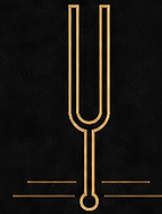


— THE —
**THREE
THRESHOLDS
TRILOGY**

— ◆ —
TWO TOOLS. ONE DESTINATION.
—

by Chris VenHaus



VHAUDIO.COM

READER'S EDITION · MMXXVI

VH AUDIO PRESENTS

The Three Thresholds Trilogy

Two Tools. One Destination.

by Chris VenHaus

PART I	The Three Thresholds	2
PART II	Measurable, Audible, Livable	12
PART III	Fidelity to What?	33

About this edition. This PDF was prepared by VH Audio for personal, offline reading and printing. The essays live online — with sources, updates, and discussion — at vhaudio.com/three-thresholds-trilogy.

© 2026 Chris VenHaus · VH Audio, Inc. All rights reserved. This document is provided for personal use only and may not be copied, distributed, reposted, or excerpted in any form without written permission. If you'd like to share the trilogy, please link to the original series so readers always get the current version.

Reader's Edition v1.1 · July 2026 · The latest version lives at vhaudio.com/three-thresholds-trilogy

The Three Thresholds

Two Tools. One Destination.

BY CHRIS VENHAUS | VH AUDIO



Someone asks, “*Can you hear the difference?*” and the room instantly divides. That’s because this single question is actually doing the work of three.

1. Is there a change in the signal that measurements indicate at the load? (If physics doesn’t allow it, you can’t hear it.)
2. Can you detect that change in short, level-matched, controlled trials such as ABX?
3. Does that change matter over time — the kind of change you’ll notice after an hour, a week, or a month, when you stop hunting for it and start listening?

This framework gives those three questions names: Threshold 1 (measurable), Threshold 2 (audible), and Threshold 3 (livable). These thresholds stack, but they do not always align, which is why one number, one sweep, or one test can end up being mistaken for the whole story.

Why These Arguments Go Straight to Nowhere

At some point, when you scan through *yet another* forum thread: “*Is it Better or Placebo!*”, you realize you’d rather discuss politics at the Thanksgiving dinner table than take the bait and hit the ‘Reply’ button. It’s usually the same loop, same cast, and the same predictable moves: someone posts a measurement, someone else counters with a listening experience, someone demands an ABX test, and then a goofy meme makes its appearance. These threads eventually devolve into a leg-lifting contest with everyone staking out their territory and daring the others to cross it. Wash, rinse, repeat.

But the interesting part is: *nobody is actually wrong...*

The measurement advocates are *right* that physics matters. The ABX camp is *right* that controlled testing can help isolate bias. And the listening crowd is *right* that some gear just feels more engaging than others.

So why does every conversation about subjectivity vs objectivity feel like we're talking past one another?

The Three Thresholds

I've been chewing on these questions for decades, hunting for the 'why.' My own experience — as well as direct feedback from thousands of fellow audiophiles — kept highlighting a gap that measurements and quick tests didn't appear to account for.

For years, I couldn't pin it down, so I kept gathering information on psychoacoustics, audio-adjacent sensory fields (especially vision), and cognitive science. The more I learned, the more it made sense to map what I found into three mental 'buckets'.

It starts when someone asks *"Is there a difference?"*

They could mean: *"Does it show up on an analyzer?"* or *"Can you pass a blind test?"* or *"Do you like it enough to want to live with it?"*

Here is how I define the Three Thresholds:

Threshold 1: The Measurable (Domain of the Analyzer)

Can instrumentation see the change? Distortion products, noise spectra, impedance behavior, time-domain ringing, load interactions — things that exist whether anyone is listening or not. If physics doesn't allow it, you can't hear it. *If a Threshold 1 claim is made, support it with the measurement(s).*

Threshold 2: The Audible (Domain of Vigilance)

Can listeners reliably tell A from B under controlled conditions? Level-matching, rapid switching, ABX — short trials built to answer a specific question. It's powerful, but it also requires vigilance: you're on the hunt for a 'tell'. *If a Threshold 2 claim is made, support it with the controlled detection method.*

Threshold 3: The Livable (Domain of the Chair)

Can you live with it longer-term? This is where fatigue, ease, attention drift, and engagement show up. You notice it in how long you stay in the listening chair, whether it becomes work, or if it is truly 'go time'. *If we're sharing a Threshold 3 experience — how a component settles into our system (and your life) after weeks in your own chair — support it with time.*

These thresholds build on each other, but they don't always agree. A change can be measurable but it might not be heard. Heard but not enjoyable. Hard to nail down in a *quick* test, but obvious over the longer term. Maybe you can't confidently pick right after a swap, but one component keeps drawing you back into your listening chair. The goal here isn't to 'dunk' on measurements or blind tests — it's to stop forcing one tool to answer *all three questions*.

Not hearing a difference in a careful short test is useful information, but it's not the same as answering what happens over time in your own chair. You don't have to turn in your 'Scientist' card to also trust your ears. There is common ground if we have a framework that includes both.

Once we name which question is on the table, we can focus on what's truly 'signal' vs static. It helps us sort what's being tested, so we can be on the same page.

- “*The specs are identical!*” — Threshold 1 claim (Measurable)
- “*I couldn't hear it in a blind test.*” — Threshold 2 claim (Audible)
- “*I eventually got bored of it.*” — Threshold 3 claim (Livable)

These examples sound nice and tidy, but it's not always so clear-cut. In practice, there are two 'gotchas' that might be lurking...

