^{\text{tail}}, q_{i,n+1}^{\text{head}}), \quad (11)$$

ANT-EQ-011—Proposed operationalization; implementation: no validated semantic-distance function is selected.

where d_Q is a versioned, dimensionless distance over compatible typed controls and stage roles. It must expose which mixins contributed to the distance and cannot assume that interpolation in embedding or prompt space is perceptually linear.

Measured acoustic-boundary distance is independent:

$$D_n^{\text{ac}} = \sum_{r \in \mathcal{F}} w_r d_r(f_r(a_n^{\text{tail}}), f_r(a_{n+1}^{\text{head}})), \quad \sum_{r \in \mathcal{F}} w_r = 1. \quad (12)$$

ANT-EQ-012—Proposed operationalization; implementation: the mock analyzer does not validate musical boundary continuity.

The feature family $\mathcal{F}$ may include normalized distances for tempo and beat phase, key or chord compatibility, loudness, spectral balance and timbre, density, and phrase-boundary alignment. The w_r are nonnegative, dimensionless, and versioned; the feature-specific d_r functions normalize otherwise incompatible units before aggregation. Thresholds for D^{sem} and D^{ac} require perceptual and technical evaluation and may differ by person or listening context. Loudness standards can anchor measurement practice without supplying a universal transition or hearing-safety threshold [31].

Within an accepted overlap window $s \in [0, 1]$, one waveform-level candidate is the equal-power crossfade

$$y(s) = \cos\left(\frac{\pi s}{2}\right) a_n(s) + \sin\left(\frac{\pi s}{2}\right) a_{n+1}(s). \quad (13)$$

ANT-EQ-013—Proposed operationalization; implementation: no adaptive continuity renderer exists in v0.

Both input signals and y share the same sample-amplitude units. The equation can reduce a gain discontinuity; it cannot repair incompatible meter, phase, harmony, phrase, timbre, meaning, or emotional pacing. Resequencing, reorchestration, continuation generation, stem mixing, and a graceful ending are separate strategies. A transition is eligible for playback only after its artifact integrity, buffer margin, semantic condition, acoustic measurements, and relevant safety constraints pass their own checks. Failure of any one check invokes an approved fallback rather than being hidden by a crossfade.

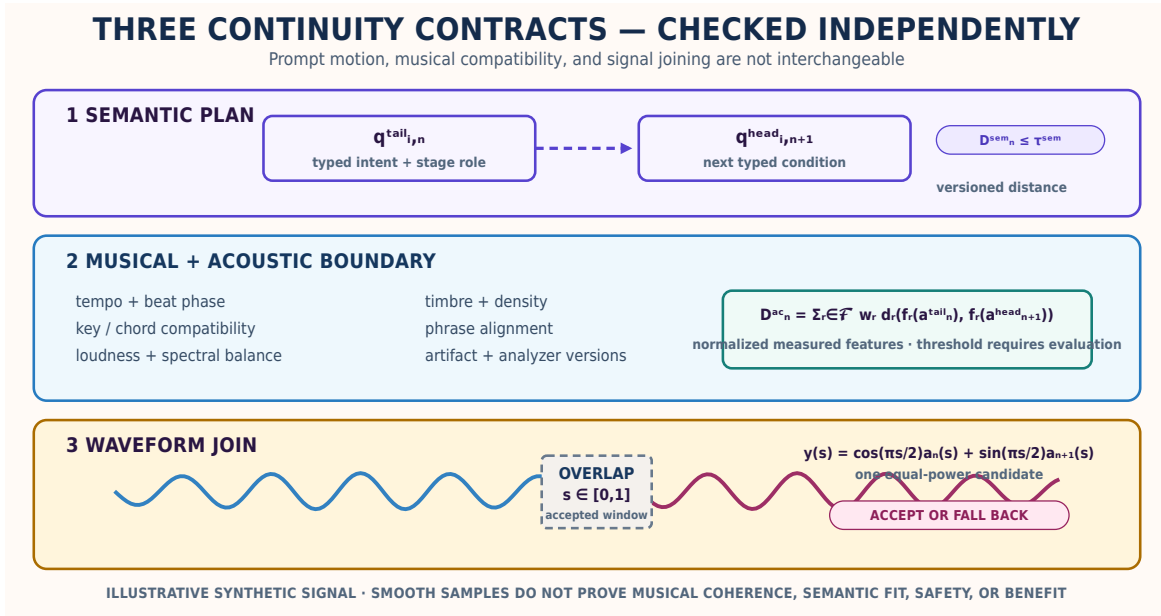


Figure 6: Independent semantic-conditioning and waveform-continuity checks across adjacent generated segments.

Figure description: The upper lane shows two versioned conditioning states and a bounded semantic transition. The lower lane shows adjacent audio segments, overlap or crossfade regions, waveform and feature measurements, and acceptance or fallback decisions. Separate indicators make clear that semantic similarity does not guarantee smooth audio and smooth audio does not prove intended meaning.

3.7 Exposure, response, and advisory updating

The generated segment a_n is not the exposure $e_{i,t}$. Exposure records which content hash was actually played, when playback began, duration heard, completion, interruption, stop reason, and safety events. Response is then a heterogeneous record linked to that exposure:

$$Y_{i,t} = \left(y_{i,t}^{\text{perc}}, y_{i,t}^{\text{felt}}, y_{i,t}^{\text{int}}, w_{i,t}^{\text{int}}, y_{i,t}^{\text{help}}, y_{i,t}^{\text{harm}}, y_{i,t}^{\text{res}}, y_{i,t}^{\text{mis}}, y_{i,t}^{\text{after}} \right). \quad (14)$$

ANT-EQ-014—Definitional; implementation: the mock slice captures a bounded subset and keeps exposure separate.

The components describe perceived musical expression, felt state, intensity, whether that intensity was wanted, immediate helpfulness, possible harm, resonance, mismatch, and a later aftereffect. They may contain text, bounded scores, categorical values, or missingness, so Equation 14 is a record declaration

rather than a numeric vector. Intense crying, for example, cannot be classified as success or harm from intensity alone; the person’s appraisal, stop behavior, and later account remain separate evidence.

A future within-person response model could estimate

$$p_{\theta_i}(\mathbf{Y}_{i,t} \mid e_{i,t}, z_{i,t}), \quad (15)$$

ANT-EQ-015—Future empirical model; implementation: no response predictor is fitted or used for generation.

where θ_i parameterizes a person-specific and versioned hypothesis. Conditioning on exposure prevents an unplayed artifact from becoming response evidence. The model would need to represent uncertainty, changed meanings, missing responses, contradictory sessions, and alternative explanations; it could not infer clinical benefit from acoustic adherence or immediate affect.

If a future protocol justified Bayesian updating, a candidate posterior form would be

$$p(\theta_i \mid \mathcal{D}_{i,1:t}) \propto p(\mathbf{Y}_{i,t} \mid e_{i,t}, z_{i,t}, \theta_i) p(\theta_i \mid \mathcal{D}_{i,<t}). \quad (16)$$

ANT-EQ-016—Future empirical model; implementation: v0 records evidence-linked proposals and prohibits autonomous updating.

Equation 16 is deliberately generic: it does not specify a valid likelihood, prior, forgetting rule, causal effect, or sufficient sample size. In the current authority model, eligible observations may support a descriptive proposal, but a new personal-model snapshot is created only after the person reviews the evidence and accepts the update. Rejection preserves both the prior snapshot and the rejected proposal. The design follows an interactive-machine-learning rationale while retaining the possibility that feedback requests can be burdensome or reduce trust [21, 32].

3.8 Provenance and failure boundaries

Every derived record links to the activity, agent, version, and source records that produced it. For generation this includes the accepted plan hash, adapter and model identity, code revision, effective parameters, seed when supported, device class, elapsed time, warnings, downgrade decisions, output hashes, and feature-analyzer versions. For exposure and response it includes the artifact hash, playback interval, interruption state, observation window, instrument version, missingness, corrections, and permission for later use. W3C PROV supplies an interoperable vocabulary for entities, activities, agents, and derivations, while Workflow Run RO-Crate demonstrates a packaging pattern for computational lineage [33, 34]. Antidote proposes to use those concepts; a provenance record remains an assertion to audit, not proof that its interpretation is true.

The system fails closed at the boundary whose precondition is missing. Invalid, expired, or ambiguous consent prevents projection. Conflicting mixins or an unsupported capability prevent compilation until the conflict or downgrade is reviewed. A worker crash, malformed response, timeout, partial output, path violation, size violation, or hash mismatch cannot become a playable verified artifact. A missed generation deadline invokes an approved fallback. A stop, reported harm, or safety event halts automatic continuation and is never scored as engagement. A failed or rejected model update preserves the prior snapshot. An export remains blocked when classification, minimization, redaction, consent, privacy review, or destination approval is incomplete.

Append-only events make those transitions reconstructable, while separately classified payloads permit future retention and deletion policies to act on sensitive content. The present repository implements integrity and replay for the synthetic path; encryption, key recovery, backup, secure deletion, projection invalidation, real-model sandboxing, and privacy-reviewed research export remain unresolved. Local

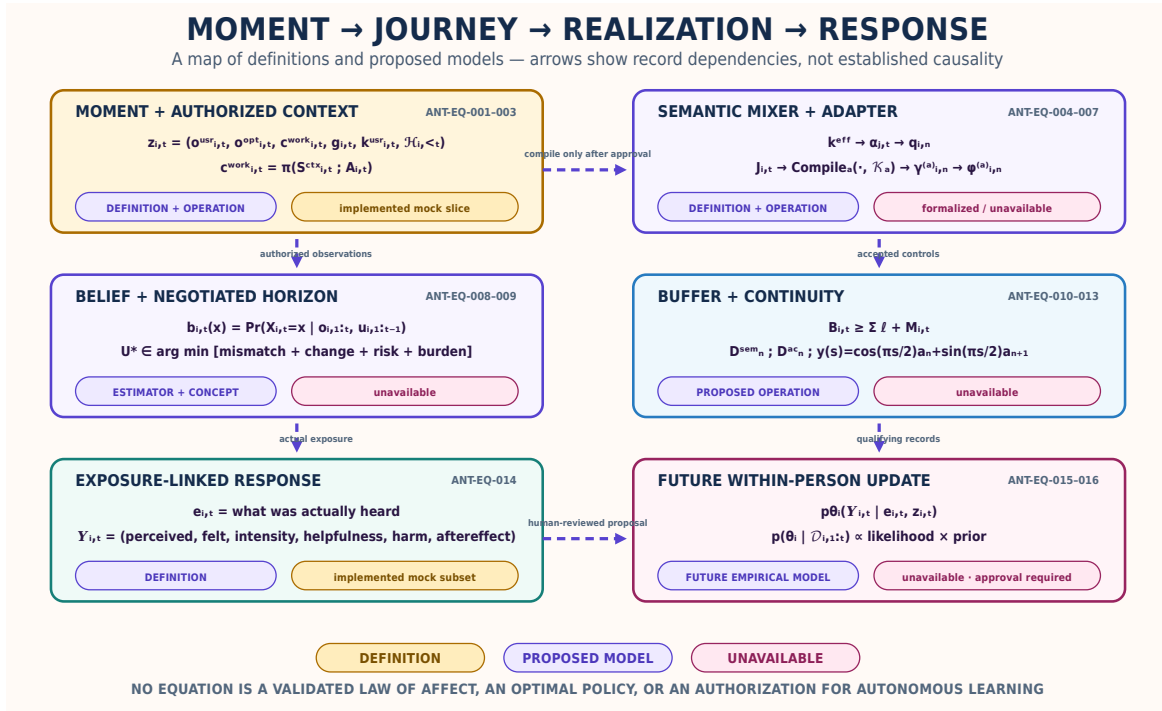


Figure 7: Conceptual map of the person-moment representation, journey decision, continuity, and cautious update equations.

Figure description: Six cards cover all 16 registered equations. They separate moment and authorized context; semantic mixing and model-specific adaptation; uncertain state belief and a negotiated planning horizon; buffer and independent semantic, acoustic, and waveform continuity; exposure-linked response; and a future within-person update. Each card states its epistemic class and implementation status. Dashed arrows show record dependencies rather than established causality, and the final warning rejects any reading of the equations as validated affect laws, optimal policies, or authorization for autonomous learning.

execution and content hashes reduce some dependencies and improve traceability, but they do not establish security, ownership, meaningful consent, safety, or benefit.

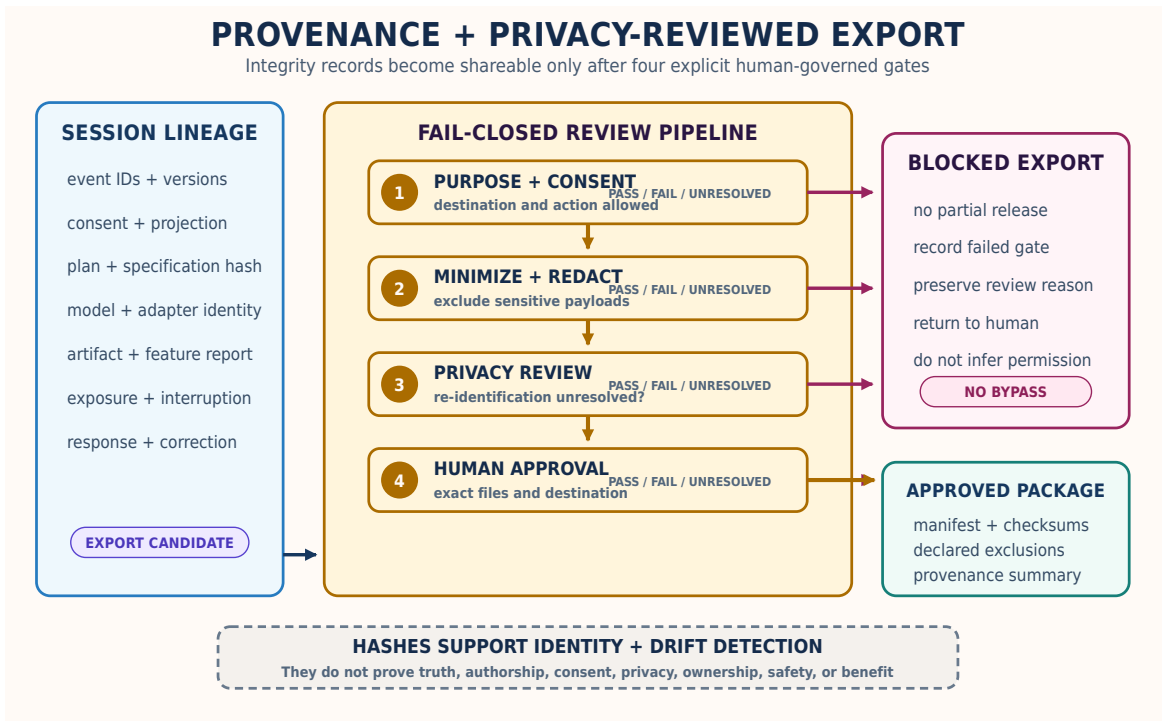


Figure 8: Provenance-linked session records pass through consent, redaction, privacy review, and explicit export approval.

Figure description: Versioned session events, model identity, control records, exposure, response, and artifact hashes feed an export candidate. Four gates check purpose and consent scope, excluded or sensitive fields, privacy review, and explicit human approval. Any failed or unresolved gate stops export and records the reason; a successful path emits a manifest and checksums without asserting truth or safety.

4 Methods

This paper specifies a prospective design and feasibility protocol rather than a clinical efficacy study. The canonical machine-readable artifact is ANT-PROT-FEAS-001, version 1.1.0, frozen on 2026-09-07 with status **frozen-design-protocol**. Issue #82 created this pre-collection corrective successor after identifying an assignment-space error and several underspecified analysis and response rules in version 1.0.0. The exact 1.0.0 bytes remain in `antidote-feasibility-v1.json` and are identified by SHA-256 `dbedbcbb74303373a7e00e95ad063b025d8cf6033864f4ad5a390d15703ba403`. No qualifying record, exposure, or human collection began under that version; accordingly, no formal human study has begun. Version 1.1.0 therefore changes no observed result. Its merge likewise grants no collection authority: all human-collection activation gates remain unsatisfied, and no ethics or safety approval is presumed. The design draws on N-of-1 protocol guidance for prospective specification and complete reporting, pilot and feasibility guidance for separating practicability from efficacy, and a digital N-of-1 platform as an implementation precedent [35–37]. Those sources inform reporting discipline; they do not validate Antidote’s measures, conditions, software, or proposed use.

