Hall of Monitors

Coordination Without Identical Information

Flyxion

August 2026

Abstract

A working band offers an unusually clean architecture for a problem most personalized systems get wrong. Each musician hears a different mix through their monitor, and the differences are not a failure of shared experience but the condition of it: the drummer needs the bassist and a little vocal, the singer needs their own voice and the keyboard, the guitarist wants almost none of themselves. This essay treats that arrangement as a general theory of coordination rather than a curiosity of stage engineering. It develops a distinction between a hall of monitors, in which personalized projections remain recoverable views of one shared state, and a hall of mirrors, in which projections quietly become the reality against which further projections are built. It then introduces a result with more reach than ordinary personalization theory: producing a signal and needing to perceive it are different scheduling problems, coordinated through their coupling rather than their identity, and a system's perception routing can be understood as local, dynamic manipulation of what each participant's noise floor allows them to detect. The essay closes by asking who, if anyone, is positioned to notice when a hall of monitors has begun to close on itself.

1 The Monitor as Three Things at Once

Stand near the stage before a show and look at what each musician is actually listening to. It is almost never the same thing the audience will hear a few minutes later, and it is almost never the same thing the musician beside them is listening to either. The drummer's monitor carries mostly bass and a thread of vocal, enough to lock the groove to the song without losing the beat under a wash of everything at once. The singer's monitor carries their own voice, pushed forward and close, along with the keyboard part that gives them their pitch. The guitarist, counterintuitively, often wants almost none of their own instrument in the mix and a great deal of the vocal and the drums, because they already know what their hands are doing and need the rest of the band to know where to land. Ask any working musician why the room sounds nothing like their in-ear mix and the answer is never confusion. It is design.

This arrangement is easy to treat as a piece of stagecraft with no application beyond itself, a practical solution to a narrow acoustic problem. The purpose of this essay is to argue that it is something more general: a working example of an architecture that most systems claiming to coordinate people, from social platforms to distributed software, either fail to build or actively avoid building, sometimes without realizing they are avoiding it.

Part of what makes the band's solution worth taking seriously is a small linguistic accident that turns out not to be an accident at all. The device each musician listens through is called a monitor, and the word already carries three distinct meanings that ordinary usage keeps separate without much effort. A monitor is something a person listens or watches through, an interface between a person and a larger signal. A monitor is also something that displays a state, a screen, a readout, a representation of a system's condition offered up for inspection. And a monitor is a process that watches another process, checking it against some expectation, the sense the word carries in computing when one program supervises another, and the sense it carries in the older and plainer institutional use, a person posted to oversee a room rather than to teach in it. This essay treats the first three senses as one idea wearing three names rather than as a coincidence of vocabulary. The fourth sense deserves more care than simply being set against the other three, because a supervisory process is often exactly what a system needs; a watchdog that checks whether something has failed is not, on its own, doing anything wrong. What separates a healthy fourth sense from an unhealthy one is not supervision itself but its direction. A monitor that watches on a participant's behalf, so that what it notices becomes something the participant can also come to perceive, remains a monitor in the fuller sense this essay wants to defend. A monitor that watches a participant without ever making anything more perceptible to them, so that the system sees the participant while the participant cannot see the system, has drifted into something else, still called by the same name. A stage monitor is simultaneously the first three: it is

what the musician listens through, it displays a version of the performance's current state, and it lets that musician check their own playing against what the rest of the band is doing. Nobody is merely being watched by it. The distinction matters enough to return to later, once this essay has the vocabulary to say precisely what separates watching for a participant from watching of one, but it is worth naming here, at the outset, that not every use of the word has carried the same direction. Separating the first three functions, treating what a person hears as unrelated to what a system displays as unrelated again to what gets watched for drift or failure, is exactly the move that lets coordination quietly reverse that direction without anyone noticing, and a large part of what follows is an attempt to show why keeping the three together is not merely convenient but necessary.

The claim this essay will build toward is that coordinated systems, of almost any kind, do not require their participants to receive identical representations of what is happening. They require something more specific and considerably harder to design for: that whatever each participant receives remain a recoverable projection of one shared event, rather than becoming, for that participant, the event itself.

2 The Shared State and Its Projections

Return to the band, now with enough distance from the stage to describe what is actually happening rather than simply what it feels like to stand near it. At any moment there is one continuing event, the performance itself, which can be written as a state evolving in time,

$$S(t).$$

This is not a mystical object. It is, in the most literal sense, the sound actually being produced in the room: every instrument, every voice, at their true relative levels, unfiltered by anyone's particular monitor mix. Nobody in the band experiences $S(t)$ directly in this pure form, and this turns out not to matter, because nobody needs to.

Each musician instead receives a projection of that state, constructed for them specifically,

$$M_j(t) = \Pi_j\big(S(t)\big),$$

where Π_j is whatever operator the monitor engineer, or increasingly an automated mixing system, applies to build participant j's particular mix from the full performance. The drummer's projection emphasizes bass and a little vocal. The singer's emphasizes their own voice and the keyboard. Two participants can receive projections that share almost nothing in common, level for level, instrument for instrument, and the band can still play together

flawlessly, because nothing about playing together ever required

$$M_i(t) = M_j(t).$$

That equation was never the goal, and treating it as though it were is precisely the mistake this essay wants to name and set aside early, because a great deal of what passes for coordination technology, in software as much as in audio, is built as though sameness of representation were the same thing as successful coordination, when the two have never actually been the same claim.[1]

What the band does require, instead of sameness, is compatibility. Each participant's projection has to remain usable as a guide back to the shared performance, close enough to what everyone else is playing that timing holds, pitch holds, and the piece stays recognizably one piece rather than several. Write this as a threshold condition on the whole set of projections at once,

$$\mathrm{Compat}\big(M_1(t), \ldots, M_n(t); S(t)\big) \geq \theta_{\mathrm{coord}},$$

a claim not about any single mix but about whether the collection of mixes, taken together, still functions as a working set of views onto the same underlying event. The threshold θ_{coord} is deliberately left unspecified here; what matters for now is only that it is a property of the whole ensemble of projections relative to $S(t)$, not a property that could be checked by comparing any two mixes to each other in isolation.[2] Two musicians could be hearing almost entirely different things and still both satisfy compatibility, because compatibility was never about resemblance between M_i and M_j. It was always about each one's continued fidelity to S.

