Persistent Sound Worlds for EEG Responsive Composition

Abstract. Physiologically responsive soundscapes must accommodate changing control signals while maintaining the identities and consequences of audible events. We present a macOS prototype that separates persistent world processes from their adaptive acoustic presentation. A scene is authored through source identities, support regions, process dependencies and a sparse cue budget. An EEG interface supplies quality-gated vendor indices for branch selection, permitted motion and bounded audio control. Local and optional cloud planners submit recordings to a common executor that checks admissibility before and after asynchronous preparation. Two artistic cases develop the approach: a cave whose inhabitants are disclosed through distributed cues, and a rainy shelter whose vessel progresses from heating to serving. Evaluation combines 34 contract checks, controlled component ablations, scene traces and archived digital renders. Across 15 selected post-admission fault cases, the complete executor prevents all invalid starts; removing both post-decode checks permits all 15. Five synthetic control profiles preserve identical shelter process times. A separate response-selection audit shows that signal-dependent decisions can improve synthetic value, while cost and uncertainty restrictions have distinct effects and hidden mechanism reversal remains harmful. These results establish execution properties of the prototype and expose the remaining perceptual and physiological questions.

Keywords: computational art · soundscape composition · physiological interaction · spatial audio · persistent worlds · runtime constraints

1 Introduction

An interactive soundscape presents evidence of a world. A breath suggests an inhabitant, a localized drip suggests a wet boundary, and boiling suggests a vessel undergoing a process. The relationships among these events matter when presentation changes: a different recording should retain its source, an object should retain its support, and a serving cue should follow heating. Physiological control introduces a further disturbance because input can change while a previously selected sound is being prepared.

We investigate how an authored sound world can retain these relationships while responding to electroencephalography (EEG). The prototype begins on a ship and opens into four scenes: a sheep cave, a goddess-associated riverbank, Circe's courtyard and Eumaeus's rainy shelter. The work is designed for sustained listening with few demands on the participant. Its composition distributes clues to place, material and inhabitants across time. Prepared recordings provide ingredients; the complete scene mix is assembled during execution.

Our authoring hypothesis is a minimum credible world (MCW): a restricted collection of sound-producing causes whose selected cues could support a coherent interpretation. The cue budget is an artistic constraint. Credibility is a perceptual question. This distinction allows the implementation to enforce identifiable source and process relationships without treating their perceptual success as already demonstrated.

The central contribution is an execution architecture for this authoring approach. World processes, object permissions and acoustic presentation share an admission path, including checks after asynchronous decoding. We examine the architecture through two worked scenes and a controlled ablation that locates distinct failure modes in the execution path. A separate synthetic audit examines the decision problem that a future learned selector would face: when does a current signal justify replacing an otherwise admissible response?

The paper asks which source and temporal properties survive adaptive execution, and under what synthetic conditions a signal improves response selection. The implemented controller uses authored mappings; the independent learner does not make its live decisions. Evaluation comprises software executions and generated episodes. Listener understanding and the value of live EEG remain empirical questions.

2 Related work

Physiological computing and passive brain-computer interfaces use observed signals as interaction inputs [1, 2]. Gilroy and colleagues connected a brain-computer interface to plan-based narrative [3]; MoodMixer used EEG-derived indices in collaborative sonification [4]. BioSynth further examines what affective measurement asks participants to feel and perform [5]. Our controller follows this interpretive caution: its inputs are vendor indices, and the artwork does not require the listener to maximize an internal score. Signal validity and generalization remain separate questions from interface smoothness [6].

Scaper specifies compositions of isolated sound events [7]. WavJourney compiles an audio script into model calls and operations [8], while SoundscapeAgent exposes planning, source acquisition, temporal layout and rendering [9]. Our implementation uses prepared recordings and changes their arrangement as a world proceeds. The distinctive execution problem occurs when a plan that was valid at admission becomes invalid before a source starts.

Scene2Sound grounds sound-emitting objects in persistent positions within 3D Gaussian worlds [10]. Its concern with spatial consistency overlaps with our source anchoring. Here, the explicit object catalog additionally specifies process phases, permitted motion and event dependencies. Scene2Sound and SoundscapeAgent are 2026 preprints. We evaluate our implementation through component ablations; no cross-system performance comparison is reported.

Interactive scores already represent temporal constraints and conditional branches. OSSIA supports authoring and execution of such scores [11], and Arias and colleagues use timed automata for multimedia-score verification [12]. Narrative planning similarly relates event order to causal progression [13]. Runtime enforcement methods provide broader accounts of action-stream restriction [14, 15]. We use application-level checks and receipts, with no claim to formal synthesis or general soundness. The contribution lies in connecting those checks to object-supported sound, persistent scene processes and physiological presentation controls.

Gaver's ecological account motivates treating everyday sounds as evidence of events [16]. Recognizing a source, however, does not establish that it belongs convincingly in a scene [17]; plausibility also differs from other aspects of presence [18]. The image method and Pyroomacoustics provide established geometric treatments of propagation [19, 20]. Our finite proxy uses Web Audio components [21] to support an authored spatial arrangement. Its parameters and digital measurements are evaluated as rendering behavior, not as measurements of perceived enclosure or comfort.