4.1 Planned measurement layers

The protocol preserves intended semantic controls, model and adapter identity, generation settings, output identity and checksum, measured acoustic features, actual exposure, baseline and target states, separated subjective responses, timestamps, protocol version, and mutations between trials. These layers are non-interchangeable. A requested control is not a realized acoustic property; a measured acoustic property is not perceived expression; perceived expression is not felt state; and emotional intensity, completion, or engagement is not evidence of usefulness, harm, or clinical benefit. Behavioral or physiological measures are optional extensions and cannot become the sole definition of a person's experienced state.

4.2 Evidence classification

Research records use five non-interchangeable classes: **source** for what external evidence establishes, **hypothesis** for what the project proposes, **observation** for what occurred, **interpretation** for a bounded explanation of an observation, and **claim** for a statement supported in the manuscript. The working claim ledger is maintained outside publication artifacts under **research/notes/**. Synthetic fixture output and repository tests remain technical observations; they are not participant data.

4.3 Research questions and feasibility stages

The protocol maps the three research questions to four evidence stages. D0 addresses RQ1 through design traceability: each transformation and authority boundary must resolve to a versioned record, equation, contract, or explicit unavailable state. T0 addresses the implemented portion of RQ2 with synthetic, deterministic conformance and recovery tests. T1 would extend RQ2 to a frozen real-model, adapter, analyzer, hardware, and mutation manifest, but is currently **blocked-no-real-model**. H1 would address the record-feasibility prerequisite to RQ3 through a prospective repeated-session within-person series, but is **blocked-no-collection-authority**. Passing D0 or T0 neither implies that T1 will pass nor authorizes H1. H1 can test only whether separated records can be collected, corrected, retained, and provenance-linked with auditable completeness, voluntary-response usability, burden, and safety-event accounting. That is a prerequisite for RQ3, not a test of useful advisory relationships. Likewise, technical controllability cannot establish that an exposure is subjectively fitting, useful, safe, or therapeutically effective.

4.4 Units, conditions, and session structure

The technical unit is one immutable generation specification and its terminal result; the human-response unit is one assigned session attempt, with an exposure counted only when a verified artifact is actually played. The proposed H1 series contains six blocks and three conditions per block, for 18 planned completed exposures and an absolute ceiling of 24 scheduled attempts. Each artifact targets eight minutes. Exposures are separated by at least 48 hours, limited to three in any seven days, and followed by an aftereffect window targeting 24 hours (acceptable range 18–30 hours).

Condition G is a generic instrumental ambient generation that receives neither current state, desired transition, nor personal context. Condition S receives only selected current-state and desired-transition categories and a frozen meet–hold–transition–integrate template; it receives no free-text history, personal descriptors, or prior-session mappings. Condition P receives the explicitly approved working projection, person-authored semantic mixins, inclusions, exclusions, and an editable journey plan. Across conditions, a qualifying execution would hold constant the immutable model revision, adapter and analyzer versions, hardware profile, duration, output format, playback setup where feasible, response-instrument

version, and file target of -23 LUFS integrated loudness with a -1 dBTP true-peak ceiling. File-level normalization neither determines sound-pressure level at the ear nor establishes hearing safety; playback begins at the lowest device level and remains under the person's control. Auditory beats, within-exposure adaptation, and personal- model updating are excluded from this protocol.

These are visibly different condition-and-interface bundles. They jointly vary context fields, desired-transition use, personal semantic content, plan visibility, editing, and—for P—negotiated approval. Randomization permits only within-person contrasts among those bundles. It cannot isolate personal-ization, a semantic field, plan editing, approval, or a psychological or neurological causal pathway.

Each prospective session follows eight governed steps: authority and readiness; baseline and desired transition; condition reveal and plan review; generation and verification; a separate, explicit playback action; immediate response within five minutes; later aftereffect report; and closure or pause. A generated but unplayed artifact never enters an exposure or human-response denominator. The current desktop MVP is not this H1 instrument, and records produced through its schema-generation v1 runtime payloads are not H1-eligible. H1 remains blocked until, at minimum, the timed session workflow, separate version 2 consent enforcement, version 2 response and correction contract, assignment and seed enforcement, participant-facing instrument, safety workflow, and minimum-data private research export have been implemented and independently reviewed.

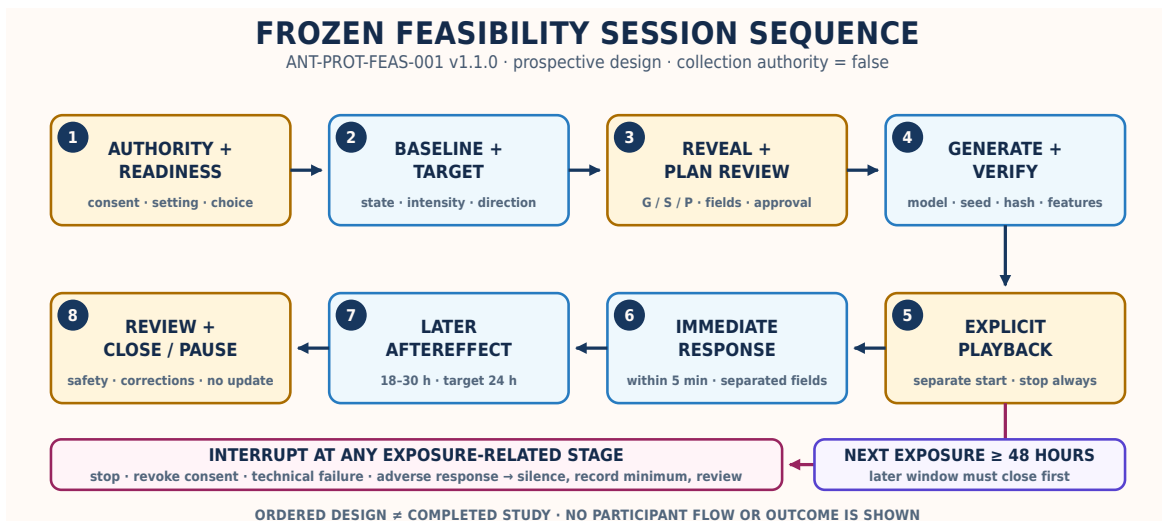


Figure 9: Prospective feasibility-session sequence with approval, exposure, response, aftereffect, and stopping boundaries.

Figure description: The timeline begins with protocol version and consent confirmation, then baseline and desired trajectory, plan review, generation and artifact verification, explicit playback approval, exposure with stop controls, immediate separated response measures, a declared aftereffect interval, and later review. Safety events may interrupt the sequence at every exposure-related stage.

4.5 Technical verification

T0 verifies the mock system's authority, schema, replay, interruption, recovery, hashing, and provenance invariants with complete denominators. T1, if activated after implementation review, freezes a real-model qualification manifest before any qualifying run. Its planned matrix contains three baseline runs, 24 one-factor mutation runs, six adjacent-segment continuity runs, and six injected failure-and-recovery runs: 39 attempts if all declared controls are supported. The fixed seeds are 101, 307, and 911.

The one-factor pairs vary tempo (70 versus 100 BPM), normalized density (0.25 versus 0.65), timbral instruction, or semantic intent while holding seed, model, adapter, hardware, duration, output format,

base plan, and all other controls constant. Tempo and density require versioned measured outputs. Timbral comparison requires a named analyzer and uncertainty. Semantic mutation is verified only as an adapter-conditioning trace unless an independently specified measure exists; it is not relabeled as acoustic adherence. Unsupported controls remain unsupported rather than being silently substituted.

The technical measures are authority-state conformance, required provenance completeness, component latency and verified buffer margin, requested-versus-measured control adherence, semantic and acoustic boundary distances, and terminal generation state. They correspond to the consent projection, adapter, generate-ahead, continuity, and response-record equations (ANT-EQ-003, ANT-EQ-007, and ANT-EQ-010–ANT-EQ-014). Hard gates require all invalid or unapproved transitions to fail closed; every playable artifact to match its declared hash, format, and bounded path; cancellation, timeout, stop, and safety halt to reach recorded terminal states; and no missed deadline, partial output, or unsupported control to be reported as success. Hardware- and adapter-specific latency, buffer, analyzer, and perceptual thresholds must be frozen in the later T1 manifest; this paper invents no universal threshold. For each registered component latency and verified buffer-margin variable, reporting fixes n , minimum, maximum, the ordinary arithmetic median, and $q_{0.50}$, $q_{0.90}$, and $q_{0.95}$. Quantiles use fixed linear interpolation: $h = (n - 1)p + 1$, $j = \lfloor h \rfloor$, $g = h - j$, and $Q(p) = (1 - g)x_{(j)} + gx_{(j+1)}$, with the upper index capped at n . Values are calculated from unrounded seconds separately for each component and terminal state. These summaries remain descriptive rather than universal thresholds.

4.6 Within-person feasibility measures

The proposed H1 instrument uses brief, near-event self-report because ecological momentary assessment can reduce recall distance while introducing its own burden, reactivity, adherence, and missingness concerns [38]. Baseline captures a person-authored description plus valence, arousal, and intensity; desired transition remains a separate nominal category with a person-authored description. Those descriptions are required by the current moment-context contract. The compact valence and arousal interaction borrows from the Affective Slider [39], but Antidote’s wording, intensity, response timing, and additional items constitute a new protocol and are not validated by that source.

After exposure, the instrument records perceived musical expression separately from felt state. It then records whether the intensity was wanted, as well as helpfulness, resonance, mismatch, harm, surprise, interaction burden, and total session burden. Helpfulness and harm are asked again in the later aftereffect window together with a yes/no/unsure ongoing-effect item and optional person-authored meaning, without overwriting the immediate record. Response descriptions are optional in the version 2 response contract; every prompted-but-absent response field instead receives its own field path and exactly one of `not_prompted`, `declined`, `missed_window`, `technical_failure`, `interrupted`, or `not_applicable`. A correction is a new record carrying the superseded response identifier and correction reason.

Every subjective scale is frozen. Valence runs from -1 “very unpleasant” through 0 “neither pleasant nor unpleasant, or evenly mixed” to 1 “very pleasant”; arousal runs from -1 “very low activation” through 0 “neither low nor high” to 1 “very high activation”; intensity runs from 0 “not at all intense” to 1 “extremely intense.” Wanted intensity and ongoing effect each use the nominal responses yes, no, and unsure. Helpfulness runs from “not at all helpful for my chosen purpose” to “as helpful as I could reasonably ask for that purpose”; resonance from “did not fit or resonate at all” to “fit or resonated completely”; mismatch from “no perceived mismatch” to “extreme perceived mismatch”; harm from “no unwanted negative effect noticed” to “extreme unwanted negative effect”; and surprise from “entirely expected” to “entirely unexpected.” Interaction burden and whole-session burden are separate items, each running from “not burdensome” to “too burdensome to repeat.” No clinical outcome is collected, and no emotional direction is prespecified as inherently better. In particular, catharsis, crying, arousal, or movement toward a target cannot substitute for the person’s appraisal of whether the experience

was wanted, helpful, mismatched, or harmful.

Physiology is disabled in version 1.1.0. A future supplemental module would require separate consent; device, calibration, clock, unit, signal-quality, artifact, dropout, and missingness records; a frozen analysis; and independent privacy and safety review. It could not overwrite self-report, diagnose emotion or need, prove entrainment, or steer generation under this protocol.

4.7 Assignment, carryover, and missingness

The six possible orders of G, S, and P are each used once. Their block order is randomized with a cryptographically generated seed committed in a separate activation record before session 1. This is a bijection of six distinct orders to six chronological blocks, so the allowable schedule space is exactly $6! = 720$. It is not six independent draws. The machine-readable protocol maps a 32-byte seed uniformly to a zero-based lexicographic schedule index using SHA-256 rejection sampling and records the counter, index, schedule, and commitment. H1 additionally fixes block seeds 101, 307, 911, 1217, 1601, and 2027 in chronological order and uses the same block seed for G, S, and P. A model or adapter that cannot honor that commitment leaves H1 blocked. The committed seed is never substituted; any pre-exposure retry retains it, and no retry is allowed after hearing or rating its output. Eligibility and baseline are confirmed before condition reveal. Participant blinding is impossible because the three interfaces disclose different context and plan operations; researcher blinding is also impossible in a self-study, and both facts must be reported. A failed assigned session remains in the attempt ledger. A replacement is allowed only before any exposure starts, retains the same block, slot, condition, and committed block seed, and never erases the failure. Once playback starts, the slot is consumed regardless of completion or response and cannot be rerun based on what was heard or reported.

At least 48 hours and completion or expiry of the later-response window separate exposures. Ongoing aftereffect, mismatch, distress, hearing discomfort, or uncertainty postpones the next attempt by at least another 24 hours and triggers a repeated readiness review. This spacing may reduce but cannot prove removal of carryover. Flagged sessions remain in primary accounting and are examined in a prespecified sensitivity analysis.

Analyses distinguish all assigned attempts, verified artifacts, started exposures, completed exposures, immediate windows, and later windows. No outcome is imputed in the primary analysis. Each prompted but absent field carries its own canonical missingness reason. Condition contrasts use only blocks containing both required observations and disclose every omitted block. Lower- and upper-endpoint values are used only for the prespecified interaction-burden progression sensitivity, never as substituted observations.

4.8 Analysis and sensitivity plan

Technical reporting includes complete state denominators, invariant outcomes, the frozen latency summaries, paired control differences by factor and seed, and every semantic and acoustic boundary component. Human-response feasibility reporting includes assigned through withdrawn counts by condition; audit completeness; usable voluntary values by field and canonical missingness reason; separate interaction- and whole-session-burden distributions; all mismatch, harm, hearing-discomfort, distress, and safety events; and the ability to link semantic, acoustic, exposure, response, and correction records through provenance. A declined response can be audit-complete but is not counted as a usable response.

Secondary within-person estimands are blockwise P-minus-S, P-minus-G, and S-minus-G contrasts for immediate helpfulness, resonance, mismatch, harm, and interaction burden and whole-session burden separately; differences between perceived-expression and felt-state valence and arousal; and

Registry ID / construct	Timing / evidence	Instrument, scale, provenance	Strongest permitted boundary
MEAS-BASELINE-STATE person-authored current state	pre-condition self-reported	Text + valence/arousal [-1,1] + intensity [0,1]. Source: moment-context.current_state	Not an objective emotion, diagnosis, or need.
MEAS-DESIRED-TRANSITION desired transition	pre-condition self-reported	Selected category + required text; nominal, not ordinal. Source: moment-context.desired_transition	Not a promised endpoint or clinical goal.
MEAS-SEMANTIC-CONTROLS requested semantic intent and plan	pre-generation self-reported-and-decided	Approved projection + plan revision; typed edits and exclusions. Source: working-context + journey-plan	Requested controls are not realized acoustics or felt response.
MEAS-ACOUSTIC-REALIZATION generated acoustic realization	post-generation-pre-exposure model-estimated-and-technical-observation	Artifact verifier + feature analyzer; feature-specific units and warnings. Source: generation-result	Measured features do not reveal perceived expression, felt state, usefulness, harm, or benefit.
MEAS-EXPOSURE actual exposure	during playback observed	Playback event; artifact hash, timestamps, seconds heard, and stop state. Source: core Exposure event	A generated but unplayed artifact is not an exposure.
MEAS-PERCEIVED-EXPRESSION emotion or meaning perceived in the audio	immediate self-reported	Optional text + valence/arousal [-1,1] + intensity [0,1]. Source: response-v2.perceived_expression	Perceived expression is not the listener's felt state.
MEAS-FELT-STATE felt response	immediate-and-later self-reported	Optional text + valence/arousal [-1,1] + intensity [0,1]; field missingness. Source: response-v2.felt_state	Felt state is not automatically helpful, harmful, or clinically meaningful.
MEAS-APPRAISAL wanted intensity, helpfulness, resonance, mismatch, harm, and surprise	immediate; helpfulness and harm repeated later self-reported	Wanted intensity yes/no/unsure; helpfulness, resonance, mismatch, harm, surprise [0,1]. Source: response-v2 appraisal fields	Intensity, crying, completion, or engagement is never substituted for helpfulness or harm.
MEAS-BURDEN interaction and whole-session burden	immediate self-reported	Interaction burden + whole-session burden [0,1]. Source: response-v2 burden fields	Low burden does not establish benefit or meaningful consent.
MEAS-AFTEREFFECT ongoing effect and optional person-authored meaning	baseline-readiness-and-later self-reported	Ongoing effect yes/no/unsure + optional later meaning. Source: response-v2 aftereffect fields	An ongoing effect is not presumed helpful, harmful, caused solely by the audio, clinically meaningful, or mechanistically explained.
MEAS-MISSINGNESS unavailable or declined response fields	every collection window observed-or-self-reported	Per-field reason: not prompted, declined, missed, failed, interrupted, or not applicable. Source: response-v2.missingness	Missing is not neutral, zero, or a successful response.
MEAS-CORRECTION append-only person-authored response correction lineage	whenever a response is corrected observed-and-self-reported	Append-only revision, superseded response ID, and person-authored reason. Source: response-v2 correction lineage	A correction preserves revision history and does not prove that either version is objectively correct.
MEAS-CLINICAL clinical outcome	not-collected unavailable	No instrument; not collected; not applicable. Source: explicit unavailable state	No feasibility field may be promoted into a clinical endpoint.