The Listener's Trap

It was the late 80's, and I'd been a Zildjian cymbal fanboy my whole life. Then I heard some sweet cymbal sounds in a recording and learned that the drummer used Paiste's RUDE series cymbals. I was totally smitten with the sound. That same summer, while browsing Daryl Stuermer's (*Genesis*) music store in West Allis WI, I spotted a set of used RUDEs: hi-hats, ride cymbal and two crashes. They just *needed* a home, and I'm charitable that way... so I snagged them.

Took them home, set them up, and... ugh. Dark. One-dimensional. Nothing like what I'd heard on that album. I stuck with them for a few days, figuring maybe I'd warm up to them — “*new cymbals, different feel, just give it time...*” I was just about ready to give up, cut my losses and trade them in, until an odd thing happened: I had an impromptu recording session in our attic rehearsal space when a guitar player buddy came over, and I flipped on the Tascam multitrack and Beyerdynamic mics just for kicks.

Days later, when I hit playback... I was listening to a completely different set of cymbals. Everything I'd wanted was right there on the tape. Rich, complex, alive, and that ride cymbal cut right through beautifully. The crashes sounded full — not the dry 'trash can lids' I heard behind the kit.

I eventually sold those cymbals with that drumset, and I still regret it. I think every drummer (and guitar player) has a *‘one that got away’* story. Those RUDEs were mine.

My best friend Brian Kelly — a guitar player who was doing session work for big names in 1960’s Cincinnati at the tender age of sixteen — told me a similar story. While chatting about ‘tone monster’ electric guitars, and how they often sound great unplugged, he flagged one notable exception: the Gibson L-5 archtop can sound *awful* unplugged, but plug it in... and it’s a different beast.

Same instruments, different verdicts when the perspective changes. This recurring pattern shows up elsewhere too...

Highway Gothic: When Better Is Worse

Take a look at the green signs on the interstate tonight. For decades, the US standard was a font called Highway Gothic. It’s a typographic detail monster: hard edges, high contrast, everything shouting LOOK AT ME. In quick identification tests, it wins.

But put that same font on a freeway at night, hit it with headlights, and the trap snaps shut. That ‘maximum data’ blooms into a glowing white blob. The light bleeds, closing up the holes in an ‘e’ or an ‘a.’ Typographers call it halation. In audio terms, it’s glare — the difference between ‘faux transparency’ and *“I can’t listen more than an hour or two.”*

This is where a font called Clearview came into the picture. It had softer edges, and slightly lower contrast. In a side-by-side ‘showroom’ test, it looked worse... less punchy, less immediately sharp. On the highway at night, legibility improved — especially by addressing halation/overglow — which is exactly the kind of ‘looks softer, works better’ tradeoff the showroom test can miss.

Audio has the same trap — some call it ‘Showroom Treble’. The demo that grabs you — the system with the sharpest attack, the most forward mids, the hyped presence — may win the quick test. You’re not wrong for hearing it. Your brain is wired to notice transients, edges, contrast. But that same wiring doesn’t predict what happens three hours later, when the sharpness becomes fatigue, and you find yourself losing interest and shutting your rig down for the evening.

The opposite trap exists too — the warm, syrupy presentation that flatters a vocal in a quick demo, but turns dense mixes into mush. Different bait, same hook.

The showroom score and the living-room score are not the same test.

The Clip Is the Sip

Audio demonstrations often use brief clips to showcase differences. The clip and the sip are structurally identical problems.

A thirty-second demo of a component swap is the ‘sip test’ of audio. It captures the moment of maximum attention, the first impression, the edge-detection response. What it can’t capture is what happens when you live with that component for two months. Does the system invite long listening sessions? Do you lose track of time? Or do you find yourself turning it off after an hour, slightly fatigued or bored?

The clip optimizes for Threshold 2. The living room optimizes for Threshold 3. They aren’t the same test.

When Play Becomes Work

A component or system tweak that imparts subtle edge or glare to the high frequencies might be undetectable in a five-second ABX switch. Your attention is consumed by the comparison task itself. But over a two-hour listening session, this edge accumulates.

Your auditory system spends extra effort compensating for the subtle wrongness. The brain is a prediction engine, working to anticipate what comes next. When artifacts don’t match natural sound, those predictions fail. Each failure has a price, and this extra burden is like running too many computer programs at once on an old machine. One program is easy, but too many running in the background bogs it down. Each cost accumulates... and this burden takes its toll.

The bench is great at steady-state; but your grey matter lives in dynamics, pattern-matching, and recognition. When its predictions fail repeatedly, the brain flags it as ‘off’ — similar to the odd feeling you get when you see a wax figure that looks remarkably lifelike, but is obviously fake. That *wax museum* effect shows that technical accuracy can also have an experiential hole that raises alarm bells.

It’s like a chair — you can sit in a showroom office chair for a minute or two and declare it comfortable. Get it home, and after a couple of days, it’s become an instrument of torture. The short test doesn’t reveal the long cost.

‘Hunting’ identifies differences. ‘Soaking’ reveals whether you can live with them.