3 Authoring a persistent sound world

3.1 Sources and permitted presentation

An MCW describes facts that a composition should support. For each sound-producing object, the author records its identity, compatible recordings, support region, permitted motion, temporal eligibility and rendering behavior. A river has water support; a seep belongs to rock; a vessel has one fixed position. Elevated sounds remain possible when the represented source has elevated support. Figure 1 shows the packaged visual references for the two cases developed here.

Fig. 1. Packaged visual references for the cave (left) and rainy shelter (right). The images identify the authored settings; they are not screenshots of an evaluated listening session or measurements of source geometry.

The author then assigns each object a role in the scene's evidence. In the cave, hearth, opening, wet boundary, flock and inhabitant are distinct roles. The scheduler ranks eligible cues using coverage, recent use, required processes and capacity. It admits one discrete foreground event and up to three scheduled continuous layers; retiring layers can overlap during fades. These authored limits distribute attention across causes; the smallest sufficient source set remains to be tested.

Let W be persistent process state, L the ingredient catalog, q the filtered control vector and H the execution history. During active, unpaused simulation, the world advances independently of EEG:

$$W_{t+h} = F(W_t,\ h). \tag{1}$$

A planner proposes a presentation, which is checked against the contracts C(W):

$$p_t = \Pi_{C(W_t)} P(L, q_t, H_t). \tag{2}$$

The projection symbol denotes validation, rejection or restriction. An accepted response may change a recording, a permitted pose or salience while preserving source identity and process consequences. This creates a compositional distinction between an event's continued existence and the moments at which its evidence is presented.

3.2 The cave as distributed evidence

The cave offers a listening position among resting animals, rock, embers and a giant. The visual reference contains many animals, but the acoustic catalog registers four sheep positions. A shared motion group binds these sound sources to their visual instances. Horizontal and depth offsets are bounded by 0.55 and 0.35 scene units, with a maximum speed of 0.12 scene units per second. The hearth, seep and giant retain fixed anchors.

The composition gives each cue a different interpretive function. Embers sustain evidence of a resting place. Faint opening air marks a boundary. A localized seep associates moisture with a particular rock surface. Sheep calls distributed among four anchors indicate occupancy beyond one point, while low breathing introduces the giant through an ongoing bodily process. The intended reading emerges from intermittent evidence: the giant remains present during the intervals between breaths. This reading is the compositional intention, pending listener evaluation.

Figure 2 pairs the registered source layout with a 400-second instrumented execution. The first seep, giant and sheep starts occur at 8.05, 20.05 and 40.10 seconds. Across the trace, seven sheep events visit all four anchors, seven giant events recur, and six seep events mark the wet boundary. The maximum foreground concurrency is one. These observed timings differ from the earliest eligible times because selection also respects spacing, source occupancy and recent use.

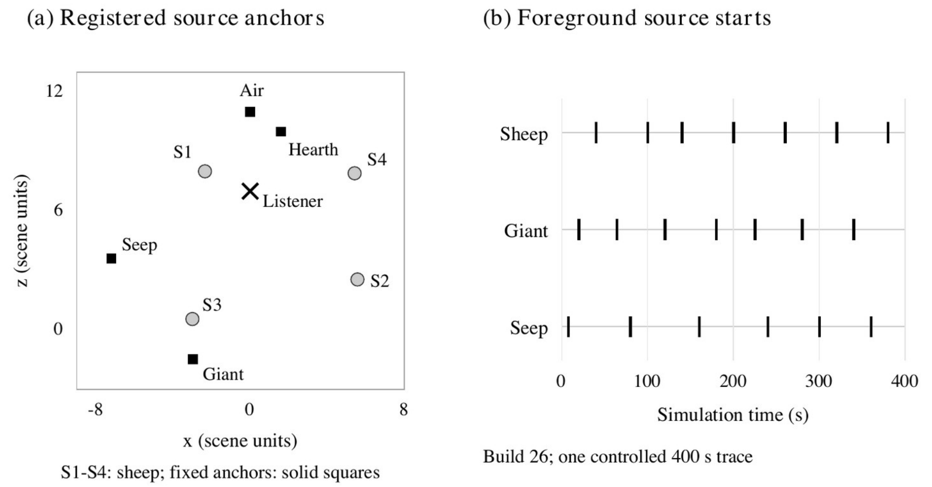

Fig. 2. Cave source layout and score. Panel (a) uses registered x-z anchors, not a reconstructed room boundary; sheep can move within permitted bounds. Panel (b) plots discrete source-start callbacks in one 400-second fixture execution. Ongoing hearth and opening layers are omitted from the event plot.

Control can alter which cause becomes salient without removing its role from the catalog. In two 35-second traces, injected coordinates (attention, relaxation, fatigue, meditation) of (0.95, 0.10, 0.05, 0.20) and (0.05, 0.95, 0.30, 0.80) yield a seep followed by a giant cue and a seep followed by a sheep cue, respectively. The flock offsets also differ within their allowed bounds. The response therefore changes the scene's current emphasis. Its appropriateness for a listener remains open.