Prospective projection of ANT-PROT-FEAS-001 v1.1.0. Collection authority is false. Scale shorthand preserves the frozen anchors; prompted but absent fields retain explicit missingness rather than a substituted value.

Table 3: Prospective measurement layers, instruments, timing, provenance, and permitted interpretation.

Table description: Each row names one non-interchangeable measurement layer. Columns identify the planned instrument, collection timing, provenance fields, missing-data treatment, and the strongest permitted interpretation. Clinical outcome is shown as outside the initial feasibility protocol rather than inferred from the other rows.

condition-specific baseline-to-immediate valence and arousal changes. Protected analysis output retains every session point; the public envelope exposes only privacy-reviewed aggregate denominators, medians, ranges, and eligible blockwise contrasts. Exploratory variable-pair searches and multivariable response regressions are not run: with at most 18 exposures, neither a search nor simultaneous condition, block, order, and baseline coefficients are defensibly supported. A later protocol would have to name its variables, model, and sample-size rationale prospectively.

For a named immediate outcome and pairwise contrast, an exact randomization distribution is available only when all three chronological position values are present in all six blocks. It holds those values fixed, enumerates lexicographically all 720 bijections of the six condition orders to the six blocks, assigns labels under each schedule, and uses the mean of the six within-block A-minus-B differences as the statistic. The two-sided tail includes every unrounded statistic with absolute value at least the observed absolute value, including ties, and reports the count divided by 720. All 720 statistics are retained in the protected analysis output. A public aggregate reports the observed statistic, tail count, denominator, exact proportion, and privacy-reviewed distribution summaries; the full vector is not released without a separate field-level disclosure review and current export scope. No reduced, asymptotic, one-sided, plus-one, or multiplicity-adjusted substitute is used. This is an exploratory reference for the single person and schedule under a sharp no-bundle-effect null, not population inference or component-level causal attribution.

Sensitivity analyses compare started exposures with those reaching 80 percent of audio (384 of 480 seconds), include and exclude named carryover flags, report blocks 1–3 and 4–6 separately without treating the split as a trend test, leave out each block in turn, and stratify by administration status only if it actually changes. For missing interaction-burden values only, the progression decision is repeated once with 0 and once with 1. Audit completeness, completion, usable-response proportions, provenance, and unresolved safety are factual and cannot be rescued by bounds. Helpfulness and resonance otherwise use 1/0 as descriptive best/worst endpoints; mismatch, harm, and both burden fields use 0/1; valence, arousal, intensity, and surprise receive endpoint ranges without calling either direction better; and nominal fields are never numerically imputed. All prespecified null, zero, opposite-direction, negative, missing, interrupted, burdensome, harmful, and adverse observations remain visible. No endpoint will be selected or renamed after observation.

The progression rule is operational rather than therapeutic. A later bounded study may be considered only if the hard technical gates pass; at least 15 of 18 exposures and at least four per condition complete; every assigned attempt and scheduled window is audit-complete; and every immediate core field has at least 80 percent of the started-exposure count, rounded upward to the next whole value, overall and at least three per condition, while every later core field has at least 70 percent of that count, again rounded upward, overall and at least two per condition. Every eligible response must preserve the full protocol-to-correction provenance chain. Observed median interaction burden and both its endpoint-bound decisions must be no greater than 0.5; no safety event may remain unresolved; and the participant and an independent reviewer must separately find another bounded study acceptable. Otherwise the decision is stop or revise. Even a “proceed” decision establishes neither advisory usefulness, efficacy, generalizability, privacy, nor safety.

Equation ANT-EQ-014 defines the separated response record used here. ANT-EQ-015 and ANT-EQ-016 remain unavailable future empirical models: this protocol fits no response predictor, specifies no likelihood, and authorizes no personal-model update.

4.9 Protocol freeze, safety, and deviations

The canonical protocol, its SHA-256 lock, and an initially empty append-only deviation log are stored under `experiments/protocols/`. Before the first qualifying record, a material change requires a new

semantic protocol version, content lock, analysis plan, and review. After collection begins under separate authority, the frozen file is never rewritten: amendments and deviations append the rationale, affected records, evidence impact, author, reviewer, and disposition. Conditions, assignment, timing, measures, estimands, exclusions, model or analyzer identity, consent, privacy, retention, safety, physiology, and missingness rules are material changes.

The person may postpone, pause, skip, stop, withdraw, or decline any response without penalty or automatic substitution. A stop request or consent revocation, hearing pain or unexpected loudness, unwanted escalating distress or disorientation, technical or provenance failure, or any loss of freely chosen participation stops playback and prevents automatic continuation. The interface must offer silence and a clear exit, retain only the minimum consented event, and defer optional questions until the person chooses. Urgent concerns belong with appropriate human or emergency support, not Antidote. Unresolved adverse responses, serious events, stop-control failure, or repeated high harm pause the protocol for independent review; these rules do not prove clinical safety.

For those pause rules, “high harm” means a non-missing immediate or later harm rating of at least 0.75, or an explicit request by the person to treat an event as high harm. “Repeated” means high harm associated with two distinct started exposures; immediate and later reports from one exposure count once. A “serious event” is operationally an event for which urgent or emergency human assessment is sought; reported hospitalization, persistent hearing change, or persistent functional impairment; incapacity preventing a self-directed exit; continued playback after a stop request; or unauthorized access, retention, or export of consent-scoped sensitive data. This is not a diagnosis or clinical severity judgment. An event is unresolved while it lacks a documented disposition, a reported unwanted effect remains ongoing or unsure, or required independent review is incomplete. Serious events and stop-control, privacy, or consent failures cannot close automatically.

Separate consent scopes govern inspection, projection, generation, playback, retention, analysis, export, and learning (future personal-model use). Raw journal, therapy-chat, health, or other private free text cannot enter the public repository or a default research export. Storage encryption and recovery, retention expiry and deletion, projection invalidation, backups, participant materials, the applicable ethics determination, assignment commitment, version 2 consent enforcement, protocol timing, version 2 response and correction handling, minimum-data private research export, instrument implementation, analysis code, and pilot stop/go review all remain unresolved activation gates. Consequently, this protocol is reproducible as a design artifact but is not permission to recruit, expose, collect, or report a human result.

5 Results

This design/protocol manuscript contains no formal human results. No qualifying formal-study or H1 package is registered. No qualifying T1 package is registered; therefore no real-model technical result is reported. The originating catharsis experience remains a hypothesis-generating observation rather than a study result. These statements describe the governed repository record, not events that might have occurred outside the protocol. The absence of a qualifying record is reported as *not run*, *not collected*, *blocked*, or *not applicable*; it is never entered as a zero.

Results reporting is governed by ANT-REPORT-RESULTS-001, version 1.1.0. It preserves the original issue #41 governance and records issue #82 as the prospective pre-collection amendment. The machine-readable contract binds each future result slot to a named source record, the byte-locked ANT-PROT-FEAS-001 protocol version 1.1.0, a frozen analysis pointer, an evidence class, and a claim-promotion rule. It currently registers no T0, T1, or H1 result package. This structure adopts the complete-accounting orientation of N-of-1 and pilot-feasibility reporting guidance—including participant

flow, missingness, harms, analysis, and progression—without describing Antidote as a completed trial or effectiveness study [36, 40].

5.1 Technical verification reporting

The D0 protocol record is available for inspection. Its identity and content lock are repository artifacts, while the reporting contract declares the intended slot, schema, equation, and activation-gate mappings. This is evidence that versioned design records and explicit mappings exist; it is not an execution result, an independent audit of implementation, or evidence of feasibility, privacy, safety, or benefit.

The repository also contains an executable synthetic session path and its component tests, as bounded by ANT-OBS-002. Those artifacts remain development and conformance evidence. The ANT-Q03 exit gate in issue #18 has not produced the reserved ANT-REC-T0-001 result package, so no T0 rate, count, latency, or success claim is promoted here. Closing an issue or observing a green workflow would not be sufficient by itself. Promotion requires the reviewed source revision, execution environment, exact commands, planned denominators, every terminal state, injected failure and recovery, artifact hashes, exclusions, and an explicit claim-ledger decision. Any accepted T0 statement must retain the labels *synthetic* and *technical* and cannot imply a human exposure or response.

No interface capture is included in the current manuscript because the qualifying ANT-REC-T0-001 package does not exist. The governed visual slot is retained as retired provenance so a later revision can reactivate it only with a reproducible reviewed build, capture metadata, an explicit synthetic-or-approved-data label, and redaction review. Interface presence could demonstrate that a control is displayed; it could not demonstrate that the underlying authority is complete or that an experience is helpful.

5.2 Control adherence and continuity

T1 remains **blocked-no-real-model**. No T1 qualification manifest has frozen a specific model, adapter, analyzer, hardware profile, technical threshold set, or run plan, and ANT-REC-T1-001 does not exist. Consequently, the governed table below contains source and analysis identities plus current evidence states, but no attempt accounting, authority or artifact-integrity invariant results, requested controls, measured features, adherence values, latencies, buffer margins, boundary distances, failure rates, or uncertainty estimates.

GOVERNED EMPTY STATE — ANT-TBL-004

Future result slot	Governance binding	Current evidence state
Mock exit-gate conformance	Source: ANT-REC-T0-001 Stage: T0 Evidence class: technical-synthetic Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-T0-STATE-001	Awaiting issue #18; no qualifying package
Real-model attempt accounting	Source: ANT-REC-T1-001 Stage: T1 Evidence class: technical Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-T1-ACCOUNTING-001	Blocked; no qualified real model or run
Authority and artifact-integrity invariants	Source: ANT-REC-T1-001 Stage: T1 Evidence class: technical Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-T1-INVARIANTS-001	Blocked; no qualified real model or run
Real-model control adherence	Source: ANT-REC-T1-001 Stage: T1 Evidence class: technical Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-T1-ADHERENCE-001	Blocked; no qualified real model or run
Generation timing and buffer margin	Source: ANT-REC-T1-001 Stage: T1 Evidence class: technical Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-T1-TIMING-001	Blocked; no qualified real model or run
Continuity, interruption, and recovery	Source: ANT-REC-T1-001 Stage: T1 Evidence class: technical Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-T1-CONTINUITY-001	Blocked; no qualified real model or run

Exact planned paths and protocol pointers are governed by ANT-REPORT-RESULTS-001. No qualifying values are registered; absent is not zero.

Table 4: Governed T0 and T1 result slots with their required source records, protocol, analyses, and current unavailable states; no qualifying values are registered.

Table description: The generated empty-state table maps each future technical result to a source package, protocol version, and analysis identifier. The mock exit-gate row awaits issue 18, while all real-model rows are blocked because no model qualification has run. The table reports no rates, latencies, adherence estimates, boundary distances, or subjective outcomes.

If T1 is later activated, its matrix reaches 39 attempts only when every declared control is supported; otherwise reporting will name each unsupported or missing attempt and the resulting denominator. Authority and artifact-integrity invariants will each be reported as pass, fail, or unavailable rather than being inferred from an aggregate success state. Requested-versus-measured differences will be paired by factor and seed; generation latency and verified buffer margin will name the hardware context; semantic and waveform boundary components will remain separate; and partial output, timeout, cancellation, failure, and recovery will stay in their denominators. Unsupported semantic controls will be reported as unsupported rather than assigned an invented acoustic measure. A technically accurate tempo, sufficient buffer, or small waveform boundary does not establish perceived continuity, affective fit, safety, or therapeutic value.

No benchmark graphic is included until the T1 source package and analysis outputs pass promotion review. Its governed visual slot is retained as retired provenance. A future reactivation must show denominators, uncertainty, missing runs, failures, model and hardware revisions, and source-package and analysis identities rather than only favorable traces.

5.3 Exposure and response completeness

H1 remains **blocked-no-collection-authority**; none of its activation gates is satisfied and no assigned attempt, generated study artifact, exposure, or response record is registered. The table shell names the future flow, immediate-response, later-response, cross-package provenance, and progression analyses. Its blocked cells are not denominators and must not be read as zero events.

GOVERNED EMPTY STATE — ANT-TBL-005

Future result slot	Governance binding	Current evidence state
Assigned attempt and exposure flow	Source: ANT-REC-H1-FLOW-001 Stage: H1 Evidence class: observed Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-H1-FLOW-001	Blocked; no collection authority or records
Immediate response completeness	Source: ANT-REC-H1-RESPONSE-001 Stage: H1 Evidence class: observed-or-self-reported Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-H1-IMMEDIATE-COMP-001	Blocked; no collection authority or records
Later aftereffect completeness	Source: ANT-REC-H1-RESPONSE-001 Stage: H1 Evidence class: observed-or-self-reported Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-H1-LATER-COMP-001	Blocked; no collection authority or records
Cross-package provenance linkability	Source: ANT-REC-H1-AUDIT-001 Stage: H1 Evidence class: observed Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-H1-PROVENANCE-001	Blocked; no collection authority or records
Stop, revise, or proceed disposition	Source: ANT-REC-H1-AUDIT-001 Stage: H1 Evidence class: decided Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-H1-PROGRESSION-001	Blocked; no collection authority or records

Exact planned paths and protocol pointers are governed by ANT-REPORT-RESULTS-001. No qualifying values are registered; absent is not zero.

Table 5: Governed H1 flow and response-completeness slots with required source records and analyses; collection is not authorized and no values are registered.

Table description: The generated empty-state table maps assigned-attempt and exposure flow, immediate response completeness, and later aftereffect completeness to their planned source packages, protocol version, and analysis identifiers. Every row is blocked because H1 collection lacks authority. The blocked state is not a zero denominator or evidence that no response occurred.

After separately authorized collection, reporting would begin with every assigned attempt and retain the transitions to attempted generation, verified artifact, started playback, completed exposure, immediate window, and later window. Conditions G, S, and P would be shown separately and together. Generated but unplayed audio would remain outside the exposure denominator; an interrupted session would remain in the attempt ledger; a replacement would not erase the original assignment; and withdrawal would not be converted into technical failure. Each response field would be classified as present or with exactly one canonical reason: `not_prompted`, `declined`, `missed_window`, `technical_failure`, `interrupted`, or `not_applicable`. The wanted-intensity and ongoing-effect responses of *unsure* are present answers, not missingness.