The Groove Test

You can put a metronome on a drum track and prove the timing is perfect, but if it doesn’t make your head nod or toes tap... something’s missing. I’ll take the groove laid down in the pocket of a drummer like Bernard Purdie over a metronome every time. Audio works the same way. Perfection on the bench isn’t always the groove in the room.

If you sat down for a ‘quick’ listening session after dinner... then realized it was 3am... or if you ‘rediscovered’ why you love music in the first place... you’ve hit paydirt.

You can count a lot in audio — and you should — but the thing you’re actually buying is the part that makes you forget you’re evaluating.

Threshold 1 is what you can measure. Threshold 3 is what draws you back for more music tomorrow. Threshold 2 sits between them — and it’s where the tool gets confused for the *destination*. That’s the Measurement Gap: the space where the instrument or process is bound to themselves, and the way you actually listen to music is bound to *you*.

The Preference Trap

Ask any guitar player about their favorite neck profile and opinions will be all over the board. A vintage Telecaster U-shaped ‘baseball’ bat, modern thin, or something in between like a 60’s C. Tall jumbo frets or thin vintage style? Everyone has their favorite, and everyone is convinced they’re right. But the ‘right’ neck depends on how you’re built, what feels comfortable, and your type of playing style. A jazz player’s dream neck might cramp a shredder’s hand by the end of the first set. If you think the objective vs subjective debates are wild in our hobby, walk into any guitar builder forum and start a topic about tonewoods used on electric guitars. You’ll see what ‘audiophile-level’ disagreement looks like — sprinkled with sawdust.

These same dynamics play out in audio, and they run deeper than “I like it” vs. “I don’t.”

Different Menus, Different Ears

I like to think of each audio manufacturer as operating similarly to a restaurateur. They build products that reflect what they think sounds good — their ‘recipes.’ Some listeners will love that product. Others will think it’s just okay. Still others will have a visceral “*this isn’t for me*” reaction. None of them are wrong.

To be clear: *Broken or inferior gear is still broken and inferior*. This is where Threshold 1 is a valuable benchmarking and screening tool.

Once you’re in the world of competently engineered designs, you’re often choosing between different priorities, not ‘right vs wrong.’ Manufacturers are well aware of this, because sometimes choosing more ‘X’ means less ‘Y’ in a design. These distinctions matter because people hear differently, value different aspects of music, and listen in different environments. It’s as personal as our palates. Our systems end up custom-fit — tuned for how/why we actually listen.

This doesn’t mean “*I like this*” is the same as “*I’ve explained this*”. One involves taste, the other is the ‘recipe’. Both matter, but each is an answer to a different question.

Now for a practical problem using my restaurant metaphor: how do you evaluate a restaurant from one bite, and how does a manufacturer communicate what their menu is really about in a 30-second sample or a quick blind test?

A friend of mine used to compete on the barbecue circuit — a competitive pitmaster who ranked in the top 1% nationally by the KCBS. His competition ribs were legendary: mahogany-dark, glazed to perfection, intensely flavored. But if you went to his house for a backyard cookout, he wouldn't dare serve that.

He told me flat out: *"This isn't what I'd cook for my family or friends for a meal."*

In competition, a judge takes exactly one bite. That single bite has to cut through the palate fatigue of tasting dozens of entries. So the recipe gets tuned for maximum immediate impact: extra salt, extra sugar, a texture dialed for a specific snap. The barbecue world calls it 'one-bite barbecue.' It's engineered for a six-second window.

Put those same ribs on a picnic table and ask someone to eat a whole rack. Different story. What dazzles in bite one... wears out its welcome by the third rib.

Competition food is a performance. Backyard food is a meal. Neither is wrong — they're optimized for different contexts.

The system that leaps out of the speakers in a showroom can win the quick audition. But the system that wins the quick audition isn't necessarily the system that wins an entire evening... or the month.

When the Metric Becomes the Target

Economists have a name for what happens when a single measure becomes the target: Goodhart's Law. *"When a measure becomes a target, it ceases to be a good measure."*

Barbecue judges look for the smoke ring — that pink band under the bark that reads as 'authentic low-and-slow.' The chemistry is real. But once judges started rewarding it, it became possible to chase the badge instead of the process. The signal of quality became a cosmetic target.

Audio has its own history of 'smoke rings'.

For a long stretch, THD was the number you could print in bold. The lower, the better; the scoreboard was public. So designers naturally chased it — cleaner sine waves at 1 kHz into a friendly load. In many cases, that work produced genuinely better gear.

But the 70's and 80's also taught us something important: you can ace the test and still miss the experience. Ojala's work on transient intermodulation distortion was one of the moments that made people realize: you don't know... what you don't know...

Once a single metric becomes the scoreboard, the designs drift towards that number — and the goalpost shifts.

Today the ‘smoke ring’ shows up in new clothes: SINAD (signal-to-noise-and-distortion) leaderboards and spec-sheet races. Useful? Absolutely. It can keep you from buying junk. But it still doesn’t answer the question most of us care about:

“Will I want to fire-up my rig on a random Tuesday night?”

Same Seat, Different View

I’ve occupied the exact same chair that well-known audio reviewers sat in at shows, listening to the exact same track they’d just auditioned, and came away with an entirely different impression. Same system. Same room. Same recording.

Two listeners in the same seat, hearing the same track, reaching different conclusions. Is this really a surprise?