3.3 The shelter as an unfolding process

The rainy shelter contrasts persistent exterior weather with an interior hearth and a changing vessel. Its acoustic composition is organized around shelter and hospitality: the environment continues while water becomes ready to serve. Rain, hearth and vessel are essential layer roles. Additional eaves and courtyard sources are available only within capacity and support constraints. The vessel remains one mono direct source with no reflections or late send. This exception deliberately privileges localization over reverberant integration.

Water starts at 52 degrees Celsius, reaches the first-boiling threshold at 99.7 degrees, and becomes eligible for serving at least fourteen seconds later. The simulation integrates in steps no larger than 0.05 seconds. Serving removes the vessel from the fire and retires its associated source. A pause freezes process time; contact loss invokes fixed-reference control while active world time continues. Restoration preserves recorded process facts, so it cannot restart heating or silently replay a serving event whose completion is uncertain.

The scene illustrates why presentation should not determine process truth. Lowering the salience of the vessel does not cool its water. Rejecting a late serving sound does not undo a completed process. This separation allows the author to preserve temporal consequences while accepting that an audience may miss some evidence. Cue omission is consequently a compositional cost that the system logs rather than conceals.

4 Planning and execution

4.1 A common admission path

SoundKnowledge stores recordings, objects, rules, reviews, provenance, plans, events and acoustic profiles. Build 26 includes 47 asset records, of which 15 are blocked, and 32 objects across eight scene identifiers. Five scenes are available through the journey; three identifiers remain for data compatibility. The 376 scene-asset review rows are versioned metadata assessments, rather than independent listening judgments.

JourneyDirector accepts local or optional cloud ScenePlans. Each plan contains a scene, epoch, catalog revision and validity interval; each item identifies an object, asset, gain scale and offset. Placement and support come from the object catalog. The local heuristic planner works without cloud configuration. The cloud path receives structured context, proposes a plan and obtains a separate critic response before local validation. Typed updates can tighten supported restrictions but cannot relax fixed geometry or inject code. Figure 3 summarizes the implemented path.

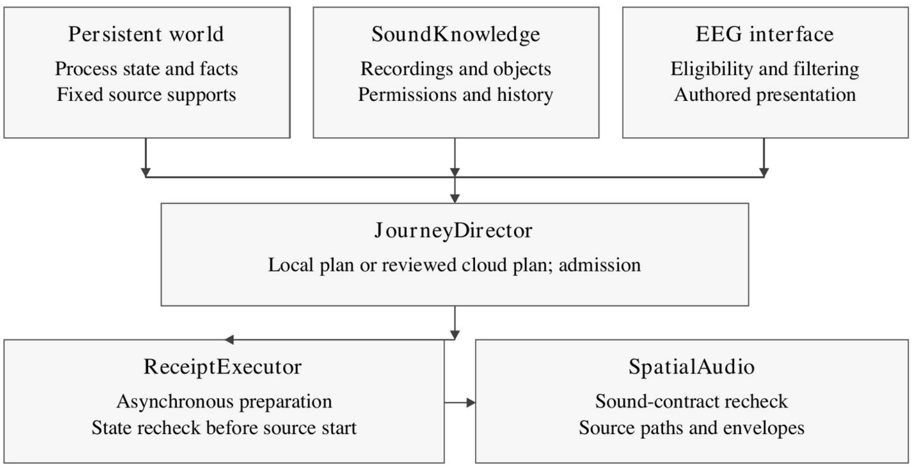

Fig. 3. Implemented execution path. Planning uses persistent state, catalog permissions and eligible EEG control. Sound-contract and executor-state checks run after asynchronous decoding. The independent learner is outside this deployed path.

The acquisition facility checks provenance, downloads an authorized asset, records its hash, decodes and inspects the signal, and requests metadata review before admission. Text review cannot determine whether a waveform contains the described source or an abrupt transient. The evaluation uses fixture planning and acquisition responses, leaving live model performance unmeasured.

4.2 Rechecking after preparation

ReceiptExecutor manages queued, preparing, started, completed, expired, cancelled, failed and interrupted events. Before preparation, it checks scene eligibility, time windows, dependencies and occupancy. Asynchronous decoding creates an interval in which these conditions can change. The voice boundary therefore validates the sound contract again, and the executor rechecks playback state, active scene, deadline, prerequisites, capacity and event spacing immediately before source start.

These two checks protect different properties. Contract validation prevents a source whose permission has been revoked from becoming audible. State revalidation prevents a still-compatible sound from starting after pause, scene replacement, expiry or loss of capacity. Preparation failures may be retried up to three attempts while eligibility remains. A started event without reliable completion becomes interrupted on restoration rather than being silently replayed.

Continuous replacement follows a related policy. The existing layer remains active while a replacement is prepared. Once an admissible replacement starts, the old layer retires gradually. Decoding failure therefore need not create an immediate gap. This policy preserves an existing cause at the cost of postponing a desired presentation change. Receipts document scheduling and callbacks; physical sound onset requires audio-clock or loopback measurement.