The progression report would show every criterion as pass, fail, or indeterminate and would keep audit completeness separate from usable voluntary response completeness. An explicit decline can complete the audit trail but cannot be counted as a response value. The public H1 audit envelope would also have to reconcile the flow, response, and safety envelopes before any H1 slot could be promoted. A stop, revise, or proceed disposition concerns only a later bounded protocol and cannot establish advisory usefulness, improved targeting, efficacy, privacy, generalizability, or safety.

5.4 Subjective response and aftereffect

No perceived-expression, felt-state, helpfulness, resonance, mismatch, harm, surprise, burden, or aftereffect observation is reported. The reporting contract reserves six distinct H1 analysis views: perceived musical expression versus felt state; baseline-to-immediate felt-state change; immediate appraisal by condition; exact randomization uncertainty; prespecified sensitivities; and immediate versus later aftereffect. The first preserves both reports without declaring either one correct. The second uses only paired present baseline and immediate valence or arousal values. The third retains every session point and prespecified within-block P-minus-S, P-minus-G, and S-minus-G contrast in the protected analysis output, while the public envelope exposes privacy-reviewed aggregate summaries only. These compare bundled conditions and cannot isolate the effect of personalization, context access, or plan editing. The sixth retains the immediate and later records as separate window-specific protected accounts rather than using later meaning to rewrite the immediate response.

Future public summaries must display exact aggregate denominators, prespecified medians, ranges, and blockwise summaries; session-level values remain in protected analysis output. The exact randomization distribution is eligible only if all six blocks satisfy its frozen analysis conditions. It enumerates the 720 assignments of the six unique condition sequences to the six chronological blocks; it is exploratory within-person uncertainty, not a population estimate or confirmatory test. No clinical outcome is collected, and movement toward a desired state cannot substitute for the person's appraisal of whether an experience was wanted, helpful, mismatched, burdensome, or harmful.

Any populated report must also retain the prespecified sensitivity views: started exposures versus those with at least 80 percent of audio heard, with versus without carryover flags, blocks one through three versus four through six, all six leave-one-block-out summaries, administration status only if both statuses occur, and the declared lower- and upper-bound burden decisions. These are reporting views rather than opportunities to select a preferred estimate after observing the data. A bound that changes the decision makes that criterion indeterminate and the final disposition *revise*; it is not resolved in the preferred direction. None has been computed.

5.5 Null, negative, missing, and adverse observations

The empty-state table gives null or opposite-direction response, mismatch or unwanted intensity, burden, harm or adverse response, missingness, and interruption the same structural visibility as a future

favorable observation. It contains no event counts because H1 has not begun.

GOVERNED EMPTY STATE — ANT-TBL-006

Future result slot	Governance binding	Current evidence state
Null or opposite-direction response	Source: ANT-REC-H1-RESPONSE-001 Stage: H1 Evidence class: self-reported Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-H1-NONPOSITIVE-RESPONSE-001	Blocked; no collection authority or records
Mismatch or unwanted intensity	Source: ANT-REC-H1-SAFETY-001 Stage: H1 Evidence class: self-reported Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-H1-NONPOSITIVE-SAFETY-001	Blocked; no collection authority or records
Interaction or session burden	Source: ANT-REC-H1-SAFETY-001 Stage: H1 Evidence class: self-reported Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-H1-NONPOSITIVE-SAFETY-001	Blocked; no collection authority or records
Harm or adverse response	Source: ANT-REC-H1-SAFETY-001 Stage: H1 Evidence class: self-reported-or-observed Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-H1-NONPOSITIVE-SAFETY-001	Blocked; no collection authority or records
Missing, declined, or unavailable	Source: ANT-REC-H1-RESPONSE-001 Stage: H1 Evidence class: observed-or-self-reported Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-H1-MISSINGNESS-001	Blocked; no collection authority or records
Stop, interruption, or technical failure	Source: ANT-REC-H1-FLOW-001 Stage: H1 Evidence class: observed Protocol: ANT-PROT-FEAS-001 v1.1.0 Analysis: ANT-AN-H1-NONPOSITIVE-FLOW-001	Blocked; no collection authority or records

Exact planned paths and protocol pointers are governed by ANT-REPORT-RESULTS-001. No qualifying values are registered; absent is not zero.

Table 6: Governed slots for null, opposite-direction, mismatched, burdensome, harmful, missing, stopped, interrupted, and adverse observations; no H1 values exist.

Table description: The generated empty-state table assigns source and analysis identities to null or opposite-direction response, mismatch or unwanted intensity, burden, harm or adverse response, missing or declined data, and stop or interruption. Every row is blocked because H1 has not begun. The table contains no fabricated zeroes and does not treat intense emotion as success.

In a qualifying report, zero movement would remain distinct from a missing measurement; opposite-direction movement would remain distinct from harm; mismatch would remain distinct from unwanted intensity; *unsure* would remain distinct from missing wanted intensity or ongoing effect; and *declined* would remain distinct from *missed_window* or *technical_failure*. Stops, interruptions, hearing discomfort, distress, consent or privacy failures, technical safety failures, and participant-described adverse responses would be listed with their denominators and review dispositions. None could be silently removed because it weakens a narrative. Conversely, the absence of an observed adverse event in a bounded series could not establish clinical safety.

5.6 Optional physiological observations

Physiology is disabled under protocol version 1.1.0. No physiological source record, analysis, panel, or result row is authorized, and no vital sign or sensor estimate may steer generation within this protocol. A future module would require a protocol amendment, separate consent, device and clock records, signal-quality and artifact handling, a frozen analysis, and independent privacy and safety review. Even then, physiology would remain supplemental: it could not diagnose emotion or need, overwrite self-report, demonstrate entrainment, or establish benefit.

5.7 Promotion and replacement rule

The current Results state is a governed empty-state specification: its three tables are deterministic version 1.1 projections bound to the final protocol lock, while the two evidence-contingent visual slots are retired and absent from the manuscript. To replace an empty state, a future change must add the reserved source package, validate its protocol and analysis identity, preserve complete denominators and non-positive records, register any deviations, and record a claim-ledger promotion decision. The reporting contract and generated projection must then be versioned together. A source package may remain available in the repository while its manuscript claim is rejected or downgraded.

Every H1 repository source is a privacy-reviewed aggregate envelope with a `.public.json` suffix. The flow, response, safety, and audit envelopes must share one evidence-set identity and opaque, nonce-bound integrity commitment. Public envelopes may contain aggregate accounting, prespecified summaries, uncertainty, public-safe artifacts, and review dispositions; they may not contain participant, session, consent-grant, or source-event identifiers, raw or free-text responses, row-level observations, private paths, per-record hashes, stable cross-package pseudonyms, or participant-event timestamps. The private evidence index and commitment nonce remain outside Git. Their commitment supports reconciliation only; it does not prove anonymity, consent, authenticity, truth, completeness, efficacy, or safety.

This gate is deliberately asymmetric: qualifying evidence may populate only the slot it supports, while failure at any provenance, authority, analysis, or claim boundary leaves the slot unavailable. Technical verification cannot unlock human-response claims, H1 feasibility cannot establish efficacy, and no result at these stages can establish a neurological mechanism.

6 Discussion

This section interprets the formalized design, considers simpler explanations and alternative architectures, and identifies the evidence gates that would be needed to answer the later feasibility questions. Its conclusions remain at the design and protocol level; a coherent record model is not an observed listener outcome.

Following AffectMachine-Classical, the term *probability-space sculpting* is used only as a computational metaphor [4]: semantic prompting and other controls constrain a conditional distribution over possible generations rather than deterministically selecting one sound. The metaphor becomes scientifically useful only if the intended controls and their realized acoustic consequences can be operationalized and tested.

6.1 Contribution at the design evidence level

Antidote's contribution at this stage is the specification of a research object and its governance, not evidence that an adaptive audio system changes a listener's state. Current-to-target selection, person-

specific modeling, generated affective trajectories, continuous musical feedback, interpretable controls, and semantic planning each have prior precedent [1–5]. The narrower contribution is to formalize how consented context, semantic intent, a time-indexed plan, adapter-specific conditioning, acoustic realization, actual exposure, separated response constructs, provenance, and a possible later update can remain individually inspectable. Our bounded review did not identify that complete governed chain in one evaluated system. This supports a candidate research-system contribution, not a global first-system claim.

Two separations carry most of the design rationale. The first is epistemic: the record that requests a property is not the measurement of its realization, and neither record determines what a person perceived, felt, found helpful, or found harmful. The second concerns timing and authority: in the proposed two-rate architecture, uncertain semantic deliberation may prepare verified future material, while a deterministic renderer would own audio deadlines and the immediate stop path. These separations make disagreements locatable in principle. A plan may be accepted while an adapter visibly downgrades one control; an artifact may satisfy its file checks while missing a requested acoustic property; or playback may be technically continuous while the person reports mismatch. Preserving those possibilities is more defensible than calling a prompt, affect label, or aggregate score “interpretable.” It does not guarantee that the records are understandable, correct, complete, or safe.

Three proposed components turn this separation into falsifiable design claims. The Semantic Intent Mixer can be tested for preservation of typed controls, exclusions, authority precedence, conflict resolution, and visible adapter downgrades; its control values do not imply linear changes in acoustics or felt response. The predictive horizon is a receding horizon over candidate controls, not a forecast of the person’s lived state: it can be tested by whether only the first approved action is committed and whether verified future material meets a frozen generation deadline. Continuity must be tested separately as semantic-plan distance, measured acoustic-boundary distance, and waveform scheduling; success at one layer cannot establish the others or a person’s perception of continuity. None of the three is implemented or validated end to end.

Five evidentiary distinctions must therefore remain non-interchangeable. Demonstrated implementation means that a named capability executes at a reviewed source revision; at that revision, the only executable session evidence is the rule-guided synthetic slice described in Section 3.4. Technical feasibility must remain stage-specific: qualifying T0 records can establish only the synthetic authority, schema, replay, interruption, recovery, hash, and provenance invariants; T1 would additionally require real-model adherence, continuity, latency and buffer-margin, and failure-and-recovery evidence under a frozen model, adapter, analyzer, hardware, mutation, and seed manifest. Experiential feasibility would require authorized H1 records showing that the procedure and its separated response measures can be completed with acceptable burden; it would not establish efficacy. Efficacy would require a materially different design capable of supporting a causal outcome claim, while a mechanism claim would require evidence identifying how an effect arose rather than merely observing change. Neither efficacy nor mechanism is tested here.

These arguments establish only the design-level boundary described by RQ1. Formal equations and a coherent record model cannot by themselves answer RQ2 or RQ3. The contribution would need revision if the record chain proves too burdensome to use, if its provenance cannot expose where material mismatches arise, if model adapters cannot preserve the approved semantic distinctions, or if simpler records support equally accountable evaluation. Inspectability is consequently a falsifiable design commitment, not an empirical success already achieved.

6.2 Moment-specific and negotiated personalization

The person–moment unit resists treating personalization as retrieval of a fixed profile or discovery of an objectively correct sound. A person may want to stay with, explore, soften, or close an experience at one time and make a different choice later. Antidote therefore represents prior evidence as eligible context rather than destiny, a desired journey as a proposal rather than a promised endpoint, and semantic language as person-extensible rather than as a universal affect vocabulary. “Moment-specific” describes this conditioning boundary; it does not imply that the system knows what the person needs.

“Negotiated” identifies an authority process rather than an optimization result. Within the proposed framework, person-authored intent, an inspectable plan, adapter downgrades, and any evidence-linked update proposal remain distinct, reviewable records; edits create new revisions, and approvals or rejections are recorded where the applicable contract permits them. Model uncertainty is not converted into authority by being presented as personalization. This process may make the system’s assumptions easier to contest, but additional interaction may itself create burden, demand characteristics, or reduced trust. An accepted plan is thus evidence of a recorded choice under the declared interface, not evidence that the plan represents a stable preference, was freely understood, or will be helpful.

The response model must be equally conditional. Musical responses may involve memory, expectancy, imagery, appraisal, cultural associations, and other pathways that are not recoverable from one acoustic feature or affect label [7]. Perceived expression and felt emotion can differ [8], while intensity may be welcome, unwanted, mixed, or harmful in ways that require the person’s interpretation [9]. The same realized artifact may therefore receive different accounts in different moments without one being treated as noise merely because it conflicts with an earlier record.

The frozen H1 design can eventually examine whether those distinctions can be collected with acceptable completeness and burden, but its P condition is a bundle: it changes contextual input, personal semantic content, plan visibility, editing, and approval together. A P-versus-S or P-versus-G contrast cannot identify which component mattered, and a favorable response cannot validate an objective need estimate. Even repeated provenance-linked records would support only a bounded within-person account unless a later protocol prospectively tests whether a named advisory relationship predicts new observations. Until then, moment-conditioned personalization remains a structured hypothesis and a human-authority requirement, not an observed advantage.

6.3 Alternative explanations and architectures

Any later favorable response would admit explanations other than successful moment-specific adaptation. Generic ambient audio might suffice; a current-state category with a fixed journey template might capture most of the difference; or preference, familiarity, novelty, expectancy, attention, editing effort, and interface-mediated agency might shape the report. Output quality, loudness, setting, time of day, baseline variation, order, or carryover could also confound the intended semantic manipulation. Repeated use may teach instrument use or reporting rather than reveal a stable audio–response relationship. The G, S, and P conditions preserve an initial comparison among interface bundles, but cannot isolate components or mechanisms.

Contradictory evidence makes those alternatives substantive rather than ceremonial. Prior systems demonstrate technically closed loops and person-specific or affective control, yet their evaluations are small, heterogeneous, or directed at different constructs. A recent preprint of a preliminary brain–computer musical-interface evaluation reported nonsignificant target-emotion and time effects, with person and trial variation dominating its observed variance [6]. A systematic review likewise found heterogeneous music-neurofeedback methods and no consensus success measure [13]; a scoping review of commercial music-based digital therapeutics found insufficient direct evidence for real-world effectiveness

or superiority over generic music [14]. These sources do not show that an individual effect is impossible. They do show why loop closure, market presence, and a beneficial response cannot substitute for one another.

The proposed architecture is also only one arrangement of the technical problem, and four simpler baselines deserve direct comparison. A population affect-label pipeline could map a small shared state vocabulary to fixed audio rules without personal context. A static recommender could select or sequence existing recordings, as prior current-to-target systems do [1, 10]. Direct physiological conditioning could drive continuous controls without a negotiated semantic layer, following earlier brain–computer music systems [2, 3]. A one-shot generator could render the complete journey offline for human review, avoiding a future-audio buffer and mid-session model deadline altogether. Interpretable rule-based and probabilistic generation [4] and human-authored segments with a fixed transition engine provide further alternatives. These arrangements may reduce latency, unsupported controls, interaction burden, or transition uncertainty at the cost of a smaller or less inspectable generation space. Antidote’s two-rate proposal should be preferred only if explicit deadlines, verification, and fallbacks measurably improve the declared technical properties without obscuring person authority.

Several outcomes would count against the proposed arrangement: adapters that frequently downgrade accepted intent, a buffer margin that cannot be maintained on the frozen hardware, fallbacks that dominate listening, semantic or acoustic boundary measures that do not correspond to reviewable failures, or an authority path that does not halt reliably. Even technical success would leave the response question open. Small boundary distances, accurate tempo, or an uninterrupted file cannot establish perceived continuity, fit, usefulness, safety, or benefit. Alternative architectures should therefore be compared under the same provenance, exposure, missingness, burden, and non-positive reporting rules rather than by selecting whichever system produces the most persuasive demonstration.

6.4 Bounded future research program

The next evidence should be acquired within the stages and gates already frozen by the protocol rather than by expanding the interpretation of the present record. At D0, the immediate question is whether every RQ1 construct and governed transformation or authority boundary resolves to a versioned record, equation, contract, or explicit unavailable state; satisfying that traceability check would not constitute an independent implementation audit. At T0, a qualifying synthetic package would have to preserve complete denominators and demonstrate the declared authority, schema, replay, interruption, recovery, hash, and provenance invariants. Only an explicit claim-ledger promotion could turn that package into a synthetic technical result, and no T0 result could describe a human exposure or response.