Our audio systems are akin to prescription glasses — each listener designing for their own listening perception. The goalposts aren’t static for everyone. What counts as the reference: the master tape? Or maybe it’s what the engineer heard in that room with their own monitors? Perhaps the musician’s perspective as they played the instrument... or an audience member in 7th row or 15th?

Right Tool for the Right Job

Early on in my audiophile journey, I often wondered if I was listening ‘the right way’. *“What should I be listening for, specifically?” — “How is the bass on this track?” — “Let me queue up a female vocal and see how that is.” — “Wait, maybe I should try another recording that doesn’t have that ‘sss’ emphasized so much.”*

I think most of us have been through this dance.

Are we even capable of hunting for ‘is it bright’ and ‘how’s the bass punch’, while simultaneously analyzing the soundstage? Can we do all of that and also let the music just take us on a ride?

Tasking vs Listening

ABX is a vigilance task. You’re not ‘just listening’ — you’re comparing. You’re holding two sonic snapshots in your head, flipping back and forth, trying to catch a tell. That’s a real skill, and it’s exactly what Threshold 2 is built for.

But it's not the same headspace as an enjoyable, long listening session. When a system is really tuned, you stop grading it. The gear disappears. You relax. You're not hunting — you're immersed inside the music.

And here's the part that gets missed: those two modes don't play nicely at the same time. Turn the comparison dial up to '11' and the dial for 'immersion' goes the other way.

Are we testing detection, or testing livability? If we run the wrong test, we can get a clean answer to the wrong question. Measurements and quick A/B tests help answer the questions they were built to ask, but there's more to the story when we look closer and see the gaps. We can find examples of these types of gaps in daily life...

A Mattress of Fact

We've all been through the 'mattress' decision. We're in the showroom — this one is 'too firm', that one is 'too soft'. It gives us a relative 'Goldilocks' read on what might work for us. You make the selection, get it home and... *"Oops — felt right in the store, but my back is sore after a week of sleeping on it."* The better deal you got, the salesman's pitch, or your spouse's insistence *"this one will be better..."* fades behind the reality of actually living with it.

Quick A/B tests are similar. They aren't flawed — they're *scoped*. They answer: *"Can I reliably hear a specific difference right now, under these conditions?"* They weed out the obviously too firm or too plush, narrowing you down into a zone that might work.

Rapid swaps are biased toward 'switch-winner' traits — contrast, edges, spotlight detail — but what grabs you in the comparison doesn't always predict what you'll want to live with. If you've ever had a component that wowed you in the demo but wore you down over time, you didn't fail to hear it — *you entered the wrong contest...*

Just like the mattress demos in a showroom, time acts as a truth serum, and you're left with a more honest reality of what's livable: *"Does this change make me want to listen longer... and will I be drawn back for more tomorrow?"* That's not a 30-second evaluation, and it doesn't live in the same mental state as 'spot the difference now.' It's a bit like your spouse measuring the quality of your sleep by waking you up every ten minutes. It's the same structural failure: the act of checking prevents the thing being checked.

The Tuesday Night Test

When I began my audiophile journey, I'd built a system around components that seemingly checked every box — Class A rated, reviewer-approved, specs that looked great on paper. I was proud of it. And then I noticed I was listening less in my 'big rig.' I'd sit down, play a couple tracks, admire the

‘resolution,’ and then calmly proceed to do something else... I blamed the recordings. I blamed my mood. I thought maybe I was just over the hobby. But the truth was simpler: I’d optimized for *someone else’s destination*. Those reviewers weren’t wrong — they heard what they heard. They were listening in a different context, for a different purpose, with a different question in mind. I imported their answers without asking whether we shared the same question. That wasn’t a failure of my ears. It was a failure of my map.

Here’s how I think about it now: Measurements are like a topographic map — contour lines, elevation data, drainage patterns. I routinely use topo maps for landscape photography. They’re incredibly useful for planning, for eliminating dead ends, for knowing what you’re getting into. But the map can’t tell you whether the light will be worth the long hike at 4 a.m., or whether the area will blow you away when you are actually there.

A/B tests are like travel photos — more vivid than the map, a little closer to the experience, but still just a representation. We’ve all seen stunning photos of places that felt ordinary in person. We’ve also visited places no one bothered to photograph that we still yearn to go back to years later.

“*I can measure it*,” (T1) and “*I can detect it*,” (T2) are tools for the journey — useful, often essential, but always in service of moving you forward to the real destination: “*I can live with it*” (T3) — the random Tuesday night that pulls you into the chair... and keeps you there.

Those are three very different questions — two of them help you navigate. The last one is where you actually want to ‘live’.

> [Sources used for this article](#) <

Published May 12, 2026. © 2026 Chris VenHaus. All rights reserved.

Measurable, Audible, Livable

BY CHRIS VENHAUS | VH AUDIO



SECTION 1:

The Question That Wouldn't Let Go

One of the biggest ‘lightbulb’ moments I’ve had in audio didn’t happen at the test bench. It was before a business meeting in the early 90’s, while killing time at an audio dealer in Cross Plains, Wisconsin. As I walked by a dimly lit demo room, my eyes caught the orange glow of vacuum tubes. The lure was set. Before I knew it, I was standing in front of Lumley M120 monoblocs lit up in all their glory. The dealer queued up the usual audiophile demo tracks. The sound? Mesmerizing. Rich midrange, incredible soundstaging, sweet highs and... that tube glow! This was the kind of demo that pries open your wallet... after a two hour round trip I made just a week later.