Stated this plainly, the claim can sound almost too modest to be worth an essay. Of course a drummer and a singer hear different things; nobody would expect otherwise, and no one thinks this constitutes a problem for the band. But this modest observation, taken out of the concert hall and applied to the systems that increasingly mediate how people encounter shared information, work, and each other, turns out to cut directly against a design

[1]The claim that shared understanding does not require identical representations has a real precedent in the study of conversation. Herbert Clark and Susan Brennan's account of grounding treats successful communication as resting on a jointly maintained common ground of mutual belief, established and updated turn by turn, rather than on both parties holding identical mental contents (Clark and Brennan, 1991). This essay's compatibility condition is a different object, a property of simultaneous, continuously updated projections of one state rather than of sequential conversational contributions, but the underlying move, replacing sameness with a weaker and more useful joint condition, is not original to this essay and should be credited to that literature.

[2]Treating θ_{coord} as a single scalar is a simplification adopted for exposition. Since $S(t)$ is properly a trajectory through a high-dimensional state rather than a single value, compatibility across projections may in practice be closer to a topological condition than a metric one, something closer to maintained phase-locking or path-consistency across overlapping local views than to a distance falling under a number. Nothing in this essay depends on resolving that question; the scalar threshold is adequate for every claim made here.

assumption those systems rarely examine: that fairness, transparency, or genuine sharedness requires giving everyone the same feed. The remainder of this essay is an attempt to show why that assumption is not merely unnecessary but sometimes the very thing standing between a system and real coordination, and, just as importantly, to show what goes wrong when the difference between a projection and a substitute for reality is allowed to blur.

3 Hall of Monitors, Hall of Mirrors

The word projection has been doing a great deal of quiet work in the previous section, and it is worth making explicit exactly what it requires, because the requirement is easy to state and easy to lose sight of in practice. For $M_j(t)$ to be a projection of $S(t)$ rather than merely something derived from it once and then left to drift, it has to remain possible, at least in principle, to move from the projection back toward the thing projected. Write this as

$$M_j \rightsquigarrow S,$$

a recoverability relation rather than an equation: a participant holding only their own monitor mix can still, given the opportunity, compare it against the room, ask why their mix differs from someone else's, request a change, and in general treat their projection as one partial view among others rather than as the whole of what there is to know.[3] A hall of monitors, in the sense this essay wants the phrase to carry, is any coordinated system in which this relation holds for every participant at once. Personalization is present, sometimes extensively so, but personalization never severs the projection from its source.

This needs one immediate qualification, because stated baldly it risks contradicting the very thing a monitor mix is for. Every projection discards information; that is what makes it a mix rather than a copy. The drummer's monitor throws away almost everything about the guitar's tone and most of the room's ambient detail, and this is not a defect to be minimized but the entire point of building the projection in the first place. So recoverability, properly understood, cannot mean recovering S exactly. It has to mean something narrower: recovering S up to whatever distinctions actually matter for the participant's continued

[3] An earlier formalization of this relation asked for an approximate left inverse recovering S itself, within some error ϵ. That asks too much, and the equivalence-class discussion below explains why: since Π_j is deliberately lossy, there need be no meaningful metric in which a reconstruction stays close to the original S, and two acoustically very different states can be entirely equivalent for a given participant's next action. The more appropriate object is the coordination quotient $\mathcal{S}/\!\sim_{\text{coord}}$, whose elements identify shared states that are interchangeable with respect to a participant's admissible continuation. Recoverability then requires a reconstruction map $\rho_j : M_j \to \mathcal{S}/\!\sim_{\text{coord}}$ such that $\rho_j(\Pi_j(S)) = [S]_{\text{coord}}$, not a reconstruction of S itself. A hall of mirrors emerges when successive projections cease to admit such a map: states belonging to different coordination classes become aliased into the same observable projection. Exact recovery of S is neither required nor, in a genuinely selective monitor, normally desirable.

ability to coordinate. Write the distinctions a given projection erases as an equivalence relation,

$$S \sim_j S' \quad \Longleftrightarrow \quad \Pi_j(S) = \Pi_j(S'),$$

two states that look identical once passed through participant j's particular mix. A good monitor is supposed to induce a great deal of this equivalence; collapsing irrelevant distinctions is what makes a mix usable at all rather than an unfiltered flood. It fails only when the distinctions it collapses turn out to be exactly the ones that mattered, so that recoverability becomes not exact reconstruction but membership in the right equivalence class,

$$\rho_j(M_j) \in [S]_{\text{coord}}.$$

This gives the hall of mirrors its sharpest available definition, sharper than the informal sequence of arrows from a moment ago. Every hall of monitors discards information; that alone never distinguishes it from a hall of mirrors. What distinguishes a mirror hall is that its successive projections discard, along with the irrelevant detail, the ability to tell which discarded distinctions were the ones that mattered. A monitor can lose enormous amounts of S and remain entirely legitimate. It becomes a mirror only once nobody, including the system itself, can any longer say what was thrown away that shouldn't have been.

This also supplies a precise name for the specific way a monitor fails, sharper than simply calling it lossy, since almost every representation worth having is lossy. A monitor is a selectively lossy projection whose discarded distinctions are exactly the ones irrelevant to a participant's admissible continuation. It stops being one, and becomes dangerous rather than merely partial, the moment it collapses two states that actually called for different responses:

$$S_a \not\sim_{\text{coord}} S_b \qquad \text{but} \qquad \Pi_j(S_a) = \Pi_j(S_b).$$

The participant receives the identical perceptible signal in two situations that required them to behave differently, and has no way of knowing, from inside their own mix, that the situations ever differed. This is a sharper failure than ordinary information loss, closer to what signal processing would call aliasing than to mere compression: not an absence of detail but a false equivalence, two distinguishable realities folded into one indistinguishable appearance. It is this specific failure, rather than data loss in general, that the rest of this essay is ultimately concerned with detecting and preventing.