T1 is the next test of the proposed generation path. Before execution it would freeze the model, adapter, hardware profile, analyzer set, mutation manifest, seed policy, output-rights review, hardware- and adapter-specific thresholds, and run plan. Its 39-attempt matrix, when every declared control is supported, would test one-factor adherence, adjacent-segment continuity, latency and buffer margins, and injected failure and recovery. A technically feasible claim would require the reserved source package, complete terminal-state accounting, reproducible analysis, and review against those frozen thresholds. Any resulting claim would remain bounded to that manifest. Unsupported controls would remain explicitly unsupported and reduce the applicable attempt denominator; missed deadlines against frozen thresholds, integrity failures, or unreliable stopping would require the corresponding technical claim to be downgraded or rejected, or the design to be revised. Passing T1 would remain specific to the tested model and hardware and would not establish perceived continuity or a beneficial response.

Optional real-time sensing and genuinely continuous generation are separate later studies, not capabilities assumed by H1. Every sensor stream would need its own consent, calibration, missingness, uncertainty, synchronization, and person-correction contract; its estimate could tailor a proposal but could not define

the person’s state. A streaming generator would likewise need a new qualified adapter and deadline study demonstrating that variable-latency model work remains outside the audio callback and that stop does not depend on model completion. Failure at either gate should default to manual input and, only after its own implementation and qualification, generate-ahead playback; it must not silently strengthen inference or timing claims.

H1 remains **blocked-no-collection-authority**: all eight activation gates—real-model/T1 qualification, privacy, consent, independent review, assignment, instrument, analysis, and pilot stop/go—remain unsatisfied. If separately authorized, its six-block G/S/P series targets 18 completed exposures within an absolute ceiling of 24 scheduled attempts and can test whether the governed records are collectable, correctable, provenance-linked, and usable with the prespecified completeness, burden, and safety-event accounting. It can describe exploratory within-person bundle contrasts while preserving missing, null, opposite-direction, interrupted, burdensome, harmful, and adverse observations. It cannot isolate personalization components, estimate a population effect, establish safety, or answer whether advisory learning is useful. Its progression rule decides only whether another bounded study may be acceptable; “proceed” is not an outcome claim.

Testing RQ3 would consequently require a new, prospectively frozen protocol after H1, not reuse of the H1 observations both to discover and confirm a preferred relationship. Such a protocol would need to name the eligible semantic, acoustic, exposure, and response variables; specify the response model, uncertainty, forgetting and contradiction rules; separate model development from evaluation on later observations; and retain human review of every proposed update. A useful advisory claim would require prospective evidence that a frozen proposal rule adds information for that person’s new records relative to declared non-adaptive alternatives, while retaining burden, mismatch, harm, and missingness. Failure to improve prospective prediction, or improvement purchased through unacceptable burden or risk, would count against the claim.

Only after those within-person gates should multi-participant work be considered. A later study would require a separate ethics determination, participant and setting rationale, sample-size justification, device and model qualification, accessible procedures, prespecified comparison conditions, and an analysis that distinguishes person variation from moment, order, and trial variation. It would also need to test rather than assume that the semantic vocabulary, control mappings, and response measures transfer across people and contexts. No stage described here is a treatment or neurological-mechanism evaluation. A claim may be promoted only when its named source package and review gate support that promotion; otherwise the relevant slot remains unavailable or the candidate remains unchanged, is downgraded, or is rejected. The unresolved validity, emotional and hearing safety, privacy, rights, accessibility, and independent-review gates in Section 7 constrain every future direction described here; publication of the design cannot satisfy them.

7 Limitations

This manuscript establishes a design and a prospective feasibility protocol, not treatment efficacy, clinical safety, successful affect regulation, learnability of a personal mapping, universal audio–state relationships, or a neurological mechanism. Its executable evidence is presently limited to a deterministic synthetic session and publication checks at a named source revision. The real-model, two-rate rendering, private-data, and formal-study paths remain incomplete or blocked. A requirement stated in an architecture or protocol is therefore classified below as *specified*, not as an effective mitigation. Table 7 summarizes the current controls beside their residual risks and blocking gates.

7.1 Construct and measurement validity

The moment representation is a bounded account produced through a particular interface; it is not the person's true psychological or physiological state. Self-report can preserve experience that an external classifier cannot observe, but wording, timing, scale anchors, recall, current arousal, and the wish to make the session coherent may all shape a response. Compact affect scales provide a measurement precedent [39], yet that precedent does not validate Antidote's desired-transition, mismatch, usefulness, harm, burden, or aftereffect items. Likewise, optional physiology or behavior would be an uncertain observation rather than an objective need: passive-monitoring research reports substantial heterogeneity and missing-data concerns, and electrodermal activity does not uniquely identify felt valence [28, 29].

Semantic intent, requested acoustic controls, measured realization, perceived expression, felt response, immediate usefulness or harm, and later aftereffect are kept in separate records. This reduces category errors; it does not prove that any record validly or reliably measures its named construct. Perceived and felt musical emotion may diverge [8], and an analyzer's tempo, key, loudness, or embedding estimate cannot decide what the audio meant to the listener. Person correction and field-level missingness are consequently necessary controls, but their usability, responsiveness, test-retest behavior, and sensitivity to meaningful within-person change remain unvalidated.

The originating cathartic experience is one retrospective, researcher-authored observation without a protocol, comparator, standardized measure, or independent interpretation. It motivates the person-and-moment question but is not a Result, an effect estimate, or evidence that emotional intensity was helpful. The framework's deliberate separation of intensity, welcome or unwanted response, harm, and later meaning must survive instrument review before any qualifying human observation can support even a feasibility claim.

7.2 Internal and statistical validity

The prospective H1 series is an unblinded self-study in which researcher, developer, participant, analyst, and author roles overlap. That overlap creates expectancy, demand, selective-attention, analytic-flexibility, and role-conflict risks that provenance cannot remove. The generic, state-aware, and negotiated personal conditions are also visibly different bundles: personal context, semantic content, plan visibility, editing, and approval change together. Neither a pairwise contrast nor exact randomization can isolate which component produced a difference, and no contrast identifies a psychological or neurological mechanism.

Baseline fluctuation, time of day, setting, prior activity, familiarity, regression toward a person's usual state, repeated instrument use, novelty, order, carryover, and learning may all covary with condition. The frozen balanced schedule and within-block seed policy constrain some researcher degrees of freedom, while near-event assessment can reduce retrospective delay [38]. They do not create blinding, washout, exchangeability, or a stable underlying affect process. A stopped or missing response may also be informative rather than random, particularly when discontinuation follows distress, burden, technical failure, or mismatch.

The target of 18 completed exposures under an absolute ceiling of 24 attempts is appropriate only to the declared record-feasibility questions. The series cannot estimate a population effect or support a clinical, safety, superiority, or causal claim. SPENT and CENT provide useful protocol and complete-reporting precedents for N-of-1 work [35, 40]; checklist alignment does not repair weak measurement, multiplicity, bundled conditions, carryover, or investigator conflict. Preserving every assignment, failure, exposure, null or opposite-direction observation, adverse response, denominator, missingness reason, and prespecified analysis is a transparency control. It is not a guarantee of conclusion validity.

7.3 External and ecological validity

Any later H1 observation would describe one person under one frozen protocol, model, adapter, device path, listening environment, semantic vocabulary, and period of life. It would not establish transfer across people, moments, languages, cultures, musical histories, sensory tolerances, hearing profiles, devices, generators, or clinical contexts. Even within one person, a mapping may change after familiarity, major events, medication or substance changes, sleep disruption, illness, or a shift in what the same semantic term means. Versioning can locate such change; it cannot make the relationship stationary.

The proposed local, self-guided setting is not interchangeable with a music therapy relationship, laboratory procedure, guided psychedelic setting, or ordinary music listening. The presence of headphones, an eye mask, a check-in, or a structured journey may itself alter attention and expectancy. Reviews of music-based digital therapeutics and music-neurofeedback systems report heterogeneous methods and limited grounds for real-world generalization [13, 14]. Those limitations neither prove that Antidote cannot work nor support transporting a favorable demonstration. A later multi-participant protocol would need a new sampling rationale, accessible procedure, device and setting qualification, and explicit tests of which constructs and controls transfer.

7.4 System, model, and continuity failures

Every transformation in the proposed chain can fail while producing a plausible-looking artifact. A context extractor may omit the relevant detail or privilege the wrong one; a state estimate may become confidently wrong; a semantic mixin may conflict with an exclusion; an adapter may ignore or downgrade a control; an analyzer may mischaracterize the output; and a generative model may reproduce, imitate, or introduce unwanted material. Human correction constrains authority but does not ensure that a person can identify every omission, understand every control, or review effectively while distressed.

The two-rate architecture is also prospective. The checked-in application can supervise a deterministic mock worker, verify a synthetic WAV, start deliberate playback, stop, recover, and record a response. It does not yet implement or qualify a real music model, uncertain state estimator, receding-horizon controller, verified future-audio buffer, or deterministic continuity renderer. Real generation can miss deadlines; supported controls may differ by adapter; crossfades can hide a waveform discontinuity while leaving harmony, rhythm, timbre, meaning, or pacing incoherent. Normative scheduling and loudness references help define tests [27, 31]; they do not establish seamlessness, perceived continuity, or an appropriate experience.

Model, code, analyzer, driver, and hardware drift may invalidate a prior result even when a prompt appears unchanged. Immutable specifications, capability negotiation, hashes, event replay, versioned analyzers, cancellation, fallbacks, and complete failure accounting can make some divergence detectable. Provenance records what the governed system says occurred; W3C PROV and workflow-run packaging do not prove semantic correctness, authenticity, consent, ownership, privacy, or benefit [33, 34]. T1 must therefore qualify control adherence, latency, buffer margin, continuity, integrity, interruption, and recovery on the exact real-model and hardware manifest before that path can inform H1. Fallbacks that dominate playback, unverifiable outputs, missed deadlines, or unreliable stopping count against feasibility rather than as invisible implementation details.

8 Ethics statement

No formal participant study, recruitment, human-subject data collection, ethics approval, exemption, or independent determination is reported. Protocol ANT-PROT-FEAS-001 remains **blocked-no-collection-authority**. Publication of this design and protocol cannot activate it.

Risk / layer	Current control	Residual risk / gate	Status
Construct and inference <i>state, intent, response</i>	Keep intended, realized, perceived, felt, usefulness-or-harm, and later fields separate; retain correction and missingness.	Measures remain unvalidated for this use, and optional passive signals can be ambiguous. <i>Owner:</i> measurement review. <i>Gate:</i> Instrument and construct-validity review before H1.	Specified
Internal and statistical validity <i>within-person comparison</i>	Freeze assignment, denominators, missingness, null and adverse retention, and the exact randomization analysis.	An unblinded self-study, bundled conditions, expectancy, order, carryover, and a short series prevent causal or population inference. <i>Owner:</i> methods and statistical review. <i>Gate:</i> H1 analysis review; new protocol for efficacy claims.	Specified
External and ecological validity <i>transfer and generalization</i>	Bound every claim to the tested person, moment, revision, model, device, and listening setting.	Transfer across people, cultures, languages, generators, devices, sensory needs, and care settings is unknown. <i>Owner:</i> future study governance. <i>Gate:</i> Separate multi-participant protocol after feasibility.	Unresolved
Model, timing, and continuity <i>technical realization</i>	Synthetic consent, hashes, cancellation, stop, and recovery are tested; real-model buffering and rendering checks are specified.	State estimates, adapters, analyzers, deadlines, transitions, and fallbacks are not qualified end to end. <i>Owner:</i> MVP and T1 qualification. <i>Gate:</i> Real-model and complete-session gates before H1.	Mixed
Emotional and hearing safety <i>exposure and adverse response</i>	Preserve playback and stop authority; record mismatch, harm, discomfort, and aftereffect separately; begin at low device level.	No control is shown to prevent overwhelm; at-ear level, cumulative dose, and crisis response remain unknown. <i>Owner:</i> safety and device review. <i>Gate:</i> Safety plan, device/dose qualification, and independent review.	Specified
Consent, privacy, and export <i>personal context and records</i>	Scoped grants and reviewed projection execute for synthetic manual context; research export remains fail-closed by contract.	Meaningful consent, encryption, keys, backup, retention, deletion, redaction, re-identification, and export are not privacy-qualified. <i>Owner:</i> privacy and data governance. <i>Gate:</i> Protected data-management and privacy review before H1.	Mixed
Rights, bias, access, and review <i>model selection and study authority</i>	Real-model selection, generated-audio release, and formal collection remain blocked pending documented review.	Licenses, training provenance, output rights, cultural bias, accessibility, misuse, and the applicable ethics determination are unresolved. <i>Owner:</i> maintainer and independent reviewers. <i>Gate:</i> Model, data, audio, accessibility, misuse, and ethics review.	Blocked

Status key. Mixed = only named synthetic behavior is implemented; Specified = required but unverified; Unresolved = no adequate control or evidence is registered; Blocked = the activity cannot proceed before the stated gate. Every row retains residual uncertainty.

Table 7: Governed risk register pairing current controls with residual uncertainty and the gate that blocks stronger use. Mixed denotes only bounded synthetic implementation; specified controls remain unverified, and formal collection remains blocked.

Table description: Seven rows cover construct and inference validity; internal and statistical validity; external and ecological validity; model, timing, and continuity failures; emotional and hearing safety; consent, privacy, and export; and rights, bias, accessibility, misuse, and independent review. Each row names the affected evidence layer, current control, residual risk, status word, owner, and blocking gate. Only some synthetic prototype behaviors are implemented. The table authorizes no collection and does not present provenance, local execution, consent records, or a planned safeguard as proof of truth, privacy, rights, accessibility, safety, or benefit.

Antidote is a research instrument, not a treatment, diagnostic, medical device, crisis service, or substitute for professional care.

8.1 Emotional safety and participant authority

Music can be unwanted or harmful as well as meaningful. A theoretical model of music-induced harm identifies affective, behavioral, cognitive, identity, interpersonal, physical, and spiritual dimensions and treats harm as an interaction among context, music, recipient, and delivery [9]. Antidote must therefore never optimize tears, intensity, duration, completion, engagement, or physiological change as a proxy for benefit. Crying, calm, activation, numbness, discomfort, catharsis, and distress require the person's interpretation in that moment and again, when appropriate, at the declared later interval.

The design preserves explicit playback, skip, stop, correction, grounding, and withdrawal authority and gives mismatch, unwanted intensity, harm, hearing discomfort, serious or unresolved events, and later aftereffect their own records. A stop must halt playback independently of model completion; a safety halt must block automatic continuation and personal-model updating. These are necessary requirements, not evidence that distress is prevented or resolved. The prototype provides no clinical monitoring, emergency response, or reliable way to distinguish tolerable challenge from escalating harm.

Before any human exposure, an independently reviewed safety plan must define eligibility and exclusions, pre-session information, low-burden stopping, grounding and follow-up, adverse-event classification, escalation contacts, unresolved-event review, and discontinuation rules. High-risk crisis use, unsupervised treatment claims, and guidance for medication or psychoactive-drug sessions remain outside v0 authority. Refusal, stopping, non-response, or an unfavorable account must not reduce access, be reframed as noncompliance, or be used as a negative reward signal.

8.2 Consent, privacy, and data governance

Consent is action- and purpose-specific. Permission to inspect one source does not authorize summarization, generation, playback, retention, learning, synchronization, research analysis, export, or publication. The current synthetic desktop path requires explicit confirmation for a manually entered, session-scoped context projection and keeps the worker outside unrestricted history. Historical journal or therapy-chat integration, revocation UI, retention enforcement, and privacy-ready payload handling are not implemented. An inspectable grant or projection also does not demonstrate comprehension, voluntariness, or freedom from demand—a particular concern when researcher and participant are the same person.

Local-first design can reduce required server transfer and keep an authoritative record available on the person's device [25]; it is not perfect security. The current store is not ready for private participant data. At-rest encryption, key creation and recovery, operating-system account separation, backup policy, retention enforcement, auditable deletion, tombstones, derived-projection invalidation, crash-log review, and a production worker sandbox remain unresolved. Device compromise, malware, cloud-synced folders, backups, screen capture, exports, and another local user can still expose sensitive context or response data.