When I got them home and fired them up with my music (densely-layered instrumental rock), that previous magic took a back seat. The lush midrange I was smitten with in the showroom now turned complex passages into a congealed, tangled mess. *The system that won the showroom failed my living room.*

The remedy? An act some would consider ‘heresy’: I swapped the highly coveted GE 6550s for Svetlana ‘Winged C’ tubes. Same amps. Same room. Same speakers. A simple tube swap... game changer! The Svetlanas resolved the nuanced layering and brought back the taut bass I expect, while also bringing harmonic richness to the party. The ‘lesser’ tube for many audiophiles was the better tube for me, in my system and with the music I love.

That gap — ‘wins the demo’ vs “*uh-oh, what happened?*” — kept reappearing. Over the years I ran into many more components and ideas that impressed on paper or in a quick listening session, but ultimately fell apart once I spent more time with them.

Sometimes the pattern flipped. The component I was certain would be transformative? Still collecting dust in the basement. It’s not just *me*. Since the mid 90’s, I’ve traded emails with thousands of fellow audiophiles telling the same story — gear that looked good on paper but the ears vetoed... despite them wanting to love it.

I eventually recognized a parallel to this pattern in another passion of mine: landscape photography.

In landscape photography, a print isn't a mirror of what I saw — what you see is shaped by the interplay of the print, the room, lighting choice, and your own eyes. The same image that sings on the gallery wall can feel 'off' in your own living room just a week later: different wall, different lighting, different mood, and different eyes on a different day. The photographer makes calculated choices, but *the final experience unfolds in conditions out of their control*.

Audio faces similar hurdles. The only *first-hand reference* of the performance lived with the people in that room — each musician with their own window, and the engineer who framed the scene.

Everything downstream is a rendering engine. We aren't re-living the event; we're interpreting its raw artifact.

If a visual art like photography requires interpretation, can there be a parallel in audio? My history studying optical illusions and photographic composition was a good 'gateway drug' here, making me wonder if audio has its own version of the classic 'face or vase' trick.

It does.

Around 2010, I started acquiring books on auditory perception (Bregman, Deutsch) because I wanted to understand why some 'demo' winners would lose over time. One example really blew me away: a pair of computer-generated tones, played one after the other. Some listeners hear this pattern climbing in pitch, while others hear the *exact same pattern* falling in pitch. This is called the Tritone Paradox... but the label matters less than the takeaway: *same input, different perception*. In other words: 'all in your head' means more than at first blush, and that's the whole point.

The signal doesn't just stop at your ears — your brain completes the picture. While measurements and parts quality still matter, the signal chain doesn't stop at the speaker's terminals, or even the room. *It runs all the way through you*. This processing isn't free — it's metabolic work that burns energy and attention. Whether what's 'in your head' is the end of it or just the beginning depends on how long you're willing to sit with it.

Other fields like food science, vision research, and sensory psychology have been working through this for decades. We don't need to rewrite physics. We just need to add the missing pages to our vocabulary.

SECTION 2:

The Fine Print on Quick Tests

I've run quick A/B swaps for over 35 years, and bench tests daily for QA and prototyping. The first 'missing page' I turned to was the test itself. What kept nagging me? The spec or A/B winner crowned

in twenty minutes... wasn't always making it to the finish line longer term. So I pulled a Gold Standard for A/B testing. Not secondhand interpretations. The source document: ITU-R BS.1116. Its scope statement reframed everything: the tool was engineered for detection, not as an arbiter of satisfaction. The science isn't wrong — the question is.

The global engineering standard for listening tests (ITU-R BS.1116) specifically states what it's intended for (and what it's not). Like any precision instrument, it was designed for a specific purpose:

“This Recommendation is intended for use in the assessment of systems which introduce impairments so small as to be undetectable without rigorous control of the experimental conditions and appropriate statistical analysis. If used for systems that introduce relatively large and easily detectable impairments, it leads to excessive expenditure of time and effort and may also lead to less reliable results than a simpler test.” — ITU-R BS.1116-3

That last sentence resonated for me: using it outside its designed scope can make results *less reliable*. It's not a universal referee... it's a precision tool for a specific class of question.

Put another way: the protocol is optimized for the ‘*detecting brain*’. But the detecting brain isn't the same mode as the ‘*enjoying brain*’.

To help understand that boundary, I view it through three distinct lenses, rather than a linear sequence:

The Three Thresholds at a Glance

T1 (Measurable): What the instruments capture on the bench.

T2 (Audible): What you can detect in a rapid switch.

T3 (Livable): What time reveals in your listening chair.

** (See [The Three Thresholds](#) for the full framework.)*

Rapid-switch testing rests upon biological constraints. Your brain's short-term audio buffer — what scientists call echoic memory — holds roughly 2 to 4 seconds of sound before it fades. Holding it longer takes active effort: you end up rehearsing the sound internally, which burns real metabolic energy. In A/B testing you're doing double duty — keeping the last sound alive while processing the new one.

To use that short-term buffer, switching must be rapid; you're comparing snapshots, not movies. Each session caps out around 15 to 20 minutes before fatigue degrades your ability to reliably catch

anything. The entire protocol is designed to keep you in *vigilance mode*: hunting for ‘tells’, flagging differences, staying alert.