The whole criterion this section has been building toward compresses into a single inclusion between two equivalence relations. Write $S \sim_j S'$ for the relation a given monitor actually induces, two states the projection cannot tell apart, and $S \sim_{\text{coord}} S'$ for the relation coordination can actually tolerate, two states a participant is genuinely permitted to treat

alike. A monitor is valid exactly when the first is contained in the second,

$$\boxed{\sim_j \subseteq \sim_{\text{coord}} \cdot}$$

Everything the projection makes indistinguishable must also be something coordination already permitted to be indistinguishable; nothing more. This also resolves a question that could otherwise make the essay sound as though it were quietly arguing for maximal information at all times, which was never the intention. A monitor that preserves too little collapses states that should have remained separate, $\sim_j \not\subseteq \sim_{\text{coord}}$, and becomes dangerous in exactly the way this section has described. A monitor that preserves too much, one whose $\sim_j$ is far finer than coordination actually requires, is not thereby safer; it has simply failed to do the one thing a monitor is for, which is selection. The target was never maximum information or minimum information. It was always the inclusion above, holding as tightly as the participant's actual coordination needs allow.[4]

The failure mode this essay is most concerned with is not the mere existence of personalized views. It is what happens when the arrow above quietly reverses, when a projection stops functioning as a view onto something shared and starts functioning as the thing itself, so that further projections are built not from S but from an earlier projection:

$$S \to M_j \to M_j' \to M_j'' \to \cdots ,$$

each step a plausible, often well-intentioned refinement of the step before it, and each step a little further from anything an outside party could check against the original. Call this a hall of mirrors. Nothing about any single step in that sequence needs to look like a malfunction. A recommendation system refining itself against a user's past clicks, a curated feed adjusting to what a user lingered on, an assistant learning a person's preferences from their own prior requests: each of these is, locally, a reasonable design choice, the kind of adaptive refinement that makes a system feel more useful over time. What accumulates across the whole sequence is not any single bad decision but the gradual disappearance of S as an independently checkable object. By the time enough steps have passed, there may be no remaining position, for the participant or for anyone else, from which the current projection

[4]This gives the essay's central criterion a natural place beside two older results it does not depend on but sits comfortably alongside. Classical observability, in control theory, asks whether a system's internal state can be reconstructed from its outputs, schematically $y(t) \rightsquigarrow x(t)$. This essay's criterion is deliberately weaker: it does not require that $S(t)$ be reconstructable from $M_j(t)$, only that $[S(t)]_{\text{coord}}$ be, observability modulo an action-relevant equivalence rather than observability of the full state. W. Ross Ashby's law of requisite variety, that a regulator must possess at least as much variety as the disturbances it is meant to control (Ashby, 1956), is the same demand read from the other direction: a monitor cannot quotient away more variety than coordination actually requires it to preserve, which is close to a representational cousin of Ashby's claim rather than a restatement of it.

could be compared against the shared state it was originally supposed to represent, because nothing outside the sequence of projections has been kept.

This gives the essay's title its full meaning, and it is worth stating the distinction as plainly as the architecture allows. A hall of monitors has windows. Standing in it, a participant can look through their own window and, with effort or assistance, look through another's, or check either window against the room itself. A hall of mirrors has no windows at all, only surfaces that reflect other surfaces, each one a plausible continuation of the one before it and none of them opening onto anything beyond the sequence. The two structures can be built from identical components, the same personalization machinery, the same per-participant weighting, the same adaptive refinement that makes a hall of monitors work so well for the band. What separates them is not mechanism but whether S remains recoverable at the end of the chain, and this is precisely why the failure is so easy to miss from inside it: nothing about a mirror hall announces itself as one. It simply stops being possible, at some point nobody marked, to ask what the room actually sounded like.

This also supplies the essay's sharpest possible statement of what a legitimate personalized system owes its participants, considerably sharper than the vague appeal to transparency that discussions of platform personalization usually settle for:

> A common world does not require a common mix, but a different mix must remain a mix of something common.

The first half of that sentence is what makes the band possible at all, and what this essay has spent its opening section defending against an unexamined intuition that fairness requires sameness. The second half is the condition a hall of monitors has to keep satisfying, turn after turn, in order not to become a hall of mirrors, and it is considerably harder to satisfy than it looks, because satisfying it requires more than good intentions on the part of whoever is building the projections. Making that relation durable, however, depends on a distinction the abstract geometry cannot supply by itself: whether the route back to the shared state is merely preserved somewhere in the system or remains practically available to the participant whose projection depends on it.

4 Retained versus Recoverable, Reprised

A companion essay to this one, concerned with personal agents, social platforms, and the conditions under which mediation becomes difficult to reverse, arrives at a distinction worth importing here directly, because it names precisely the gap that separates a hall of monitors from a hall of mirrors in practice rather than only in principle:

$$\text{retained by the system} \neq \text{recoverable by the participant.}$$

A platform can preserve an enormous, faithful, internally consistent record of $S(t)$ on its own servers while giving every individual participant no practical way to reach that record themselves. Nothing about the earlier description of a mirror hall required that S actually vanish. It only required that S stop being reachable by anyone standing inside the hall. A system can therefore satisfy every engineering standard for data retention, keep flawless internal logs, run sophisticated ranking models with fully reconstructable histories, and still produce exactly the epistemic fragmentation this essay has been describing, because retention was never the property at stake. Recoverability was.