Public provenance presents a second tension: detailed lineage aids audit but can increase linkability and re-identification risk. The reporting contract permits only privacy-reviewed aggregate public envelopes and keeps row-level human records, free text, private paths, stable identifiers, record hashes, and the private evidence index outside the repository. Even a reconciled aggregate commitment does not establish anonymity, consent, authenticity, truth, or completeness. A person must be able to decline retention and export separately; the exact deletion effect on events, payloads, projections, analyses, backups, and already shared artifacts must be specified before collection. No public release may proceed

solely because a schema validates or a hash matches.

8.3 Licensing, bias, accessibility, and hearing safety

No real generator has been selected or qualified for the study. Before a model or audio artifact can enter T1, H1, a release, or the magazine, review must cover the exact code and weight licenses, redistribution and commercial restrictions, training-data documentation and consent, required attribution, model-card limitations, remote-code execution, recognizable reproduction or imitation, third-party samples, output rights, and the intended distribution context. Generative-audio research has historically discussed positive implications far more often than negative ones and identifies copyright, agency, bias, privacy, misuse, labor, access, and resource concerns as material review categories [41]. That review does not prove that a particular model or output violates rights; conversely, an open weight license cannot clear its training data or every generated artifact.

Training corpora, control vocabularies, affect labels, and musical evaluation may privilege well-represented genres, languages, cultures, instruments, and listener assumptions. A person-extensible semantic layer may reduce pressure to adopt one vocabulary, but adapter mappings and model outputs can still flatten or appropriate meaning. The prototype has automated keyboard and semantic checks; it has not established usability with screen readers, reduced-motion or low-vision settings, hearing differences, motor constraints, cognitive load, trauma-related sensitivities, or users outside the developer context. More knobs may increase agency for one person and burden or suggestion effects for another. Accessibility and cultural review must precede recruitment rather than be inferred from a nominally local interface.

File-level integrated loudness and true-peak constraints can bound some digital-output errors but do not determine sound pressure at the ear. The WHO-ITU safe-listening standard treats exposure as level over duration and recommends dosimetry, accessible information, and volume-limiting options [42]. Antidote currently has no calibrated device path, at-ear measurement, cumulative-dose accounting, or proof that its lowest volume is appropriate. Human exposure therefore requires device-aware qualification, conservative defaults, continuously available volume and stop control, warnings that do not rely on sound alone, and a response path for tinnitus, pain, dizziness, or hearing discomfort. None of those controls alone establish hearing safety.

8.4 Independent review and collection gate

The project has not recorded an institutional review board, research ethics committee, or other independent determination, and this manuscript does not decide whether a future self-study is exempt or outside a particular regime. That determination depends on the final protocol, institution, jurisdiction, recruitment, data, claims, and intended use. Self-use, local execution, public source code, or the absence of payment does not by itself resolve research ethics. WHO health-AI guidance provides a relevant normative foundation for autonomy, well-being, transparency, accountability, inclusiveness, and responsiveness [43]; alignment with those principles is not an approval or safety certification.

Before H1 can change state, independent review must document the applicable ethics determination and approve or otherwise disposition the exact protocol, role conflict, consent process, eligibility and exclusion rules, emotional and hearing-safety plan, adverse-event and crisis escalation, compensation if any, privacy and data-management plan, retention and deletion behavior, public-export boundary, accessibility procedure, device qualification, model and output rights audit, analysis plan, and stopping rules. The separate real-model/T1, privacy, consent, independent-review, assignment, instrument, analysis, and pilot stop/go gates must each be satisfied by named evidence; no gate is inherited from the paper or another project. Any material amendment requires a new version and renewed review

before collection continues.

Until those records exist, formal collection, recruitment, therapeutic use, autonomous personalization, sensor inference, and public release of private or row-level response data remain prohibited. A later approval would be limited to its exact scope and would not prove benefit, privacy, accessibility, or safety.

9 Data and code availability

This article is a design/protocol manuscript. As of version 0.1.0, no formal participant dataset, qualifying T1 real-model package, or H1 within-person result package exists. The public repository does, however, contain the manuscript and an executable synthetic prototype. The prototype implements a rule-guided local session with a deterministic mock worker; it is not a real generative-audio model, a human-outcome dataset, or evidence of usefulness, safety, efficacy, or mechanism.

The canonical source is the [Antidote repository](#). The repository records the versioned manuscript, bibliography, literature atlas, source assessments, claim ledger, architecture decisions, protocol and reporting contracts, application code, JSON Schemas, synthetic fixtures, tests, and project-owned publication scripts. The source revision printed on the title page identifies the exact Git state from which each paper projection was built.

9.1 Public artifacts and reproducibility boundary

The canonical review surface is [antidote.ego hygiene.io](#). It exposes the [accessible web paper](#), the [PDF](#), and a [download index](#). The download index links the independently compilable, arXiv-oriented source archive; the publication manifest; the site catalog; build and site-suite provenance; and the complete SHA-256 inventory. The source archive is a packaging target, not an arXiv submission. No DOI, arXiv identifier, archival release, or peer-reviewed status is asserted for this preprint.

For each deployed revision, `publication.json` binds the web, PDF, provenance, and source-archive artifacts to their Git revision and SHA-256 digests; `site.json` records the stable routes and their availability states; and `SHA256SUMS` covers the staged publication tree. These records can detect byte drift and trace an artifact to its declared build. They do not establish source correctness, scientific validity, authorship, consent, privacy, ownership, safety, or clinical efficacy.

The project-owned Make and Task entry points invoke the same build scripts. `make check-all` and `task check-all` build and validate the neutral and Ego Hygiene projections. The reproducibility target performs two clean builds and byte-compares the PDF, accessible HTML, provenance JSON, and arXiv-oriented source archive. Continuous integration repeats the publication checks on Ubuntu 24.04; the Beacon research-paper source and Holon site-suite inputs are pinned to immutable commits, while language package graphs use committed lockfiles. The operating-system packages installed by the publication workflow are not frozen as a complete container image, and cross-platform or perpetual byte identity is therefore not claimed. The documented developer baseline records Rust 1.97.1, Node 24.19.x, npm 11.19.x, Python 3.12.13, uv 0.11.33, and Tauri 2 for the synthetic application workspace; these pins do not constitute end-user platform or real-model support.

9.2 Excluded and access-controlled material

No private journal, therapy-chat, health, credential, or other personal content is included. No formal participant record has been collected under protocol ANT-PROT-FEAS-001, whose H1 stage remains `blocked-no-collection-authority`. Any later qualifying H1 material would require separately re-

viewed consent, security, retention, de-identification, and release decisions. The public reporting contract permits only reviewed aggregate envelopes and audit metadata; it does not promise row-level records, raw text, direct identifiers, or protected provenance payloads.

The repository does not distribute third-party article full text, training datasets, proprietary media, model weights, real-model packages, or generated audio. Source records provide citations, links, and project-authored assessments without transferring the rights of the underlying works. Future model code, weights, datasets, and audio outputs must retain their own verified licenses and output-rights records before use or release.

Licensing follows the artifact boundary rather than the repository location. Repository software, automation, and non-manuscript documentation are provided under the root MIT license. The reviewable manuscript source and manuscript figures remain copyright Alan Szmyt, all rights reserved, under `paper/LICENSE.md`; public access does not grant a reuse license. Linked or quoted third-party works retain their original rights. The root MIT license does not relicense the manuscript, source literature, model assets, training data, or future generated audio. A later archival release or venue submission may select a different manuscript license through an explicit reviewed change.

10 Acknowledgements

This work received no external funding and was not produced under institutional sponsorship. Ego Hygiene is the author's independent project affiliation; its appearance does not imply university, employer, clinical, or other institutional endorsement. No additional individual acknowledgement is declared for version 0.1.0.

Generative AI tools, including OpenAI ChatGPT and Codex, assisted with literature discovery, code and document drafting, editing, test construction, and visual ideation. Alan Szmyt selected the research direction, reviewed and revised the claims and artifacts, approved the manuscript, and remains accountable for the work. AI systems are not authors or contributors under the CRediT taxonomy. Vendor-hosted model revisions and the complete interaction history are not frozen in the repository, so exact regeneration of AI-assisted intermediate drafts is not claimed.

11 Contributor statement

Alan Szmyt is the sole human author. Contributor roles were reviewed against the CRediT taxonomy and are recorded in `beacon-project.toml`. **Alan Szmyt:** Conceptualization; Data curation; Investigation; Methodology; Project administration; Software; Validation; Visualization; Writing—original draft; Writing—review and editing.

Alan Szmyt originated and maintains Antidote and Ego Hygiene and has a direct intellectual and project-role interest in the proposed system. The prospective H1 protocol is a self-study in which researcher, developer, participant, analyst, and author roles would overlap; no H1 collection is authorized, and independent review of that role conflict remains an entry gate. No external commercial sponsorship or other financial competing interest is declared for version 0.1.0. The author reviewed and approved the manuscript and accepts responsibility for its content.

12 Conclusion

Antidote asks how a local research instrument can represent, generate, and evaluate an audio journey for one person in one moment without collapsing the records that give the journey its meaning or evidentiary status. At the design level, RQ1 is answered by treating consented source context, a reviewed working projection, explicit semantic intent, an accepted time-indexed plan, adapter-specific conditioning, measured acoustic realization, actual exposure, separated response observations, and any later advisory update as distinct, versioned, provenance-linked objects. Human approval, correction, interruption, and stop authority remain explicit transitions throughout that chain. This answer specifies what an inspectable instrument must preserve; it does not show that the instrument correctly understands a person or produces a useful, continuous, safe, or beneficial experience.

The manuscript makes five bounded contributions at different evidence levels. It situates the problem among prior personalized, current-to-target, closed-loop, controllable, physiological, and longitudinal music systems rather than claiming those ingredients as new. It formalizes the end-to-end inspectability contract and the non-interchangeability of intended controls, realized acoustics, perceived expression, felt response, immediate usefulness or harm, and later aftereffect. It proposes a two-rate architecture in which person-correctable deliberation, typed semantic mixins, short-horizon planning, and variable-latency generation remain outside deterministic playback and stop deadlines. It defines a local-first human-authority model, with only a bounded synthetic path presently implemented. Finally, it freezes a prospective staged feasibility and reporting protocol that preserves null, negative, missing, interrupted, burdensome, and adverse records rather than manufacturing a favorable Results narrative.

RQ2 and RQ3 remain unanswered. No qualifying real-model technical package or formal human-study package exists, and collection authority remains false. RQ2 requires promoted synthetic conformance evidence followed by a frozen real-model, adapter, analyzer, hardware, timing, continuity, integrity, and recovery qualification. The record-feasibility prerequisite to RQ3 additionally requires privacy, consent, independent-review, assignment, instrument, analysis, and pilot stop/go gates. A useful advisory relationship would then require a later prospectively frozen evaluation on new within-person records; it cannot be inferred from the planned H1 series. Antidote’s central proposition is therefore not that a system knows what a person needs, but that a person can remain the authority over a moment-specific audio hypothesis whose transformations are visible enough to inspect, challenge, and study.

References

- [1] Joris H. Janssen, Egon L. van den Broek, and Joyce H. D. M. Westerink. Tune in to your emotions: A robust personalized affective music player. *User Modeling and User-Adapted Interaction*, 22(3):255–279, 2012. doi: 10.1007/s11257-011-9107-7. URL <https://doi.org/10.1007/s11257-011-9107-7>.
- [2] Ian Daly, Duncan Williams, Alexis Kirke, James Weaver, Asad Malik, Faustina Hwang, Eduardo Miranda, and Slawomir J. Nasuto. Affective brain–computer music interfacing. *Journal of Neural Engineering*, 13(4):046022, 2016. doi: 10.1088/1741-2560/13/4/046022. URL <https://doi.org/10.1088/1741-2560/13/4/046022>.
- [3] Stefan K. Ehrlich, Kat R. Agres, Cuntai Guan, and Gordon Cheng. A closed-loop, music-based brain–computer interface for emotion mediation. *PLOS ONE*, 14(3):e0213516, 2019. doi: 10.1371/journal.pone.0213516. URL <https://doi.org/10.1371/journal.pone.0213516>.
- [4] Kat R. Agres, Adyasha Dash, and Phoebe Chua. AffectMachine-Classical: A novel system for

- generating affective classical music. *Frontiers in Psychology*, 14:1158172, 2023. doi: 10.3389/fpsyg.2023.1158172. URL <https://doi.org/10.3389/fpsyg.2023.1158172>.
- [5] Yimeng Zhang, Yueru Sun, Haoyu Gu, and Zhanpeng Jin. MindMelody: A closed-loop EEG-driven system for personalized music intervention, 2026. URL <https://arxiv.org/abs/2605.01235>. Version 2, revised 2026-05-16.
 - [6] Pablo A. Monroy-D’Croz, Rafael Ramirez-Melendez, and Julian Cespedes-Guevara. A minimalist brain–computer musical interface for real-time emotion-driven sonification: System design and preliminary evaluation, 2026. URL <https://arxiv.org/abs/2606.01473>. Version 1.
 - [7] Patrik N. Juslin and Daniel Västfjäll. Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31(5):559–575, 2008. doi: 10.1017/S0140525X08005293. URL <https://doi.org/10.1017/S0140525X08005293>.
 - [8] Marcel Zentner, Didier Grandjean, and Klaus R. Scherer. Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, 8(4):494–521, 2008. doi: 10.1037/1528-3542.8.4.494. URL <https://doi.org/10.1037/1528-3542.8.4.494>.
 - [9] Michael J. Silverman, Lori F. Gooding, and Olivia Yinger. It’s...complicated: A theoretical model of music-induced harm. *Journal of Music Therapy*, 57(3):251–281, 2020. doi: 10.1093/jmt/thaa008. URL <https://doi.org/10.1093/jmt/thaa008>.
 - [10] Adiel Mallik and Frank A. Russo. The effects of music and auditory beat stimulation on anxiety: A randomized clinical trial. *PLOS ONE*, 17(3):e0259312, 2022. doi: 10.1371/journal.pone.0259312. URL <https://doi.org/10.1371/journal.pone.0259312>.
 - [11] Jiyang Li, Wei Wang, Kratika Bhagtani, Yincheng Jin, and Zhanpeng Jin. BEAMERS: Brain-engaged, active music-based emotion regulation system, 2022. URL <https://arxiv.org/abs/2211.14609>. Version 1.
 - [12] Thomas C. T. Ho and Xiang Chen. iHeartLift: A closed loop system with bio-feedback that uses music tempo variability to improve heart rate variability. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1181–1184, 2011. doi: 10.1109/IEMBS.2011.6090277. URL <https://doi.org/10.1109/IEMBS.2011.6090277>.
 - [13] Alexandre Sayal, Bruno Direito, Teresa Sousa, Neomi Singer, and Miguel Castelo-Branco. Music in the loop: A systematic review of current neurofeedback methodologies using music. *Frontiers in Neuroscience*, 19:1515377, 2025. doi: 10.3389/fnins.2025.1515377. URL <https://doi.org/10.3389/fnins.2025.1515377>.
 - [14] Tara Venkatesan, Andrew M. Demetriou, Audrey Hempel, and Daniel L. Bowling. A scoping review of music-based digital therapeutics for stress, anxiety, and depression. *Frontiers in Human Neuroscience*, 20:1602004, 2026. doi: 10.3389/fnhum.2026.1602004. URL <https://doi.org/10.3389/fnhum.2026.1602004>.
 - [15] Jan Melechovsky, Zixun Guo, Deepanway Ghosal, Navonil Majumder, Dorien Herremans, and Soujanya Poria. Mustango: Toward controllable text-to-music generation. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 8293–8316. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.naacl-long.459. URL <https://doi.org/10.18653/v1/2024.naacl-long.459>.