4.3 Physiological presentation control

The EEG102 BLE and SDK guides specify attention, relaxation, fatigue and meditation as dimensionless indices in the range 1-100 [25, 26]. The meditation descriptions use the opposite direction from the other displayed coordinates. The bridge maps attention, relaxation and fatigue as (r-1)/99 and meditation as (100-r)/99. It rejects a packet when any byte falls outside 1-100. Contact, connection, a new sequence identifier and a sample age no greater than three seconds are also required. These are input-eligibility checks, not validation of the named psychological constructs.

Optional calibration uses at least 30 valid seconds and 15 distinct packets. For each oriented coordinate, it estimates a median center and an interquartile-range scale, with a floor of 0.15. It then maps the centered coordinate to its fixed reference plus one quarter of the standardized deviation and clips to [0,1]. A five-update median and a 0.025 deadband precede continuous smoothing:

$$q_{t+h}=q_t+\left(1-\exp(-h/2.5)\right)(m_t-q_t). \tag{3}$$

Here m is the accepted target and the time constant is 2.5 seconds. Invalid or stale input returns control toward the fixed reference (0.30, 0.50, 0.10, 0.50), preserving active world processes. Selection, position and parameter control can be enabled independently.

The opening router evaluates about thirty seconds of filtered target features, requires at least fifteen packets, and rejects a window whose standard deviation exceeds 0.12 in any coordinate. It ranks branches with the linear weights in Table 1. The cave mapping emphasizes attention, the riverbank mapping the oriented meditation coordinate, Circe's courtyard relaxation, and the shelter fatigue. These affinities encode different authored modes of attending to the work; they do not classify which setting is beneficial for a measured state.

Table 1. Authored branch affinities on oriented coordinates. Meditation is reversed relative to the raw vendor index.

Branch	Attention	Relaxation	Fatigue	Meditation
Cave	0.70	0.15	-0.25	-0.05
Riverbank	-0.10	-0.10	-0.25	0.80
Circe	-0.10	0.75	-0.35	-0.15
Shelter	-0.10	0.10	0.95	-0.10

The winning score must exceed the runner-up by 0.08 and remain eligible for fifteen seconds; the opening lasts at least sixty seconds. A stable but ambiguous window can propose the shelter after 120 seconds, subject to the same hold. Persistent variability can prevent automatic entry indefinitely. Manual selection is available separately. Once entered, the branch remains active until the user changes it, avoiding repeated narrative switching in response to feature fluctuations.

4.4 Spatial rendering and envelopes

Acoustic profiles specify metres per scene unit, finite surfaces, openings, material estimates and decay limits. The cave uses a 24 by 12 by 23 metre chamber proxy, and the shelter an 8.1 by 3.24 by 5.25 metre local structure. Direct and reflected path lengths determine delays using a sound speed of 343 metres per second. First-order reflections are admitted only when their bounce lies on a finite solid surface outside an opening. Intervening walls attenuate paths, and material filters reduce high-frequency energy.

For enclosed and semi-open profiles, an Eyring-style volume-to-absorption estimate sets a bounded decay target [27]. Three filtered stochastic bands centered nominally at 250, 1000 and 4000 Hz form the convolution response. The cave's low, mid and high decay parameters are 2.400, 1.398 and 0.804 seconds; the shelter's are 0.344, 0.241 and 0.167 seconds. These are authored parameters, not measured room reverberation times. The early-reflection gain budget is 0.22 in the cave and 0.12 elsewhere. Vessel localization is an explicit exception to the room treatment.

The active composition excludes clacks, hard impacts, footsteps, thunder and discrete wave crashes. Sheep, giant and seep families have minimum spacings of 24, 38 and 18 seconds, with additional jitter. Continuous layers enter and retire over ten seconds; scene transitions last at least sixteen seconds. Foreground envelopes scale to recording duration. Waveform preparation checks ten-millisecond blocks against peak and RMS ceilings of 0.055 and 0.018. Ten prepared water loops use constant safety gain to preserve their circular seams. The final mix has a numerical ceiling of 0.22. These controls regulate digital changes and do not determine headphone sound pressure or guarantee comfort.

5 Evaluation of the implemented system

5.1 Materials and protocol

Evaluation uses version 9.5.2 build 26. Four instrumented JavaScript suites contain 34 contract checks. The distributed subset also includes prepared-loop, layer-continuity, feature-stability and branch-router programs. All were executed successfully. Their audio nodes and decoded PCM are fixtures, and their feature streams are synthetic. The supplementary archive distinguishes these runnable checks from archived full-source Web, native WebKit, package and digital-render reports.

Two analyses use the same application source: a component ablation of post-decode checks isolates execution mechanisms, and deterministic scene and routing traces expose compositional behavior.

5.2 A controlled ablation of post-decode checks

We admit a valid cave sheep event, suspend its decoding, and change one condition before resolving the prepared buffer. The five challenges are pause, active-scene change, deadline expiry, occupied foreground capacity and revocation of the sheep sound kind. Each is exercised at logical preparation delays of 0.1, 1 and 5 seconds. A sixth condition leaves the event valid at each delay, giving 18 cases per execution variant and 72 executions in total.

Four variants use identical admission, fixtures and source code apart from the specified removals: both post-decode checks removed, state revalidation removed, contract revalidation removed, and the complete implementation. The deadline challenge advances scene time past a deadline halfway through the selected delay. Other cases remain inside the event window, isolating their respective condition. Source-start calls, rather than error labels, determine whether an invalid event was launched.