This matters because it is tempting to read the hall of monitors versus hall of mirrors distinction as a claim about how much a system remembers, and to conclude that a well-resourced platform, one that keeps everything, must therefore be safely on the monitor side of the line. The opposite can easily be true. A system can be an excellent archive and a terrible hall, holding a perfectly coherent S in a form nobody outside its own infrastructure could ever consult, compare, or use to check their own projection against. From a participant's seat, a platform with superb internal recoverability and zero external recoverability is operationally indistinguishable from one that kept nothing at all. Both leave the participant with only M_j, no path back to S, and no way of knowing whether their projection has quietly become their entire world.

The relationship between the two essays is worth stating plainly rather than leaving implicit. The companion essay is concerned with what happens to a single person's relationship to a single activity once mediation has gone far enough that returning to independent practice is no longer practically available to them: it asks about a state, a threshold crossed, a margin that has gone negative. This essay is concerned with something that happens earlier and at the level of the whole system rather than any one person's history: the architecture by which a shared state is distributed into projections in the first place, and the point at which that architecture stops preserving the path back for anyone at all. One essay describes the condition of having already lost the way home. This one describes the floor plan that determines whether a way home was ever built into the structure to begin with. A hall of monitors that has quietly become a hall of mirrors is exactly the kind of architecture in which the companion essay's failures become common rather than exceptional, because once S is unreachable for everyone, there is no longer any independent position from which a participant's drift away from it could even be measured, let alone reversed.

5 Two Scheduling Problems, Not One

Return once more to the band, this time to ask a question the earlier sections have quietly skipped over. Somebody has to decide what goes into $S(t)$ in the first place, which is a

different question from how $S(t)$ gets projected into each participant's monitor. The guitarist has to actually play the guitar part; the drummer has to actually keep time. Call this production: the assignment of who realizes which component of the shared state. Separately, and this is the part most personalization systems collapse into the same question without noticing they have done so, somebody has to decide what each participant needs to perceive in order to do their own part well, which need not have anything to do with who is producing what. These are two different scheduling problems, and treating them as one is where a great deal of naive personalization design goes wrong before it ever reaches the harder questions this essay is building toward.

Production scheduling can be written as an assignment problem. Let $\mathrm{Aff}(p_i, x_k, r)$ measure participant p_i's affinity for realizing component x_k in role r, capturing whatever mix of skill, practice, and current capacity makes one person better suited than another to actually produce a given part. Production scheduling then selects

$$P(x_k, r) = \arg\max_{p_i} \mathrm{Aff}(p_i, x_k, r),$$

assigning each component of the performance to whoever is best equipped to realize it. This is the ordinary sense in which a band divides labor, and nothing about it is especially novel; every ensemble, human or otherwise, solves some version of this problem constantly, usually without needing to name it.

Perception scheduling asks a genuinely different question, and it is here that the earlier discussion of monitors as displays and monitors as supervisory processes becomes useful again rather than merely historical. Define $Q(p_i, x_k, t)$ as the usefulness, to participant p_i at time t, of having component x_k made salient in their own projection, regardless of whether p_i has anything to do with producing x_k. The resulting monitor mix is then a weighted combination,

$$M_i(t) = \sum_k w_{ik}(t)\, x_k(t), \qquad w_{ik}(t) \propto Q(p_i, x_k, t),$$

and there is no reason at all to expect

$$P(x_k, r) = \arg\max_{p_i} Q(p_i, x_k, t).$$

The drummer does not produce the vocal line, and yet a thread of vocal in the drummer's monitor is often exactly what lets the drummer's own playing stay locked to the song. The singer does not produce the keyboard part, and yet the keyboard, pushed forward in the singer's ear, is often what gives the singer their entrance. A well-functioning ensemble depends on production and perception routinely disagreeing about who needs what, not on the two graphs happening to coincide. This gives the essay two distinct structures over the

same set of participants,

$$G_{\text{produce}} \neq G_{\text{perceive}},$$

with successful coordination arising from how the two are coupled rather than from either one alone,

$$G_{\text{coord}} = G_{\text{produce}} \bowtie G_{\text{perceive}}.$$

[5] Most systems that describe themselves as coordinating people, in software as much as in music, only ever build the first graph, deciding who does what, and quietly assume the second graph will take care of itself, or worse, silently collapse the two together, on the unexamined premise that whoever is doing a thing must therefore be the person who most needs to perceive it. The band's monitor engineer knows better, because the cost of getting it wrong is audible within seconds.

It is worth making the perception side of this precise rather than leaving $Q(p_i, x_k, t)$ as an unexplained usefulness score, because it turns out to have a concrete mechanical interpretation rather than a merely intuitive one. Each participant can be understood as carrying a local detection threshold, a noise floor N_j, below which a signal, however real, fails to register as usable information for that participant at that moment. This is not a metaphor borrowed loosely from audio engineering for the occasion; it names a structure that recurs across a wide range of otherwise unrelated systems, wherever something can be present without being detectable. A fault-related signal in a mechanical system can exist below the threshold that separates it from ordinary operating noise, so that a healthy machine and a failing one are, for a time, observationally indistinguishable to an instrument that is not specifically tuned to notice the difference. A weak but genuine signal in a saturated information environment can exist without being locatable against a background of louder, more confident, less reliable material. In each of these cases the same distinction is doing

[5]The nearest existing framework to this distinction comes from human-automation research on function allocation and levels of automation, beginning with Sheridan and Verplank's scale of human supervisory control (1978) and extended by Parasuraman, Sheridan, and Wickens into a model spanning separate stages of information acquisition, analysis, decision, and action (2000). That literature asks, stage by stage, how much of a task should be allocated to a human versus a machine. It does not, so far as this author is aware, separate who produces a component of a shared task from who most needs that component made perceptible, as two independently varying graphs coordinated through their coupling rather than collapsed into one allocation decision. That separation, and the claim that a well-functioning system depends on the two graphs disagreeing rather than coinciding, is this essay's own contribution.

the work:[6]

$$\text{existence} \neq \text{transmission} \neq \text{detectability} \neq \text{recoverability}.$$

Something can be real, and even successfully transmitted into a shared environment, without crossing the threshold that would make it detectable to any particular observer, and detectability itself is not the same claim as recoverability, since a signal noticed once can still be lost again if nothing preserves the path back to it. The room in this essay is simply one more setting in which that same structure appears, not a novel discovery specific to audio.