- [16] Or Tal, Alon Ziv, Itai Gat, Felix Kreuk, and Yossi Adi. Joint audio and symbolic conditioning for temporally controlled text-to-music generation, 2024. URL <https://arxiv.org/abs/2406.10970>. Version 1.
- [17] Patrick E. Hutchings and Jon McCormack. Adaptive music composition for games. *IEEE Transactions on Games*, 12(3):270–280, 2020. doi: 10.1109/TG.2019.2921979. URL <https://doi.org/10.1109/TG.2019.2921979>.
- [18] Inbal Nahum-Shani, Shawna N. Smith, Bonnie J. Spring, Linda M. Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A. Murphy. Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine*, 52(6):446–462, 2018. doi: 10.1007/s12160-016-9830-8. URL <https://doi.org/10.1007/s12160-016-9830-8>.
- [19] Carlos E. García, David M. Prett, and Manfred Morari. Model predictive control: Theory and practice—a survey. *Automatica*, 25(3):335–348, 1989. doi: 10.1016/0005-1098(89)90002-2. URL [https://doi.org/10.1016/0005-1098\(89\)90002-2](https://doi.org/10.1016/0005-1098(89)90002-2).
- [20] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2):99–134, 1998. doi: 10.1016/S0004-3702(98)00023-X. URL [https://doi.org/10.1016/S0004-3702\(98\)00023-X](https://doi.org/10.1016/S0004-3702(98)00023-X).
- [21] Saleema Amershi, Maya Cakmak, W. Bradley Knox, and Todd Kulesza. Power to the people: The role of humans in interactive machine learning. *AI Magazine*, 35(4):105–120, 2014. doi: 10.1609/aimag.v35i4.2513. URL <https://doi.org/10.1609/aimag.v35i4.2513>.
- [22] Ruth Maria Ingendoh, Ella S. Posny, and Angela Heine. Binaural beats to entrain the brain? a systematic review of the effects of binaural beat stimulation on brain oscillatory activity, and the implications for psychological research and intervention. *PLOS ONE*, 18(5):e0286023, 2023. doi: 10.1371/journal.pone.0286023. URL <https://doi.org/10.1371/journal.pone.0286023>.
- [23] Mendel Kaelen, Bruna Giribaldi, Jordan Raine, Lisa Evans, Christopher Timmermann, Natalie Rodriguez, Leor Roseman, Amanda Feilding, David Nutt, and Robin Carhart-Harris. The hidden therapist: Evidence for a central role of music in psychedelic therapy. *Psychopharmacology*, 235(2):505–519, 2018. doi: 10.1007/s00213-017-4820-5. URL <https://doi.org/10.1007/s00213-017-4820-5>.
- [24] T. Rowe, T. Hurzeler, E. Towers, E. Louie, and K. C. Morley. Psychedelic therapy and the role of music: A scoping review of quantitative evidence on subjective and objective outcomes. *Brain and Behavior*, 16(6):e71533, 2026. doi: 10.1002/brb3.71533. URL <https://doi.org/10.1002/brb3.71533>.
- [25] Martin Kleppmann, Adam Wiggins, Peter van Hardenberg, and Mark McGranaghan. Local-first software: You own your data, in spite of the cloud. In *Proceedings of the 2019 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software*, pages 154–178. Association for Computing Machinery, 2019. doi: 10.1145/3359591.3359737. URL <https://doi.org/10.1145/3359591.3359737>.
- [26] Yin Lin, Elaine Ang, Erkang Zhu, Bolin Ding, and Jingren Zhou. Context as an environment: Programmatic context management for long-horizon agents, 2026. URL <https://arxiv.org/abs/2608.21690>. Version 1.
- [27] World Wide Web Consortium. Web audio api 1.1. W3C First Public Working Draft, 2024. URL <https://www.w3.org/TR/webaudio-1.1/>. 5 November 2024; latest published version checked 2026-08-29.

- [28] Valeria De Angel, Serena Lewis, Katie White, Carolin Oetzmman, Daniel Leightley, Emanuela Oprea, Grace Lavelle, Faith Matcham, Alice Pace, David C. Mohr, Richard Dobson, and Matthew Hotopf. Digital health tools for the passive monitoring of depression: A systematic review of methods. *npj Digital Medicine*, 5:3, 2022. doi: 10.1038/s41746-021-00548-8. URL <https://doi.org/10.1038/s41746-021-00548-8>.
- [29] Tomás Ariel D’Amelio, Lorenzo Ariel Galán, Emmanuel Alesandro Maldonado, Agustín Ariel Díaz Barquinero, Jerónimo Rodríguez Cuello, Nicolás Marcelo Bruno, Enzo Tagliazucchi, and Denis Alexander Engemann. Emotion recognition systems with electrodermal activity: From affective science to affective computing. *Neurocomputing*, 651:130831, 2025. doi: 10.1016/j.neucom.2025.130831. URL <https://doi.org/10.1016/j.neucom.2025.130831>.
- [30] Baisen Wang, Chenxi Bao, and Qisong Han. Real-time interactive music generation via data-free streaming consistency distillation, 2026. URL <https://arxiv.org/abs/2606.24307>. Version 1.
- [31] European Broadcasting Union. EBU R 128: Loudness normalisation and permitted maximum level of audio signals. Recommendation R 128, version 5.0, 2023. URL <https://tech.ebu.ch/publications/r128>.
- [32] Donald R. Honeycutt, Mahsan Nourani, and Eric D. Ragan. Soliciting human-in-the-loop user feedback for interactive machine learning reduces user trust and impressions of model accuracy. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, 8(1):63–72, 2020. doi: 10.1609/hcomp.v8i1.7464. URL <https://doi.org/10.1609/hcomp.v8i1.7464>.
- [33] Luc Moreau and Paolo Missier. PROV-DM: The PROV data model. W3C Recommendation, 2013. URL <https://www.w3.org/TR/prov-dm/>. 30 April 2013.
- [34] Simone Leo, Michael R. Crusoe, Laura Rodríguez-Navas, Raúl Sirvent, Alexander Kanitz, Paul De Geest, Rudolf Wittner, Luca Pireddu, Daniel Garijo, José M. Fernández, Iacopo Colonnelli, Matej Gallo, Tazro Ohta, Hirotaka Suetake, Salvador Capella-Gutierrez, Renske de Wit, Bruno P. Kinoshita, and Stian Soiland-Reyes. Recording provenance of workflow runs with RO-Crate. *PLOS ONE*, 19(9):e0309210, 2024. doi: 10.1371/journal.pone.0309210. URL <https://doi.org/10.1371/journal.pone.0309210>.
- [35] Antony J. Porcino, Larissa Shamseer, An-Wen Chan, Richard L. Kravitz, Aaron Orkin, Salima Punja, Philippe Ravaud, Christopher H. Schmid, Sunita Vohra, and SPENT Group. SPIRIT extension and elaboration for n-of-1 trials: SPENT 2019 checklist. *BMJ*, 368:m122, 2020. doi: 10.1136/bmj.m122. URL <https://doi.org/10.1136/bmj.m122>.
- [36] Sandra M. Eldridge, Claire L. Chan, Michael J. Campbell, Christine M. Bond, Sally Hopewell, Lehana Thabane, Gillian A. Lancaster, and PAFS Consensus Group. CONSORT 2010 statement: Extension to randomised pilot and feasibility trials. *BMJ*, 355:i5239, 2016. doi: 10.1136/bmj.i5239. URL <https://doi.org/10.1136/bmj.i5239>.
- [37] Stefan Konigorski, Sarah Wernicke, Tamara Slosarek, Alexander M. Zenner, Nils Strelow, Darius F. Ruether, Florian Henschel, Manisha Manaswini, Fabian Pottbäcker, Jonathan A. Edelman, Babajide Owoyele, Matteo Danieleto, Eddy Golden, Micol Zweig, Girish N. Nadkarni, and Erwin Böttinger. StudyU: A platform for designing and conducting innovative digital n-of-1 trials. *Journal of Medical Internet Research*, 24(7):e35884, 2022. doi: 10.2196/35884. URL <https://doi.org/10.2196/35884>.
- [38] Saul Shiffman, Arthur A. Stone, and Michael R. Hufford. Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4:1–32, 2008. doi: 10.1146/annurev.clinpsy.3.022806.091415. URL <https://doi.org/10.1146/annurev.clinpsy.3.022806.091415>.

- [39] Alberto Betella and Paul F. M. J. Verschure. The affective slider: A digital self-assessment scale for the measurement of human emotions. *PLOS ONE*, 11(2):e0148037, 2016. doi: 10.1371/journal.pone.0148037. URL <https://doi.org/10.1371/journal.pone.0148037>.
- [40] Sunita Vohra, Larissa Shamseer, Margaret Sampson, Cecilia Bukutu, Christopher H. Schmid, Robyn Tate, Jane Nikles, Deborah R. Zucker, Richard Kravitz, Gordon Guyatt, Douglas G. Altman, David Moher, and CENT Group. CONSORT extension for reporting n-of-1 trials (CENT) 2015 statement. *BMJ*, 350:h1738, 2015. doi: 10.1136/bmj.h1738. URL <https://doi.org/10.1136/bmj.h1738>.
- [41] Julia Barnett. The ethical implications of generative audio models: A systematic literature review. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, pages 146–161, 2023. doi: 10.1145/3600211.3604686. URL <https://doi.org/10.1145/3600211.3604686>.
- [42] World Health Organization and International Telecommunication Union. Safe listening devices and systems: A WHO–ITU standard. Technical report, World Health Organization and International Telecommunication Union, Geneva, 2019. URL <https://www.who.int/publications/i/item/9789241515276>.
- [43] World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Technical report, World Health Organization, Geneva, 2021. URL <https://www.who.int/publications/i/item/9789240029200>.

A Notation and symbol glossary

This glossary is canonical for the first design/protocol manuscript. Units are explicit because tuples and typed records are not silently added as numerical vectors.

- i — Individual identifier; never a fixed person feature vector. *Unit or type: index.*
- t — Decision point within or between sessions. *Unit or type: decision index.*
- n — Journey stage or adjacent generated-audio segment. *Unit or type: stage index.*
- $z_{i,t}$ — Consent-bounded moment information available to the system. *Unit or type: heterogeneous record.*
- $o_{i,t}^{\text{usr}}$ — Explicit person-authored check-in and correction observations. *Unit or type: observation record.*
- $o_{i,t}^{\text{opt}}$ — Optional separately consented measurements or derived observations. *Unit or type: observation record.*
- $g_{i,t}$ — Desired direction or time-indexed journey proposal, not a guaranteed state. *Unit or type: journey representation.*
- $\mathcal{H}_{i,<t}$ — Prior evidence eligible under current retention, consent, and relevance review. *Unit or type: record set.*
- $S_{i,t}^{\text{ctx}}$ — Context state separating source events, projections, and provenance. *Unit or type: record tuple.*
- $L_{i,t}$ — Consented lossless source-event record. *Unit or type: event set.*

- $P_{i,t}$ — Replaceable derived context views. *Unit or type: projection set.*
- $V_{i,t}$ — Lineage and version information for context derivations. *Unit or type: version and provenance set.*
- $A_{i,t}$ — Active source-, purpose-, action-, time-, and retention-scoped consent grant. *Unit or type: authority set.*
- $c_{i,t}^{\text{work}}$ — Reviewed working context permitted for the current purpose. *Unit or type: semantic record.*
- π — Consent-constrained derivation from retained context to a working view. *Unit or type: projection operator.*
- m_j — Versioned semantic mixin with purpose, bounds, neutral position, and exclusions. *Unit or type: typed control.*
- $k_{j,t}^{\text{usr}}, k_{j,t}^{\text{acc}}, k_j^0$ — Current user value, explicitly accepted suggestion, and neutral default. *Unit or type: dimensionless in $[0,1]$.*
- $\alpha_{j,t}$ — Smoothed effective mixin authority for an ordinary non-safety control. *Unit or type: dimensionless in $[0,1]$.*
- ρ_j — Visible mixin-specific smoothing rate. *Unit or type: dimensionless in $(0,1]$.*
- $q_{i,n}$ — Resolved, model-independent semantic condition for stage n. *Unit or type: typed semantic condition.*
- $\oplus$ — Versioned semantic composition; not numerical addition or text concatenation. *Unit or type: typed composition operator.*
- $\Xi_{i,t}$ — Active exclusions, safety constraints, and prohibited transitions. *Unit or type: constraint set.*
- $\prec_{i,t}$ — Visible conflict-resolution and authority order. *Unit or type: precedence relation.*
- $J_{i,t}$ — Accepted model-independent stages, semantic intent, and acoustic controls. *Unit or type: journey-plan record.*
- $\mathcal{K}_a$ — Controls and limits declared by model adapter a. *Unit or type: capability set.*
- $\gamma_{i,n}^{(a)}$ — Adapter-bounded conditioning state with uncertainty, conflicts, and provenance. *Unit or type: conditioning record.*
- $\varphi_{i,n}^{(a)}$ — Optional prompt projection of the conditioning state; not canonical intent. *Unit or type: adapter-native text or tokens.*
- $b_{i,t}(x)$ — Uncertain system belief assigned to candidate represented state x. *Unit or type: probability.*
- $U, u_{i,t}$ — Candidate horizon and one semantic or musical control proposal. *Unit or type: control sequence and control record.*
- H_p — Finite planning horizon. *Unit or type: decision steps.*
- $\bar{L}_g, \bar{D}_\Delta, \bar{R}, \bar{C}_{\text{burden}}$ — Journey mismatch, control change, risk or uncertainty, and interaction burden terms. *Unit or type: normalized dimensionless costs.*

- $B_{i,t}$ — Duration of verified playable audio already buffered. *Unit or type: seconds.*
- $\ell_{\text{observe}}, \ell_{\text{plan}}, \ell_{\text{generate}}, \ell_{\text{verify}}, \ell_{\text{schedule}}$ — Measured component latencies along the generate-ahead path. *Unit or type: seconds.*
- $M_{i,t}$ — Hardware- and adapter-specific latency uncertainty margin. *Unit or type: seconds.*
- $D_n^{\text{sem}}, D_n^{\text{ac}}$ — Semantic-plan and measured acoustic-boundary distances. *Unit or type: dimensionless.*
- $\mathcal{F}, f_r, d_r, w_r$ — Versioned components of the acoustic-boundary comparison. *Unit or type: feature set, extractor, normalized distance, and dimensionless weight.*
- $a_n(s), y(s)$ — Adjacent input signals and renderer output over an overlap window. *Unit or type: sample amplitude.*
- $e_{i,t}$ — What content was actually played, for how long, and with what interruption state. *Unit or type: exposure record.*
- $Y_{i,t}$ — Separated perceived, felt, evaluative, harm, mismatch, and aftereffect observations. *Unit or type: heterogeneous response record.*
- θ_i — Versioned person-specific response-model hypothesis, if a later protocol fits one. *Unit or type: model parameter set.*
- $\mathcal{D}_{i,1:t}$ — Qualifying provenance-linked observations through decision point t. *Unit or type: evidence set.*

B Conceptual and operational equation classification

Each equation below is a design artifact at its declared epistemic and implementation state. No equation is presented as a validated law of affective response.

- **ANT-EQ-001 — definitional.** Equation 1. Names the consent-bounded information available for one individual at one decision point; it is not a fixed representation of the person. *Implementation: implemented mock slice. Prohibited inference: The tuple does not reveal a true emotion, diagnosis, need, or successful outcome.*
- **ANT-EQ-002 — definitional.** Equation 2. Separates consented source events, replaceable projections, and their versions and provenance. *Implementation: implemented mock slice. Prohibited inference: The partition does not prove privacy, truth, retention compliance, or correct interpretation.*
- **ANT-EQ-003 — proposed operationalization.** Equation 3. Derives the minimized working view permitted by the active consent grant for one declared purpose. *Implementation: implemented mock slice. Prohibited inference: A projection is not permission to use excluded, expired, revoked, or out-of-purpose context.*
- **ANT-EQ-004 — proposed operationalization.** Equation 4. Makes current user direction authoritative over accepted suggestions and neutral defaults. *Implementation: unavailable. Prohibited inference: A knob value does not directly control acoustics, emotion, usefulness, or benefit.*
- **ANT-EQ-005 — proposed operationalization.** Equation 5. Low-pass filters ordinary normalized control changes to reduce jitter while leaving stop, revocation, and exclusion immediate.