Table 2. Post-decode component ablation. Counts are source starts in selected instrumented cases, not estimated deployment failure rates.

Execution variant	Invalid starts (15 fault cases)	Valid starts (3 controls)
Both post-decode checks removed	15/15	3/3
State revalidation removed	12/15	3/3
Contract revalidation removed	3/15	3/3
Complete execution path	0/15	3/3

Table 2 shows that the complete path blocks all 15 invalid starts while permitting all three valid starts. Removing state revalidation permits the twelve pause, scene, deadline and capacity failures. Removing contract revalidation permits the three permission-revocation failures. Removing both permits all fifteen. These selected counterexamples establish a need for post-admission checks, with valid-start controls ruling out suppression of every event. Their frequency in deployment and physical onset latency remain unmeasured.

5.3 Process invariance and routing behavior

Five constant synthetic coordinate profiles drive the shelter for 105 seconds each. Four strongly favor the different branches and one is ambiguous; direct scene entry isolates within-scene control from opening selection. Every trace reaches first boiling at 66.45 seconds and serving at 80.50 seconds, and every serving receipt completes. These times are identical across profiles within numerical precision. Together with the different cave cue selections described in Section 3.2, they exhibit the intended separation between changing presentation and persistent process progression.

At two feature packets per second, the four stable branch profiles select their intended branches at 60 seconds. The ambiguous profile selects the shelter at 135 seconds. Attention alternating between 0.05 and 0.95 in five-second blocks produces no selection over the 600-second observation horizon. Replacing that variability with constant attention 0.95 after 300 seconds produces cave selection at 341 seconds. Invalid input produces no selection. These cases characterize routing boundaries; human waiting times and manual-intervention rates remain unmeasured.

The cave trace in Figure 2 covers all four sheep anchors and the warmth, opening, wet-boundary, flock and inhabitant roles. Existing continuity tests retain the previous layer under delayed or failed replacement decoding. Voice tests reject unsupported water placement, diffuse vessel rendering and incompatible recordings. Acoustic checks verify expected parameter directions and reject reflection paths outside finite surfaces or through openings. Together, the tests establish contract execution and trace coverage.

5.4 Digital renders and audibility limits

Five archived renders use a two-channel OfflineAudioContext at 24 kHz. Ship and shelter last 200 seconds each; cave, riverbank and courtyard last ninety seconds each. Attention alternates between 0.9 and 0.2 while the other injected coordinates remain fixed. Rendering suspends approximately every 0.25 seconds for controller updates and awaits pending preparation before resuming. This produces deterministic digital traces but removes preparation latency from real-time contention.

Table 3. Archived build-26 digital renders. Linear amplitudes and derived RMS in dBFS; no calibrated sound-pressure or equal-loudness comparison.

Scene	Length (s)	Peak	RMS	RMS (dBFS)	Max. 100 ms RMS rise
Ship	200	0.001943	0.000324	-69.79	0.000173
Cave	90	0.003598	0.000716	-62.90	0.000333
Riverbank	90	0.002141	0.000265	-71.54	0.000137
Circe	90	0.004649	0.001108	-59.11	0.000520
Shelter	200	0.008409	0.000914	-60.78	0.000275

All outputs are nonzero, remain below the 0.22 peak ceiling, and have adjacent 100-ms RMS increases below 0.012 and foreground counts at most one. Table 3 expresses RMS in dBFS as $20\log_{10}(RMS)$. The low levels, from -71.54 to -59.11 dBFS, help satisfy upper-bound assertions but do not establish audibility, equal loudness or cue retention at the ear.

Prepared-loop checks inspect constant attenuation, seam energy and the use of preparation metadata. None of the ten prepared water files contains a 100-ms block below the larger of three percent of its median block RMS and 10 to the power of -8. This detects processing gaps; it is not a naturalness criterion. Perceptual evaluation requires controlled playback gain, matched materials and listening judgments.

6 A separate audit of response selection

6.1 The decision problem

An executor determines whether a response is permissible. It does not determine whether that response is preferable for a listener. We examine this second question using an independent binary learner. H denotes available history, E a pre-decision feature and Q the availability of that feature and both candidate responses. For an outcome Y in [0,1], the conditional response contrast and its history-only average are:

$$\begin{gathered}\tau(H,\ E) = E[Y(1) - Y(0) \mid H,\ E,\ Q = 1] \\ \bar{\tau}(H) = E[\tau(H,\ E) \mid H,\ Q = 1].\end{gathered} \tag{4}$$

A signal has incremental oracle decision value when the preferred response can change with E within a given history. For example, if E predicts comfort equally under both responses and one response always dominates, observing E does not change the optimal choice. This distinction motivates comparing a signal-aware policy with a history-only policy under identical eligible alternatives.