Perception routing, understood this way, is threshold manipulation rather than mere emphasis. The role of Π_j is not only to select which components of $S(t)$ appear in participant j's mix, but to decide, moment to moment, which components cross that participant's own local floor:

$$x_k < N_j \quad \Longrightarrow \quad \Pi_j(S) \text{ may amplify } x_k \quad \Longrightarrow \quad x'_k > N_j.$$

The information was never absent from $S(t)$. What determines whether it becomes perceptible to a given participant is the projection built for them, not the presence or absence of the signal itself. A good monitor mix does not make everything louder; that would only raise the floor along with everything above it and leave the relative problem unsolved. It selectively raises what a particular participant needs to cross their own threshold at a particular moment, and lets everything else recede, which is a considerably more demanding design target than uniform amplification and a considerably more useful one. This is the sense in which perception scheduling is not a secondary feature bolted onto production scheduling but a genuinely separate discipline. Perception scheduling therefore has its own logic and its own failure modes, but its most important property may be that a correct solution at one moment need not remain correct at the next. The signal that has to cross a participant's noise floor now may be precisely the signal that should recede once coordination has been restored.

[6]The technical vocabulary of a detection threshold below which a real signal fails to register has a precise origin. Signal detection theory, developed by David Green and John Swets, formalizes the separation between a signal's presence and an observer's ability to detect it against background noise, treating detection as a statistical decision problem rather than a fixed sensory limen (Green and Swets, 1966). The auditory case specifically traces to Harvey Fletcher's work on critical bands, which established that a tone can be rendered undetectable by masking noise confined to a narrow band around its own frequency, independent of noise anywhere else in the spectrum (Fletcher, 1933, refined 1940). This essay's noise floor N_j borrows the concept, not the mathematics, from that literature, and generalizes it to an arbitrary participant's threshold for an arbitrary component of a shared state, which is this essay's own extension rather than a claim already made in the psychoacoustic literature.

6 The Monitor as Control Surface, Not Profile

It would be easy, having established that each participant needs a different projection, to conclude that the design problem is now solved once each participant has been assigned a suitable weighting and left alone. This is the point at which most personalization systems stop, and it is also the point at which they begin to drift, quietly, toward the mirror hall this essay has been warning against, because a fixed weighting is not actually what makes the band's monitor system work. What makes it work is that the weights change, continuously and in response to the moment, rather than settling into anything that could reasonably be called a profile.

Write the weights as depending not only on the participant but on time, on the present state of the shared performance, and on that participant's own accumulated history:

$$w_{ij}(t) = w_{ij}\big(t, S_t, \mathcal{H}_i\big).$$

When a singer approaches a difficult entrance, the instrumental cue that historically helps them find it can rise sharply in their monitor for the few seconds that matter and recede once the entrance has landed. When a participant begins drifting from the tempo, whatever reference signal has, for that particular person, reliably helped them recover can be pushed forward without anyone needing to diagnose the drift in the moment; the system can simply have learned, from $\mathcal{H}_i$, what tends to work for this person when this kind of deviation appears. None of this resembles a static preference profile, the kind of thing a conventional personalization system builds once from past behavior and then applies uniformly thereafter. It resembles something closer to an instrument a skilled engineer is actively playing, a control surface responding to what is happening right now rather than a fixed setting responding to what happened, on average, at some point in the past.

This distinction connects directly to a criterion developed elsewhere in this author's work, in an essay concerned with what allows a process to continue moving through a sequence of admissible states rather than becoming stuck at one. The relevant question there is not simply what a process currently has access to, but which of its possible next moves remain genuinely available to it, a criterion oriented toward the future rather than only toward the present. Applied here, this reframes what a monitor mix is actually for. The design question is not, and should not be, what would this participant enjoy hearing more of. It is something considerably more demanding:

Which projection of the shared state best preserves this participant's next admissible continuation?

A monitor mix judged only by immediate preference will tend, over time, toward whatever

is easiest to enjoy in the moment, which is not at all the same target as whatever keeps a participant able to do what they need to do next. The drummer does not necessarily want more vocal in their monitor; they need it, because without it their next several bars will drift out from under the song. A system optimizing for stated or inferred preference alone has no reliable way to distinguish these two cases, and a system that cannot distinguish them will, eventually, optimize away the very cues that kept participants capable of continuing rather than merely comfortable in the present. Treating the monitor as a control surface rather than a profile is what keeps this from happening: the projection is built new, each moment, from where the participant actually is and where they need to be able to go next, rather than inherited wholesale from where they have already been.

7 Who Holds the Faders?

Everything so far has treated the projection operator Π_j as though it simply exists, arriving at the participant's monitor already built. This was a useful simplification for developing the geometry, and it is no longer an honest one to leave standing, because a projection is always made by somebody, and who that somebody is turns out to matter as much as anything this essay has formalized about the projection itself.

In a band, Π_j has an author: a sound engineer, or increasingly software standing in for one, adjusting levels at a console. But the crucial fact about the band is not merely that an author exists. It is that the participant has a channel back to that author. The drummer can ask for more vocal. The singer can ask for less of their own reverb. A musician who finds their monitor unworkable can pull an earpiece out entirely and rely on the room. Write this as a loop rather than a one-way arrow,

$$p_j \longrightarrow \alpha \longrightarrow \Pi_j,$$

where α denotes whoever, or whatever, holds the authority to shape a given projection, and the loop closes because p_j, the participant, has a standing way to act on α. This loop is what makes the earlier claim that personalization in a band is negotiated rather than imposed actually true, and it is easy to take for granted precisely because it costs the band nothing to provide. The engineer has expertise the musician lacks; the musician has situated, moment-to-moment knowledge of what they need that the engineer cannot fully anticipate. The mix that results belongs to neither party alone.