That’s where the ‘*Eureka!*’ moment was for me: The state of mind required to detect a difference competes with the state of mind required for musical immersion. You can move between them, but leaning hard into one inevitably pulls you away from the other.

Under the hood it’s a see-saw: when I climb onto ‘*detect*,’ the ‘*just listen*’ side lifts off the ground.

When we listen for ‘tells’ in 4-second loops, we’re effectively proofreading the audio. But nobody reads a novel by proofreading it, and we certainly don’t *enjoy* reading a book that way. If you stop every sentence to check for typos, you’ll find the typos... but you’ll likely miss the plot (and be bored to tears along the way).

The test works because your attention is hunting for a difference. Spotting one just means ‘different’ — not necessarily ‘better.’ But... you can’t hunt and surrender at the same time.

Expert listeners can predict what might bother people eventually... they can catch it faster than casual listeners. However, they’re doing it through *trained attention in short bursts*, not by living with the system for hours, days, and weeks. *It’s a snapshot*... great for catching what sticks out, but less useful for what settles in.

So the question isn’t whether BS.1116 (a *gold-standard blind test protocol*) works... it’s *which question are we trying to answer*. Detection under ideal conditions? That’s T2. Living with it over time? That’s T3. One is about *vigilance*. One is about *immersion*. We’re not just taking snapshots of the signal... we’re *marinating* in the sound.

I use measurements every day for quality control and testing new design ideas. But over time, my listening experience weighs what the bench data can and cannot tell me. We can’t listen to measurements... We listen to the music. *Measurements are the tool... they aren’t the destination*.

The legendary Bob Ludwig — arguably the most respected mastering engineer ever — put it bluntly: “*A/B testing, while the only scientific method we have, does not reveal too much with short-term back-and-forth comparisons due to the anxiety the brain is under doing such a test.*” (Ludwig, Tape Op #105, 2015).

What happens when we stop compressing time and allow it to unfold?

SECTION 3:

What Time Reveals

It won the A/B test. The ‘winner’ felt obvious. I can usually trust my own ears...

A month later? “*Nope, not this time...*”

The sound that impressed during the demo... laid bare with more time. Nothing was obviously ‘broken’. It still had the attributes that pulled me in... but somewhere along the line, my system was no longer luring me into my listening chair.

That gap... between the sound that initially impressed and what I wanted to live with, kept reappearing.

The disparity wasn’t coming from the gear: It was the way I was testing it. A rapid A/B swap acts like a zoom lens in the woods: zoom in, and the motion of a squirrel shaking its tail captures your attention... while you lose the whole forest around you. My attention was getting grabbed by the zoomed-in lens I was told to use. The question wasn’t whether my hearing was lying: it was whether the test format itself was hijacking my perception.

The mechanism behind that hijack clicked when I read Bregman’s *Auditory Scene Analysis*: Your brain has a ‘let it fall into place’ mode where it quietly sorts the scene, and a ‘compare and label’ mode that kicks in when you force the analysis. Rapid A/B switches drag you into that ‘auditor’ mode. But as Bregman showed, your brain needs several seconds to lock onto who’s playing what... so every A/B switch resets and defaults to ‘zoom-in’ again.

That is the *Vigilance Tax*.

When the signal drifts from what your brain is predicting — the tonal balance it mapped, the soundstage it locked, the attack it expects on every transient — it clamps onto the mismatch... and *the music becomes the task*.

Ground Already Covered by Audiology

If the snapshot only shows us what changed... how long does it take to hear what stays?

Audiology demonstrated this principle long ago. In 1992, Stuart Gatehouse tracked new hearing-aid users and found that perceived benefit kept increasing for six to twelve weeks after fitting. He concluded it “*calls into question short-term methods of evaluation.*”

In a quick session, we’re judging the handshake, not the marriage.

A version of this shift often plays out in our own listening chair... compressed into a single session. You begin on alert, switching back and forth, spotlighting a specific layer or instrument: blunted attack here; soundstage recessed there; “*cymbals not quite right with this one*”... then you quietly stop ‘proofreading’ the sound, and forget you’re there to compare. You came to evaluate, but instead got lost in the music. When I dug into the research, I found a reason for this...

See-Saw... to Swing

You can hear it... when you stop hunting for it: the ‘pieces of sound’ turn into *music*. If you focus on a small slice of the sound, your brain locks onto it — at the expense of everything else around it. This focus may have been a choice... *but the collapse isn’t*.

I found the reason in Bregman’s work (1990), where two hearing modes explain the split: one is primitive, fast and *involuntary*, while the other slowly collects and assembles the whole scene over time. LeDoux (2000) found your brain’s shortcut underneath — a primal pathway that fires well before your ‘evaluating’ brain shows up. It’s like a smoke detector versus your nose. The detector is crude and fast, triggering on burnt toast the same way as it does on a house fire. Your nose takes longer... but it can tell if you simply ruined your breakfast.

Every A/B switch trips the alarm first — then your brain has to rebuild the scene before the sound turns back into music.

You know when a soloist locks-in and everything else just... *disappears*? Limb and Braun (2008) used fMRI to scan jazz pianists while improvising: their inner critic (DLPFC) shut down. Analysis ‘whacked’ receptivity.

That’s the performer, but what about the *listener*?