The reference implementation fits separate linear residualized contrast models using bases (1,h,e,he) and (1,h), two-fold outcome prediction, randomized propensity 0.5 and no regularization [22]. Let $\hat{a}_H$ and $\hat{a}_E$ be the selected actions, with estimated improvement $\hat{d} = (\hat{a}_E - \hat{a}_H)\hat{\tau}(H, E)$. A gate retains the history-only action unless the estimated improvement exceeds an override cost c and a pointwise uncertainty allowance b:

$$\pi_G(H,\ E,\ Q) = \hat{a}_H + 1\{Q = 1, \hat{d} > b + c\}(\hat{a}_E - \hat{a}_H). \tag{5}$$

Defaults are $c = 0.02$ and $b = 2.576\,s$, where s is the HC1 standard error. This is a reference selection rule with an approximate pointwise margin, not a uniform improvement certificate [24]. Costs concern overrides, excluding the costs of acquiring EEG. Independent randomized records also support doubly robust evaluation of a frozen policy [23]. For repeated decisions, proximal effects under logged histories would need to be distinguished from the consequences of changing an entire journey [28].

6.2 Synthetic mechanisms and paired analysis

Each of thirty seeds generates 1,200 training and 20,000 independent test episodes per regime. Variables h, e and z are independent uniform draws on [-1,1]; action A is randomized with probability 0.5. Outcomes follow $Y = 0.5 + 0.03h + A\tau + \varepsilon$, with uniform noise ε on [-0.08,0.08], and remain in [0,1]. The generator contrasts are 0.25e+0.05h for crossover, 0.18+0.04h+0.03e for same action, 0.25z+0.05h for uninformative input, and zero for the null. Missing input removes forty percent of test features independently. Hidden reversal changes the test coefficient of e from +0.25 to -0.25 while preserving feature marginals.

Table 4 reports net value relative to the learned history-only policy, calculated from known potential-outcome contrasts minus 0.02 times the override rate. The raw threshold chooses response 1 for e>0; the ungated method uses the full fitted contrast. Unavailable features retain the history-only action. These generated variables are not EEG samples, and the binary alternatives are not the application's four branches.

Table 4. Net synthetic value relative to the learned history-only policy. Mean and standard deviation across 30 seeds; higher is better and the baseline is zero.

Regime	Raw threshold	Ungated	Gated
Crossover	+0.0402 ± 0.0008	+0.0420 ± 0.0008	+0.0423 ± 0.0008
Same action	-0.0925 ± 0.0005	+0.0000 ± 0.0000	+0.0000 ± 0.0000
Uninformative	-0.0223 ± 0.0008	-0.0014 ± 0.0014	+0.0000 ± 0.0000
Null	-0.0100 ± 0.0001	-0.0042 ± 0.0022	+0.0000 ± 0.0000
Missing input	+0.0241 ± 0.0004	+0.0251 ± 0.0004	+0.0253 ± 0.0004
Hidden reversal	-0.0846 ± 0.0007	-0.0812 ± 0.0009	-0.0759 ± 0.0010

The crossover gate improves on the ungated policy by a paired mean of 0.000268 synthetic outcome units, with a 95 percent t interval across seeds of [0.000210,0.000326]. This is a small simulation increment. The corresponding interval describes Monte Carlo variability of the mean difference, not uncertainty about human benefit. Under hidden reversal, gated net value is -0.075911. Restricting overrides reduces some loss relative to the ungated policy but does not restore an advantage over history-only decisions.

To separate cost from uncertainty, we use identical draws for a two-by-two comparison of c in {0,0.02} and b/s in {0,2.576}. Table 5 reports crossover net value and null override frequency. Without either restriction, the null policy overrides in 20.93 percent of episodes on average despite zero true effect. Either restriction alone suppresses all overrides in these thirty runs. At c=0.02, the cost-only crossover net value is 0.042413, compared with 0.042290 using both restrictions. The default margin therefore provides no crossover gain over the cost-only alternative in this experiment. This comparison prevents attributing the default null result exclusively to uncertainty estimation.

Table 5. Cost and uncertainty comparison on identical draws. Entries are means across 30 seeds. Cost changes both the selection rule and the net-value objective.

Cost c	Multiplier b/s	Crossover net value	Crossover overrides (%)	Null overrides (%)
0	0	+0.051047	45.12	20.93
0	2.576	+0.050938	43.14	0.00
0.02	0	+0.042413	41.15	0.00
0.02	2.576	+0.042290	39.04	0.00

The six learner tests pass. An independently shuffled test signal yields negative value in the crossover regime. A deliberately post-action variable correlates with assigned action despite zero response contrast, illustrating why a response's control signal cannot also establish its benefit. With correct propensity and deliberately constant outcome predictions, the doubly robust audit differs from known gross value by at most 0.002035 across the 180 regime-seed runs. This is an observed simulation discrepancy, not a bound.

6.3 Connection to the artwork

The implemented architecture offers a place to insert a learned decision: choosing between validated presentation plans before execution. A future pair could retain the same sheep recording, source identity, duration and world consequences while varying the flock's permitted offset, or retain the shelter's sources while varying rain-to-hearth salience. Their waveforms would still differ in masking, spectrum and loudness. Eligibility alone does not make them perceptually comparable.

Matched response pairs, human randomization and an outcome pipeline are not implemented. The synthetic learner therefore diagnoses a future selector's decision problem. Learning useful differences between real presentation alternatives requires an outcome independent of the post-response control feature and evaluation on held-out data.