Many of the larger halls this essay has gestured toward do not preserve this loop. A platform, a curated feed, an automated classroom scheduler can each build a Π_j of considerable sophistication while offering the participant on the receiving end of it no analogous channel

back,

$$\alpha \longrightarrow \Pi_j \longrightarrow M_j,$$

authority flowing one direction only. This is a profound architectural difference even in a case where the resulting M_j happens, at some particular moment, to look identical to what a negotiated loop would have produced. Two systems can present a participant with the same mix today and differ entirely in what happens when that participant wants something different tomorrow.

It is worth being more precise than simply asking whether a participant has access to their own projection, because access itself is not one capability but several, and a system can grant some while withholding others in ways that matter enormously. Define the set of actors capable of modifying a given projection,

$$\mathcal{A}(\Pi_j) = \{\text{actors with the standing ability to change } \Pi_j\},$$

and then distinguish, for a given participant, whether they can merely inspect their own projection, modify its parameters directly, contest a projection they believe is wrong, bypass it in favor of something closer to S itself, or compare it meaningfully against another participant's. A participant might know with complete clarity that their feed is personalized, might even be shown a plausible explanation of why, and still possess none of the other four capabilities, which is a considerably weaker position than the drummer's, however sophisticated the explanation offered to them. Knowing that a projection exists is not the same as having any standing over it, and a hall of monitors that has quietly stopped offering the loop above has, whatever its other virtues, stopped being a band and started being something closer to an audience being told, with increasing precision, what it is currently allowed to hear.

8 Who Verifies the Hall Hasn't Become a Mirror

Section 1 named a fourth sense of the word monitor and asked what separates watching for a participant from watching of one. That distinction can now be given its full weight, because it names exactly the missing piece in everything this essay has built so far. A hall of monitors, in the generous sense this essay has been developing, requires more than good projections and more even than the negotiated authority described in the previous section. It requires that someone, or something, remain positioned to check whether the projections are still projections at all, or whether the architecture has already begun sliding toward the mirror hall described earlier, and this final requirement turns out to be the hardest one to satisfy honestly.

A working band solves this problem almost for free, in a way that is easy to overlook precisely because it costs nothing extra. Anyone in the room, a sound engineer, a bandmate between parts, an audience member near the stage, can simply listen to what the room actually sounds like and compare it against what any individual musician reports hearing in their monitor. The shared state $S(t)$ is not hidden behind the projection architecture; it is sitting in the open air, available to any ear willing to attend to it. This is what makes the band such a forgiving case to reason about, and it is also what makes the band a misleading model for larger or more thoroughly mediated systems, because this free, ambient verification is not a general property of hall-of-monitor architectures. It is a specific gift of acoustics in a small room, and nothing in the formal structure of $M_j(t) = \Pi_j(S(t))$ guarantees anything like it elsewhere.

Consider a much larger hall: a platform serving millions of personalized feeds, a distributed software system routing information among processes that never directly observe one another, a classroom in which each student receives a differently paced version of the same material, or a multi-agent system in which no single component ever holds the complete state at once. In each of these, the room-air verification that saves the band is simply unavailable. No participant can casually listen past their own projection to check it against $S(t)$, because there is no equivalent of standing near the stage. Whatever confirmation exists that the architecture remains a hall of monitors rather than a hall of mirrors has to be built deliberately into the system, as an actual capability held by someone, rather than assumed as an ambient byproduct of the setting.

It helps to notice that recoverability was never a single property to be had or lacking; it is indexed by who is doing the recovering, and different indices can come apart in ways that matter enormously. Distinguish at least three relations. Participant recoverability, $M_j \rightsquigarrow_j [S]_{\text{coord}}$, asks whether the person holding the projection can themselves reach its coordination class. System recoverability, $M_j \rightsquigarrow_{\text{sys}} [S]_{\text{coord}}$, asks only whether the platform that built the projection retains enough to reconstruct it internally. Externally situated recoverability, $M_j \rightsquigarrow_{\text{ext}} [S]_{\text{coord}}$, asks whether some party independent of whoever produced M_j in the first place can also reach it.

Recoverability is indexed by who is doing the recovering.

The distinction retained versus recoverable, imported earlier from this essay's companion piece, is one instance of this larger principle rather than the whole of it: a system can satisfy system recoverability magnificently while offering the participant none, and a system can even satisfy participant recoverability, letting an individual inspect their own unfiltered feed, while offering no external party any way to confirm that the inspection reflects what was actually shown at the time in question. This is precisely why auditing has to be situated

correctly to count as auditing at all. A record, however complete, only functions as evidence if whoever is checking it occupies a position independent of whoever produced it; a platform auditing its own personalization pipeline against its own retained history satisfies $\rightsquigarrow_{\text{sys}}$ while saying nothing whatsoever about $\rightsquigarrow_{\text{ext}}$, which is the relation actually at stake whenever the question is whether a hall can be trusted from outside itself. A credible architecture does not need to specify which particular mechanism, independent auditors, cryptographic attestation, some other arrangement entirely, provides this; it only needs to make some genuine $\rightsquigarrow_{\text{ext}}$ path available, one not wholly determined by the same authority responsible for Π_j in the first place. Short of that, whatever confirmation a system offers remains the mirror hall's own reflection confirming itself.

This gives the essay's warning its sharpest and most concrete form. The danger is not that a hall of monitors will announce its transformation into a hall of mirrors, because it will not; nothing about the sequence $S \to M_j \to M_j' \to M_j'' \to \cdots$ contains an alarm. The danger is specifically that the system will continue doing the one thing it was designed to do, raising signals above participants' local noise floors, while gradually narrowing which signals ever get the chance. A system under no obligation to preserve independent verification will tend, for entirely ordinary reasons of engineering convenience and cost, to keep amplifying whatever is already loud enough to be easy to detect and easy to justify amplifying further, while whatever remains below threshold stays there, not through malice but through the simple absence of anyone positioned to notice that it never crossed. The room does not get quieter. It gets louder in exactly the places it was already loud, and the quiet thing that mattered, whatever it was, never has to be silenced, because it was never given the chance to be heard in the first place.