Ever notice how the moment you start grading the sound, you’re no longer in the music? That shift is measurable: the brain regions that ‘listen’ and the regions that ‘reward’ effectively disconnect when the task becomes evaluative (Liu et al., 2017). So the problem isn’t *focus*... it’s *evaluation*. Even the mental effort of tagging a sound passage as ‘happy’, ‘sad’, or ‘fearful’ — tilts the brain’s see-saw from feeling to ‘tasking’ (Bogert et al., 2016).

It’s not just music. When viewing artwork, simply changing our frame of mind between *appreciation* (how it makes you feel) and *analysis* (picking out objects) wakes up different parts of the brain (Cupchik et al., 2009).

The analysis finally stops. You wish the music would just pull you back in... except it doesn’t. This common frustration led me to a listening study from Jäncke’s lab that blew my mind. The setup: 51 listeners listened to music under two different conditions. When just listening (for pleasure), all five

EEG bands — theta through upper beta — lit up. When they *added the task of rating the music* as it played, those bands dropped. Here's the mind-blowing part: when the rating task came first, that 'just-listening' state wasn't recaptured later in the session (Markovic, Kühnis & Jäncke, 2017).

That's five distinct studies showing the same pattern: evaluation and immersion tilt the brain in opposite directions (Limb & Braun, 2008; Liu et al., 2017; Bogert et al., 2016; Cupchik et al., 2009; Markovic et al., 2017).

This all aligns with my experience: Begin the audit... and you're listening *at the music*, instead of *into it*.

Sometimes the gear just disappears, time is suspended as you sync up with the music, trading analysis for that pure experience. *Feeling*, instead of thinking. Other times the spell snaps — your 'analytical' mind crashes the party like an unwanted sales pitch. The vibe is gone; you're put in the corner, running the same checklist you thought you were done with.

Your brain's see-saw has a bias: bear down on the 'analytical' and immersion lifts off the ground — not because you chose to stop feeling, but because the two modes compete for the same bandwidth. When the analytical side finally eases up and the see-saw tips back, detection gives way to what musicians simply call *swing*.

Field Notes: The Drummer's Perspective

After 40+ years behind a drum kit, I've learned much of the 'groove' lives between the notes, not on the 'click'. Guys like Purdie, Bonham, Paice, Gadd... they don't operate like an atomic clock. They slide a little ahead or behind, while staying 'in the pocket'. Hendrix did the same thing on guitar. Transpose his playing to a perfect grid and you'll suck the life out...

When you're truly locked in, you stop thinking about the *task of playing*. You're not counting. You're not analyzing. You're just 'in it', and the music is coming through you rather than coming from you. The best sessions feel like the band becomes one living organism, and the music seems to feed itself.

In that state, the *quiet stuff* starts doing more work than you think. Drummers have a name for the taps that don't beg to be noticed: **Ghost notes**. These are subtle flutters on the drum tucked inside the backbeats. Isolate the track and they're obvious; anyone can hear them. But in the full mix, you don't experience them as '*a thing to identify*.' You experience them as the mortar between the bricks. Take them out and the tempo's still technically correct... but the *groove goes flat*.

That's the curious part for me: they're often below our attention threshold, but NOT our hearing threshold. The 'subtle stuff' does its work in support of the main beat to form the full groove, without

the need to spotlight it. As a drummer, when locked in... it's as if my hands know what the groove needs before my brain issues the command.

Simon Hoeffding gives this a name: 'performative passivity' — it's essentially *when you stop playing the music and the music starts playing you...*

The instant I snap back to check my timing, this state collapses. I can't analyze the beat *and* ride the groove at the same time. Psychologists call this *flow* — some musicians call it *absorption*. For the listener, it's a state of *total reception* — letting the music take over. But you can't reach that state while auditing your gear.

The Spec-Sheet Trap

Back in the listening chair, I can be a few minutes into a track and then — snap — I'm grading my system again. It feels like an errant house light came on mid-concert. I'm glad my ears caught it, but in this same instant... the music loses balance. Pinning down that "thing" is like audiophile catnip... but it's also like a mousetrap. Once that inner scorecard opens, I reach for available 'handles' I can readily track and register. This experience was at odds with sessions where I came in with no agenda. That made me wonder: perhaps the issue isn't a change in my hearing; maybe it was my approach that cast a different frame?

In 1996, a paper by Christopher Hsee nailed this feeling — with CD changers, of all things. He described two hypothetical models on paper: one held five CDs but listed 0.003% THD; the other held twenty CDs but showed 0.01%. No listening was involved — just specs and a question: would you pay? When shown side by side, people leaned toward the lower-THD unit, as if that tiny fraction-of-a-percent gap was an obvious handle. Naturally, people had no way of knowing if either distortion number would map directly to audibility or not — yet the *comparison made it feel* like the clean way to pick the winner.

In isolation, the preference flipped. Twenty discs vs five needs no explanation: capacity became the *intuitive winner*. That THD number only appeared important when anchored with the same spec sitting right next to it — the comparison frame itself changed what mattered.

An A/B swap is like a measuring stick: great when you're chasing a narrow question like, "*is this change rolling off the highs?*" But the moment you pick up that measuring stick, the task quietly changes — from "*what do I want to live with?*" to "*what's different?*" And when that difference is subtle, your brain doesn't settle for ambiguity; it reaches for whatever's pre-loaded.

You put the gear *in the way...* by checking if it's in the way.