Implementation: unavailable. Prohibited inference: Control smoothing does not establish musical, waveform, or emotional continuity.

- **ANT-EQ-006 — definitional.** Equation 6. Combines typed semantic mixins under explicit conflict, exclusion, and priority rules. *Implementation: formalized only. Prohibited inference: Typed composition is not raw text concatenation or evidence of perceptually linear mixing.*
- **ANT-EQ-007 — proposed operationalization.** Equation 7. Compiles a model-independent journey into a capability-bounded adapter conditioning state and then an optional native prompt. *Implementation: formalized only. Prohibited inference: Compilation does not guarantee control adherence, model quality, or felt response.*
- **ANT-EQ-008 — estimated model.** Equation 8. Represents alternative provisional state hypotheses conditioned on authorized observations and prior controls. *Implementation: unavailable. Prohibited inference: The belief distribution is not the person's true inner state, an objective need, or a diagnosis.*
- **ANT-EQ-009 — conceptual.** Equation 9. Transfers constrained receding-horizon comparison to short, inspectable journey proposals while retaining human approval. *Implementation: unavailable. Prohibited inference: The equation does not model a person as a known plant or establish stability, optimality, or intervention benefit.*
- **ANT-EQ-010 — proposed operationalization.** Equation 10. Requires verified playable audio to cover observation, planning, generation, verification, scheduling, and uncertainty time. *Implementation: unavailable. Prohibited inference: Meeting the inequality does not establish musical continuity, subjective fit, or real-time performance before measurement.*
- **ANT-EQ-011 — proposed operationalization.** Equation 11. Measures bounded change between adjacent typed semantic conditions. *Implementation: unavailable. Prohibited inference: Semantic proximity does not guarantee perceptual linearity, smooth audio, or desired emotion.*
- **ANT-EQ-012 — proposed operationalization.** Equation 12. Combines normalized boundary distances for measured musical and acoustic features. *Implementation: unavailable. Prohibited inference: A low aggregate distance does not prove semantic coherence, seamlessness, safety, or benefit.*
- **ANT-EQ-013 — proposed operationalization.** Equation 13. Defines one gain curve for joining two verified signals over an accepted overlap window. *Implementation: unavailable. Prohibited inference: A crossfade cannot repair incompatible rhythm, harmony, phrase, timbre, meaning, or emotional pacing.*
- **ANT-EQ-014 — definitional.** Equation 14. Keeps perceived expression, felt response, intensity appraisal, usefulness, harm, resonance, mismatch, and aftereffect as separate fields linked to exposure. *Implementation: implemented mock slice. Prohibited inference: Intensity, catharsis, completion, or engagement is not automatically success or harm.*
- **ANT-EQ-015 — future empirical model.** Equation 15. States the future within-person response distribution that would require qualifying exposure-linked observations. *Implementation: unavailable. Prohibited inference: No response predictor is fitted, validated, or authorized to steer generation.*
- **ANT-EQ-016 — future empirical model.** Equation 16. Shows a generic future posterior update

while preserving evidence and a prior model snapshot. *Implementation: unavailable. Prohibited inference: The form supplies no valid likelihood, causal effect, forgetting rule, sample size, or autonomous update authority.*

C Working protocol checklist

The following reader-facing checklist is generated from the byte-locked machine-readable protocol rather than maintained independently. It records the current evidence stages, prospective schedule, and unresolved activation gates; it is not a recruitment notice, an ethics determination, or collection authority.

Protocol identity. ANT-PROT-FEAS-001, version 1.1.0, was frozen as a **frozen-design-protocol** artifact on 2026-09-07. Its byte-exact SHA-256 is shown below in two contiguous halves.

```
8d6848f148a627424c566f0438110377
9debd65b4563bd4712652e09cd715024
```

Protocol succession. This is the pre-collection corrective successor to version 1.0.0. The superseded protocol remains preserved at `experiments/protocols/antidote-feasibility-v1.json` with byte-exact SHA-256 `dbedbcbb74303373a7e00e95ad063b025d8cf6033864f4ad5a390d15703ba403`. No qualifying record or human collection began under the predecessor.

Authority status. Collection authority is **false**; no formal human study has begun. The amendment and deviation log contains 0 entries.

Evidence stages

- D0 Design traceability.**
Current state: `available-for-design-review`. Unit: one governed transformation or authority boundary.
- T0 Synthetic conformance and recovery.**
Current state: `executable-mock-slice`. Unit: one deterministic fixture, command trace, or injected failure path.
- T1 Real-model technical qualification.**
Current state: `blocked-no-real-model`. Unit: one immutable generation specification and terminal generation result.
- H1 Repeated-session within-person feasibility.**
Current state: `blocked-no-collection-authority`. Unit: one assigned session attempt, with exposure analyzed only when verified audio was actually played.

Prospective H1 schedule

- 6 blocks $\times$ 3 conditions = 18 planned completed exposures, with an absolute ceiling of 24 scheduled attempts.
- Each proposed exposure is 8 minutes; at least 48 hours separate exposures, with at most 3 in seven days.
- The later aftereffect window targets 24 hours and accepts 18–30 hours.

- The balanced block-order assignment contains exactly 720 schedules (6!), selected by the frozen uniform algorithm; the six H1 block seeds are committed before session 1.
- G generic generation.** Create an eight-minute instrumental ambient listening piece with no lyrics, sudden transients, or intentionally intense effects.
- S structured non-personalized journey.** Use the selected current-state and desired-transition categories with one frozen meet-hold-transition-integrate template; do not use free-text history, personal descriptors, or prior-session mappings.
- P negotiated personal semantic journey.** Use only the explicitly approved working projection, person-authored semantic mixins, inclusions, exclusions, and accepted journey-plan revision.

Human-collection activation gates

- **GATE-REAL-MODEL — not satisfied.** Real model, adapter, artifact, cancellation, and T1 technical qualification complete.
- **GATE-PRIVACY — not satisfied.** Encryption, retention, deletion, backup, projection invalidation, minimum-data private research export, and separately minimized public audit-envelope controls reviewed for the intended setting.
- **GATE-CONSENT — not satisfied.** The version 2 scope-specific consent contract and participant-facing consent, withdrawal, adverse-response, and data-use materials are implemented and approved for the actual study context.
- **GATE-INDEPENDENT-REVIEW — not satisfied.** An independent qualified reviewer records the applicable ethics, safety, and researcher-participant determination; no approval is presumed.
- **GATE-ASSIGNMENT — not satisfied.** Assignment seed, all 720 eligible schedules, selected complete schedule, and the six fixed block seeds are frozen and their enforcement is verified before session 1.
- **GATE-INSTRUMENT — not satisfied.** The timed session workflow, exact prompts and anchors, version 2 response and append-only correction contract, playback setup, accessibility path, and implementation match this protocol.
- **GATE-ANALYSIS — not satisfied.** Analysis code and synthetic validation fixtures reproduce the frozen estimands, denominators, missingness, and sensitivity rules.
- **GATE-STOP-GO — not satisfied.** A documented pilot stop/go decision authorizes only the explicitly reviewed next step.

Freeze and audit rule

- Any material change creates a new semantic protocol version, a new content lock, an updated analysis plan, and renewed review before collection.
- Never rewrite the frozen file. Append a timestamped amendment or deviation with rationale, affected records, evidence impact, author, reviewer, and disposition.
- Report zero movement, null contrasts, opposite-direction movement, mismatch, unwanted intensity, burden, harm, stops, failures, and adverse events with the same visibility as favorable observations.

D Claim and evidence classification index

This reader-facing index projects the disposition of the governed claim ledger at the source revision printed on the title page. It is a navigation aid rather than a second claim authority; the complete statement, evidence, permitted sections, and wording boundary for every identifier remain in `research/notes/CLAIM_LEDGER.md`.

External evidence (ANT-SRC-001–007). Reviewed sources and source syntheses establish prior-system, measurement, architecture, protocol, and risk-review context. Their findings remain attributed to those sources and do not become Antidote outcomes.

Open hypotheses (ANT-HYP-001–003). Within-person advisory relationships, responsive generate-ahead journeys, and improved interpretability through separated records remain questions to test. None is reported as learned, seamless, useful, or beneficial.

Recorded observations (ANT-OBS-001–004). The originating cathartic self-report is hypothesis-generating only. Repository implementation and reporting-contract states are technical observations at named revisions; they are not formal human results.

Bounded interpretations (ANT-INT-001–002). The reviewed material supports end-to-end inspectability as a defensible contribution boundary. The bounded search did not identify the complete governed chain in one evaluated system, but that absence is not proof of global novelty.

Accepted design and protocol claims (ANT-CLM-001–007). The manuscript may state that it formalizes the governed record chain, preserves optional observations as uncertain and correctable, separates evidence layers, requires explicit human authority, proposes the two-rate architecture, and freezes versioned protocol and reporting contracts. These claims establish neither feasibility nor outcome validity.

Unresolved candidate (ANT-CAN-001). Whether the complete combination constitutes a novel research-system contribution remains unresolved. Only the narrower bounded-review wording is permitted.

Rejected claims (ANT-NEG-001–004). First-system, efficacy, deterministic mechanism, guaranteed transition, privacy-or-safety proof, and synthetic-as-human-result statements are prohibited.

A stronger future statement requires its own stable ledger identity, exact source or observation, evidence class, permitted wording, review disposition, and an explicit downgrade or rejection path. No neighboring record promotes it automatically.

E Consent and context-projection record

Any future qualifying session record must make the consent boundary auditable without placing private content in this public appendix. The protected record must identify the protocol and source revision, decision point, separately authorized action, purpose, consent-grant revision, permitted source types and fields, explicit exclusions, retention mode, expiry or revocation state, and the person who reviewed the grant. Inspect, project, generate, analyze, play, retain, learn, and export remain separate authorizations; one action cannot be inferred from another.

The protected context-projection record must identify every contributing source event, semantic item, derivation rule and version, uncertainty, conflict, omission, person edit, and final review disposition. It must preserve a digest and provenance for the exact approved projection while keeping raw private text and row-level identifiers outside public artifacts. The model-worker record must show that only the

approved semantic projection, accepted journey-plan revision, and immutable generation specification crossed the process boundary. A later correction, revocation, deletion, or expiry must create an attributable event and invalidate affected derived projections rather than rewrite the original record silently.

Public reporting may expose only the privacy-reviewed aggregate envelope and the minimum revision and integrity metadata approved for that destination. The private record must retain the export decision, redactions, excluded fields, reviewer disposition, and the effect of withdrawal or deletion on payloads, projections, analyses, backups, and previously shared artifacts. This checklist specifies what must be recorded; it does not establish meaningful consent, privacy, secure deletion, or current collection authority.

F Provenance and reproducibility checklist

The checklist below states the minimum record needed to reproduce or audit one reported Antidote artifact. A field marked unavailable remains unavailable; it must not be inferred from a neighboring record or replaced with a zero.

Source and build identity. Record the full Git revision, paper version and stage, build theme, source-date epoch, operating system, build commands, and the versions or immutable commits named by `dependencies/beacon.lock.toml`, `publication/antidote-site-suite.lock.json`, language toolchain files, and package lockfiles. Retain `provenance.json`, `publication.json`, `site.json`, and `SHA256SUMS` with the generated PDF, HTML, and source archive.

Protocol and reporting identity. Name protocol `ANT-PROT-FEAS-001`, its semantic version, content-lock digest, freeze date, deviation log, collection-authority state, and the exact `ANT-REPORT-RESULTS-001` version. Register every source package and promotion decision; an empty registry is an unavailable result state, not evidence of no effect.

Model and artifact identity. For a synthetic record, identify the deterministic mock worker, fixture, schema version, seed, request, output digest, warnings, interruption state, and test revision. A future T1 record must additionally identify the real model, adapter, weights and licenses, hardware profile, analyzers, generation parameters, mutation manifest, seed policy, attempted outputs, failures, exclusions, and output-rights review. Those T1 records are currently unavailable.

Exposure and response identity. A future H1 record must link the approved plan and verified artifact to actual playback, stop or interruption events, response windows, corrections, missingness reasons, adverse observations, and the applicable consent, retention, and review decisions. Raw personal content and row-level identifiers remain protected; only a separately approved aggregate public envelope may enter a public result package. No qualifying H1 exposure or response package exists at this revision.

Claim and exclusion audit. Record which source, observation, or result supports each reported statement; the claim-ledger disposition; unresolved contradictions; and every omitted, failed, null, negative, mismatched, missing, or adverse record. Confirm that synthetic fixtures are not labeled human outcomes and that provenance or checksum agreement is not presented as truth, consent, privacy, ownership, safety, or efficacy.

Audit row	Version / state	Identity / source	Evidence class	Boundary or gate
Protocol lock	ANT-PROT-FEAS-001 v1.1.0	SHA-256 8d6848f148a62742. . .	Decided: frozen prospective design	Collection authority false; no qualifying record started.
Reporting contract	ANT-REPORT-RESULTS-001 v1.1.0	SHA-256 33fdf8a64f6ccec4. . .	Repository observation: governed empty state	No T0, T1, or H1 source package is registered; unavailable is not zero.
D0: Design traceability	available-for-design- review	manuscript + ledgers	decided	The manuscript contract, equation registry, claim ledger, and protocol are synchronized.
T0: Synthetic conformance and recovery	executable-mock-slice	ANT-REC-T0-001 absent	repository-observation	The checked-in mock worker and test revision are recorded.
T1: Real-model technical qualification	blocked-no-real-model	ANT-REC-T1-001 absent	technical-observation	A model, adapter, hardware profile, analyzer set, mutation manifest, seed policy, and output-rights review are frozen before the first qualifying run.
H1: Repeated-session within-person feasibility	blocked-no-collection- authority	public H1 envelopes absent	self-report-and-exposure- observation	All H1 activation gates require independent satisfaction under a new activation record.
Claim promotion	ANT-CLM-006 / ANT-CLM-007	claim ledger	Accepted artifact claims only	RQ2 and RQ3 remain unanswered; later evidence needs separate promotion review.

Digests identify the exact local artifacts used by this projection. They support integrity inspection, not truth, privacy, safety, ownership, or efficacy. No private or participant record is represented.

Table 8: Current protocol, reporting-contract, stage, source-package, evidence-class, and claim-promotion audit; no qualifying T0, T1, or H1 result package is registered.

Table description: The audit begins with the frozen ANT-PROT-FEAS-001 version 1.1.0 identity and the ANT-REPORT-RESULTS-001 version 1.1.0 governed empty state. Separate D0, T0, T1, and H1 rows show their current state, expected source, evidence class, and entry gate. A final row limits accepted promotion to protocol- and reporting-artifact claims while RQ2 and RQ3 remain unanswered. Digests support integrity inspection only; the table contains no private record, participant observation, result value, or claim of safety or benefit.

G Further reading and research shelf

The numbered References list contains works cited directly by the manuscript. The index below preserves verified but currently uncited bibliography entries as additional reading; inclusion does not imply endorsement or support for a paper claim. The complete living atlas, including discoveries that have not entered the verified bibliography shelf, remains a repository supplement.

43 directly cited works; 5 additional verified bibliography entries; 96 total living-atlas records.

- **gong2026acestep**

ACE-Step 1.5. Catalog class: preprint; state: metadata-verified. Domains: generative-audio, architecture.

- **aws2026eventsourcing**

Event sourcing pattern. Catalog class: implementation; state: background-only. Domains: architecture, provenance.

- **jain2019intent**

Probabilistic human intent recognition for shared autonomy in assistive robotics. Catalog class: system; state: metadata-verified. Domains: adaptive-control, measurement, human-authority.

- **copet2023musicgen**

Simple and controllable music generation. Catalog class: system; state: metadata-verified. Domains: generative-audio.

- **kothe2025lsl**

The Lab Streaming Layer for synchronized multimodal recording. Catalog class: system; state: metadata-verified. Domains: psychophysiology, architecture.

H Migration provenance

The initial research snapshot originated at `egohygiene/empathy/research/antidote`. The immutable extraction commit, source tree, origin commit, file-level Git blobs, and dispositions are recorded in `MIGRATION.md` and `research/inventory/empathy-source-tree.tsv`.