The fourth sense of monitor, watching of a participant rather than for one, is what a hall of monitors degrades into once this verification has quietly disappeared: a structure that watches for gross malfunction while doing nothing to expand what anyone inside it can actually perceive, the system able to see the participant while the participant remains unable to see the system. That asymmetry, not supervision itself, is what separates a monitor that serves its participant from one that merely oversees them. The difference was never a difference in mechanism. It was always a difference in whether someone, somewhere, remained able to check, and in which direction the checking ran.

9 Beyond Audio

Nothing in the architecture developed here depends on sound specifically, and it is worth naming a few other settings briefly, not to argue each case in full but to indicate how far the same structure travels. A laboratory runs on exactly this pattern: instruments produce

readings, and different members of a research team need different subsets of those readings made salient at different moments, the technician watching for a specific failure signature, the principal investigator watching for a pattern across many runs, neither needing the other's full view but both needing their own view to remain answerable to the same underlying data. A cockpit or an operating room works the same way under far higher stakes, with checklists, alarms, and role-specific displays functioning as monitors in every sense this essay has used the word: something looked through, something displaying a state, something watching for deviation, ideally always watching for the people relying on it rather than merely of them, because in these settings the cost of that reversal is measured directly in harm.[7] A classroom, at its best, is a hall of monitors rather than a single broadcast: different students productively receiving different pacing, different emphasis, different scaffolding, all answerable to the same underlying curriculum rather than to isolated, uncheckable tracks. A distributed software system routing messages among processes faces the identical design choice between building real perception scheduling, so that each process receives what it needs to remain correctly synchronized with the others, and defaulting to broadcast or silence, the software equivalent of either flooding every monitor with everything or trusting that whatever isn't explicitly routed will simply not matter. And a multi-agent system, artificial or otherwise, in which no single component holds the complete state at once, faces the sharpest version of every question this essay has raised, because there the shared state $S(t)$ may never be directly observable by anyone, human or machine, making the discipline of preserving recoverable projections not a nicety but the entire design problem.[8]

The point of this brief survey is not to work out any of these cases in the depth they would individually deserve. It is to establish that the band was never a special case dressed up as

[7]The term supervisory control, in its technical sense, comes from Thomas Sheridan and William Verplank's scale describing how much of a task a human operator delegates to a machine versus retains direct control over (Sheridan and Verplank, 1978), later developed at length in Sheridan's own account of telerobotics and human oversight of automation (Sheridan, 1992). That literature is centrally concerned with where authority sits along a spectrum from fully manual to fully autonomous operation. This essay borrows the word monitor's supervisory connotation from the same general territory, cockpits, control rooms, human oversight of automated systems, but is asking a different question than that literature typically asks: not how much authority a human retains, but whether the supervisory function, whatever its level of authority, is also doing the separate work of keeping a shared state recoverable for the people depending on it, which is a claim about architecture rather than about the human-machine authority boundary that literature was built to describe.

[8]The observation that a working system's knowledge can be distributed across people, instruments, and procedures, with no single participant holding the complete representation, is the central finding of Edwin Hutchins's ethnographic study of ship navigation teams (Hutchins, 1995), where the navigating "mind" belongs to the whole sociotechnical assembly rather than to any one officer. This essay's shared state $S(t)$, never directly received by any participant in its entirety, is consistent with that picture and indebted to it. What this essay adds, rather than restates, is the specific claim that coordination under those conditions depends on keeping projections recoverable against S rather than merely on distributing the cognitive labor competently; Hutchins's framework describes the distribution, not the failure mode of that distribution collapsing into a hall of mirrors.

a metaphor. It was the clearest available instance of a structure that shows up wherever a group of participants needs to act on one shared reality without needing, or being able to bear, identical access to it.

10 Conclusion

Return, one last time, to the room before the show. Nothing about what the drummer hears, the singer hears, or the guitarist hears is the same, and nothing about that difference threatens the performance. It is what makes the performance possible. The room does not achieve coordination by handing everyone the same signal. It achieves coordination by ensuring that whatever signal each person receives remains, at every moment, a way of reaching the same underlying event, checkable against it, correctable against it, and never mistaken, by the system or by the participant, for the event itself.

That is the whole of what this essay has argued, worked out across a shared state and its projections, a distinction between a hall that keeps its windows and a hall that has quietly become only mirrors, a companion claim about what a system retains for itself versus what it actually returns to the people depending on it, two scheduling problems too often collapsed into one, a monitor treated as something played rather than something merely set, a question of who holds the authority to shape a projection and whether the participant it serves has any standing over that authority, and finally the hardest and least glamorous requirement of all: that somebody, positioned independently enough to matter, remain able to notice the difference between a hall and its own reflection, since nothing about that difference announces itself from inside.

One extension is worth naming here without being argued out, since it belongs to a more specific piece than this one. The band's monitor engineer hears everything and is trusted with that access because the reason for the sensing and the beneficiary of the sensing are the same person: the musician. A wearable device built to augment ordinary perception, hearing what its wearer hears, seeing what its wearer sees, need not preserve that alignment, and the architecture developed here gives a precise way to see why. Alongside G_{produce} and G_{perceive} sits a third graph this essay has not needed until now, G_{extract}: not who produces a signal or who needs to perceive it, but who is positioned to derive value from having sensed it at all. Nothing guarantees that the answer to that third question follows from either of the first two, and a platform whose commercial interest lies mainly in the third graph can build an excellent monitor, in every sense developed here, while quietly also becoming something else. The question this essay has spent its length teaching a reader to ask, who verifies the hall hasn't become a mirror, has a companion worth carrying forward: who is the monitor actually monitoring for. Appendix A collects the additional formalism this

extension requires, both for the single-wearer case and for what happens once several people, each wearing their own audio-only assistant, need to coordinate through a shared physical scene without any of them receiving the same signal.