7 Discussion

The cases expose a tradeoff between sparse presentation and persistent causes: the cave implies inhabitants beyond its current cues, while the shelter preserves processes during reduced salience. Contracts execute these choices; listeners must establish whether their meaning survives. Perceptual evaluation should measure source identity, spatial relations and process ordering separately from plausibility and comfort. Cue removal and density comparisons would test the MCW hypothesis.

Deployment evaluation should measure physical onset, rejection, cue omission and recovery under recorded preparation delays, extending the present logical-delay counterexamples. The router's prolonged waiting also motivates measuring input availability and manual selection in ordinary use. Fixed-reference control preserves continuity during unavailable input; it does not demonstrate the value of live EEG over matched replay.

After material validation, fixed versus adaptive presentation, local versus reviewed-agent planning, and live versus time-aligned replay signals require separate comparisons. Micro-randomization can estimate proximal response effects across pre-decision contexts [30], with prespecified outcomes, delivery-failure handling, participant-level uncertainty and controls for order, carryover and expectancy [31]. MUSHRA may support sound-quality comparisons [29]; narrative understanding requires distinct measures.

The present acoustic model also omits diffraction, full wave propagation, individualized HRTFs and measured material coefficients. Its small recording catalog and text-only metadata reviews limit variation and semantic assurance. The complete system should therefore be understood as an authored computational artwork with demonstrated execution properties and unresolved perceptual outcomes.

Ethics and development history. Informal headband use by two individuals and subsequent human engineering retests informed development. No formal ethics approval or written exemption was obtained. Consent was reported for the two initial users; the available account does not independently establish its form or documentation. The subsequent retest participant count is not specified in the materials used for this paper. Human recordings and experience reports are excluded from every analysis and from the supplement. This exclusion does not determine whether the earlier activities required review. Prospective participant research requires the applicable institutional determination and informed consent before recruitment and participation.

Generative AI assistance. OpenAI Codex assisted drafting, methodological comparison, translation, revision and verification-script preparation. The author is responsible for the manuscript, evidence and final claims.

8 Conclusion and artifact availability

Persistent sound worlds connect adaptive presentation to explicit causes and consequences. The prototype realizes this connection through a shared catalog, object contracts and revalidation after asynchronous preparation. Controlled ablations distinguish state and permission failures, while scene traces exhibit changing presentation alongside invariant process progression. The independent response-selection audit shows why signal relevance, override cost and mechanism stability must be evaluated separately before learned adaptation is deployed.

The anonymous supplement includes build-26 contract source, fixtures, ablation and scene traces, and the learner's per-run paired and factorial results. Its README maps results to commands or archived sources and documents the environment. The subset runs without a headband, network or cloud credentials. Original recordings, full native dependencies, manufacturer PDFs and participant records are excluded. Repeating archived digital renders, WebKit execution and package verification requires the omitted distribution. The subset reproduces execution logic, not the full listening experience.

References

1. Fairclough, S. H.: Fundamentals of physiological computing. Interacting with Computers 21(1–2), 133–145 (2009). https://doi.org/10.1016/j.intcom.2008.10.011