A common world does not require a common mix. It requires only that every mix remain, in the end, a mix of something common.

A Extraction and Multi-User Audio Coordination

This appendix collects formal material gestured at in the essay's closing paragraph but not developed in the main text, since it belongs to a more specific treatment of wearable and multi-user audio systems than the general architecture above requires.

A.1 A Third Graph

Section 5 distinguished production, who realizes a component of $S(t)$, from perception, who needs a component made salient. A third question becomes unavoidable once the projecting device is not a mixing console but something a person wears continuously through the ordinary business of their day: who is positioned to extract value from the fact that sensing occurred at all, independent of whether that value ever returns to the wearer. Write this as a third allocation problem alongside the two already developed,

$P(x_k)$: who should produce this information? $Q(p_j, x_k)$: who needs to perceive it? $X(d_k, a)$: who

Nothing in this essay's architecture guarantees that the answer to X tracks the answers to P and Q. A system can be an excellent perception scheduler for its wearer, raising exactly the right signals across exactly the right thresholds, while the same sensor stream that makes this possible is simultaneously valuable to some other party for reasons that have nothing to do with the wearer's own admissible continuation. The trust the band's monitor engineer earns for free, because the reason for the sensing and the beneficiary of the sensing coincide, cannot be assumed here and has to be established architecturally instead. This gives the essay's verification question its sharpest possible form:

Who is the monitor actually monitoring for?

A monitor that helps a wearer hear, see, or navigate more effectively is one architecture. A monitor that additionally, or primarily, allows some other party to accumulate a model of the wearer, the people around them, and the spaces they move through is another. Nothing about the device's outward appearance need distinguish the two.

A.2 Coordination Among Private Audio Projections

A visual projection competes for the same channel it augments; a purely auditory one need not, since the world remains visually unmediated while computation occupies a separate perceptual channel. This makes an audio-only architecture worth treating on its own terms rather than as a lesser version of a visual one. Suppose several people, each wearing an independent audio assistant, share one physical scene, $S(t)$. Each receives a private auditory projection,

$$A_j(t) = \Pi_j^{\text{audio}}(S(t)),$$

and it is not enough for each assistant to optimize independently for its own wearer,

$$\max_{\Pi_j} U_j,$$

since individually sensible interventions can be collectively incompatible: one person's assistant advising them to move exactly where another's is about to guide its own wearer. The group instead needs

$$\max_{\Pi_1,\ldots,\Pi_n} \sum_j U_j(\Pi_j, S) \qquad \text{subject to} \qquad \text{Compat}(A_1, \ldots, A_n; S) \geq \theta_{\text{coord}},$$

the same compatibility condition Section 2 introduced for the band, now applied across independent devices rather than independent musicians. The assistants have effectively become monitor engineers for one another, and one signal worth routing is not always the wearer's own observation returned to themselves, but a wearer's observation routed to someone else's assistant instead, on the grounds that the second person is the one who actually needs to act on it.

This creates a natural privacy principle that follows directly from the equivalence-class criterion developed in Section 3, and it can be stated with the same precision the rest of this appendix has aimed for. An assistant coordinating with another need not, and should not, transmit its complete model of its own wearer, H_i, built from everything it has sensed on that participant's behalf. What participant j's assistant actually needs is only the coordination-relevant class of whatever has been observed, the smallest message sufficient to keep the two of them jointly coordinated,

$$m_{ij} = \arg\min_m I(m; H_i) \qquad \text{subject to} \qquad \text{Compat}\big(A_i, A_j; S \mid m\big) \geq \theta_{\text{coord}},$$

rather than

$$H_i \to \text{complete internal model} \to H_j.$$

In prose: transmit the least about oneself that is sufficient for the two of you to remain coordinated.

Coordinate conclusions without synchronizing private states.

This is not a separate privacy constraint bolted onto the architecture. It is the same operation as monitor mixing itself, applied to the channel between assistants rather than to the channel between a shared scene and one wearer: preserve whatever distinctions the recipient's continuation actually depends on, and quotient away everything else, whether the everything else is irrelevant acoustic detail or irrelevant private state. This gives the local-versus-centralized distinction real architectural teeth. A system built around this principle can perform coordination locally, exchanging only small, coordination-relevant messages between devices,

$$S_i \to H_i \to m_{ij} \to H_j,$$

while a centralized alternative has every incentive to send full sensory streams to a single point that performs coordination on the platform's own infrastructure and retains everything beyond the coordination message as a byproduct,

$$(S_1, \ldots, S_n) \to \text{platform} \to (A_1, \ldots, A_n).$$

The two architectures can, in principle, produce the same sentence in every wearer's ear at every moment. They remain opposites underneath. The first builds a hall out of monitors that coordinate with one another. The second places a single monitor above the entire hall, and everyone inside it, wearer and bystander alike, becomes something that architecture can see without any corresponding way of seeing it back.

Good mediation preserves differences in futures, not differences in representations.

Two projections can look nothing alike and still be equivalent, provided they leave the same continuations open to whoever receives them. Two projections can look nearly identical and still be dangerously inequivalent, if some difference they conceal is exactly the one that would have called for a different next action. This is the same criterion, $\sim_j \subseteq \sim_{\text{coord}}$, doing its work one final time: not at the boundary between a shared scene and a single monitor, but at the boundary between one private mind and another.[9]

[9] A companion notebook of synthetic computational explorations, illustrating projection compatibility, aliasing, noise-floor routing, and disclosure/coordination frontiers among other things, is available separately at the author's repository. It makes several of this essay's formal distinctions computationally manipulable;

it does not constitute empirical validation of the essay's claims about real systems.