2. Zander, T. O., Kothe, C.: Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. Journal of Neural Engineering 8(2), 025005 (2011). https://doi.org/10.1088/1741-2560/8/2/025005
3. Gilroy, S. W., Porteous, J., Charles, F., et al.: A Brain-Computer Interface to a Plan-Based Narrative. In: IJCAI, pp. 1997–2005 (2013). https://www.ijcai.org/Proceedings/13/Papers/295.pdf
4. Leslie, G., Mullen, T.: MoodMixer: EEG-based Collaborative Sonification. In: NIME, pp. 296–299 (2011). https://www.nime.org/proceedings/2011/nime2011_296.pdf
5. Gee, E. M.: The BioSynth—an affective biofeedback device grounded in feminist thought. In: NIME, pp. 479–485 (2023). https://www.nime.org/proc/nime2023_66/index.html
6. Lotte, F., Bougrain, L., Cichocki, A., et al.: A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. Journal of Neural Engineering 15(3), 031005 (2018). https://doi.org/10.1088/1741-2552/aab2f2
7. Salamon, J., MacConnell, D., Cartwright, M., Li, P., Bello, J. P.: Scaper: A Library for Soundscape Synthesis and Augmentation. In: WASPAA, pp. 344–348. IEEE (2017). https://doi.org/10.1109/WASPAA.2017.8170052
8. Liu, X., Zhu, Z., Liu, H., et al.: WavJourney: Compositional Audio Creation with Large Language Models. IEEE Transactions on Audio, Speech and Language Processing 33, 2830–2844 (2025). https://doi.org/10.1109/TASLPRO.2025.3574867
9. Zhang, H., Zhao, Y., Zhang, Y., Shao, Y., Yves, S.: SoundscapeAgent: Agentic Soundscape Construction for Controllable Synthesis and Scalable Audio-Language Supervision. arXiv:2607.21857 (2026), preprint submitted to SLT. https://arxiv.org/abs/2607.21857
10. Yoshida, M., Togo, R., Ogawa, T., Haseyama, M.: Scene2Sound: Auditory-Grounded Soundscape Generation for 3D Gaussian Worlds. arXiv:2608.00463 (2026), preprint. https://arxiv.org/abs/2608.00463
11. Celerier, J.-M., Baltazar, P., Bossut, C., Vuaille, N., Couturier, J.-M., Desainte-Catherine, M.: OSSIA: Towards a Unified Interface for Scoring Time and Interaction. In: TENOR, pp. 81–90 (2015). https://www.tenor-conference.org/proceedings/2015/13-Celerier-OSSIA.pdf
12. Arias, J., Celerier, J.-M., Desainte-Catherine, M.: Authoring and Automatic Verification of Interactive Multimedia Scores. Journal of New Music Research 46(1), 15–33 (2017). https://doi.org/10.1080/09298215.2016.1248444
13. Riedl, M. O., Young, R. M.: Narrative Planning: Balancing Plot and Character. Journal of Artificial Intelligence Research 39, 217–267 (2010). https://faculty.cc.gatech.edu/~riedl/pubs/jair.pdf
14. Bloem, R., Könighofer, B., Könighofer, R., Wang, C.: Shield Synthesis: Runtime Enforcement for Reactive Systems. In: TACAS 2015. LNCS, vol. 9035, pp. 533–548. Springer, Berlin, Heidelberg (2015). Extended version: arXiv:1501.02573v2. https://doi.org/10.1007/978-3-662-46681-0_51
15. Ligatti, J., Bauer, L., Walker, D.: Edit Automata: Enforcement Mechanisms for Run-Time Security Policies. International Journal of Information Security 4(1–2), 2–16 (2005). https://doi.org/10.1007/s10207-004-0046-8
16. Gaver, W. W.: What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception. Ecological Psychology 5(1), 1–29 (1993). https://doi.org/10.1207/s15326969eco0501_1
17. Gygi, B., Shafiro, V.: The Incongruency Advantage for Environmental Sounds Presented in Natural Auditory Scenes. Journal of Experimental Psychology: Human Perception and Performance 37(2), 551–565 (2011). https://doi.org/10.1037/a0020671

18. Slater, M.: Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philosophical Transactions of the Royal Society B 364(1535), 3549–3557 (2009). https://doi.org/10.1098/rstb.2009.0138
19. Allen, J. B., Berkley, D. A.: Image method for efficiently simulating small-room acoustics. Journal of the Acoustical Society of America 65(4), 943–950 (1979). https://doi.org/10.1121/1.382599
20. Scheibler, R., Bezzam, E., Dokmanić, I.: Pyroomacoustics: A Python package for audio room simulation and array processing algorithms. In: ICASSP, pp. 351–355. IEEE (2018). https://doi.org/10.1109/ICASSP.2018.8461310
21. World Wide Web Consortium: Web Audio API. W3C Recommendation, 17 June 2021. https://www.w3.org/TR/2021/REC-webaudio-20210617/
22. Nie, X., Wager, S.: Quasi-oracle estimation of heterogeneous treatment effects. Biometrika 108(2), 299–319 (2021). https://doi.org/10.1093/biomet/asaa076
23. Dudík, M., Langford, J., Li, L.: Doubly Robust Policy Evaluation and Learning. In: ICML (2011). https://arxiv.org/abs/1103.4601v2
24. Thomas, P., Theocharous, G., Ghavamzadeh, M.: High-Confidence Off-Policy Evaluation. Proceedings of the AAAI Conference on Artificial Intelligence 29(1) (2015). https://doi.org/10.1609/aaai.v29i1.9541
25. Kingsense Electronics: KS EEG102 Module BLE Communication Protocol V2.x (in Chinese). UG-KSEEG102W-BLE-V2.2-CN(2026C), rev. 2.2, pp. 6, 14 (June 2026).
26. Kingsense Electronics: KSEEG Module SDK Documentation V2.x (in Chinese). UG-KSEEG-SDK-V2.1-CN(2025A), rev. 2.1, pp. 5, 11 (November 2024).
27. Eyring, C. F.: Reverberation Time in "Dead" Rooms. Journal of the Acoustical Society of America 1(2A), 217–241 (1930). https://doi.org/10.1121/1.1915175
28. Robins, J. M., Hernán, M. A., Brumback, B.: Marginal structural models and causal inference in epidemiology. Epidemiology 11(5), 550–560 (2000). https://doi.org/10.1097/00001648-200009000-00011
29. International Telecommunication Union: Method for the subjective assessment of intermediate quality level of audio systems. Recommendation ITU-R BS.1534-3 (October 2015). https://www.itu.int/rec/R-REC-BS.1534-3-201510-I/en
30. Qian, T., Walton, A. E., Collins, L. M., et al.: The microrandomized trial for developing digital interventions: Experimental design and data analysis considerations. Psychological Methods 27(5), 874–894 (2022). https://doi.org/10.1037/met0000283
31. Ros, T., et al.: Consensus on the reporting and experimental design of clinical and cognitive-behavioural neurofeedback studies (CRED-nf checklist). Brain 143(6), 1674–1685 (2020). https://doi.org/10.1093/brain/awaa009