The feeling is familiar — some sparkle or transient ‘bite’ jumps out in a quick swap — the kind of ‘wow’ you can easily hear in ten seconds... That’s contrast doing its job, and it feels like expertise: “*I caught that — ‘Golden Ears’ are still working...*”

But when the switching stops and I’m just living with my setup, a different ledger opens. It shows up in my behavior before I can name it: sessions get shorter, I start surfing the playlist for ‘showoff’ tracks instead of digging deeper into the rest of my library — or God forbid... *new music*. I might catch myself leaning forward to listen *at* the system, rather than relaxing back and letting the music wash over me. The body registers the verdict before the mind can name it. Hard to evaluate doesn’t mean inaudible — it means hard to score in a ten-second switch. The quick-swap winner and long-haul winner might not be the same.

If this disparity of evaluation over time is baked into our judgment, I just had to see if there were parallels outside our hobby. Where else do quick tests pick one winner, while time quietly votes for another?

What the Tongue Already Knew

Food science hit this dilemma decades before audiophiles did. Tongue and ear are different intake devices, but the processor upstairs — our brain — is the same. Their version of a quick A/B is the *Central Location Test* (CLT): small samples, tight controls, quick verdict.

Carol Dollard, the PhD biochemist who directed flavor development across Pepsi’s entire product line, learned this the hard way: “*I’ve seen many times when the CLT will give you one result and the home-use test will give you the exact opposite... sometimes a sip tastes good and a whole bottle doesn’t.*” (Quoted in Gladwell, 2005)

If the tongue can be fooled this reliably in quick tests — same brain, same shortcut — can we map this to our ears in rapid listening tests? Audio researchers went looking... and they found numbers that are hard to ignore.

When the Tool Bends the Ruler

This biological susceptibility to ‘contrast’ isn’t just limited to the palate. When audio researchers put their own evaluation formats under the microscope, this same pattern surfaced — short tests systematically weighted attributes that longer exposure rebalances. The sip-vs-bottle mismatch isn’t just a food problem. It’s a measurement problem.

Here’s the first surprise I learned: *the test itself can bias the verdict.*

In 2008, audio scientist Slawomir Zielinski published a review of biases in modern listening tests. What stood out was his take on the ‘recency effect’ — our tendency to weight the last thing we heard

over what came before it. With longer, time-varying material, *the quality of the ending dominated the listener's overall rating*. In telephone speech studies, this recency bias also shifted scores by up to 23% of the full scale range; in video evaluation, it reached as high as 50%.

Not 5%. Not a rounding error. Enough to question whether we're really judging the gear... or what we heard last.

The tool bends the ruler.

Recency isn't the only thumb on the scale. Choisel and Wickelmaier (2007) found that when two equally enveloping sounds were compared, listeners chose the second one 36% more often than the first — a bias baked into the presentation order itself. If you remember Bregman's scene-analysis reset from earlier: every rapid switch forces the brain to rebuild its auditory landscape, and the last scene assembled sits sharpest in the frame. Not because it's better, but because it's freshest.

Gold standards like BS.1116 are meticulously built to control for this: randomization, counterbalancing, many listeners, careful level-matching, etc... When you spread first/second positions evenly across a group, the order tilt smooths out in the aggregate, and you get a clean read on what changes are detectable. That's good engineering for the question the protocol was designed to answer. But here's the wrinkle: the math works after the fact — *across the group* — not inside a single listener's head in real time.

In the real world... our audiophile world: *You aren't an 'average.'*

We're each a listener in our room, running one trial at a time with no counterbalancing. The lab can neutralize that recency tilt; your Tuesday night can't. An 'average' helps a shoe factory stock shelves — it doesn't stop your feet from blistering in the wrong pair. If your experience diverges from the lab mean, that isn't irrational or a 'bug'; it's expected. This is precisely where the format injects something subtle: it doesn't just bend the ruler — it actually *makes the bending feel productive*.

Proofreading the Music

Why does A/B testing feel so... productive? I found a clue in a 1991 study by Timothy Wilson and Jonathan Schooler — not about audio, but about what happens when people are asked to *analyze* a preference instead of simply having one. Their test vehicle happened to be strawberry jams. One group simply rated what tasted good. The other had to explain *why* they liked what they liked. The 'why' group didn't just pick different jams: their choices drifted away from the sensory experts' taste rankings. Thinking about the preference warped the preference. My takeaway: every A/B comparison is — at its core — an invitation to explain why one sounds better.

In the listening chair, this invitation lands on the biology that's pre-loaded. Trained ears don't just hear more — they snap to anomalies involuntarily, before conscious attention shows up. That speed is the skill... and it has a price: once the 'snap' fires, it's harder to let go of what it found.

This trigger isn't just limited to flaws. Any hyper-contrast — sparkle that leaps out of the mix, or slam that punches above the track — pulls your brain's see-saw balance toward the analytical side. Immersion is still at the other end, but the tilt is already there before the first note plays.

Here's the allure: catching a difference feels good. It lands as expertise, so the brain keeps scanning — it keeps getting 'paid'. No matter what sense is being used, attention hunts for what 'pops' first, and then charges you for the effort. Each sense has its own scale... but the biology rhymes.

But later, does that difference feel like a comfortable pair of worn-in slippers, or am I just compensating down, stepping around a flaw I caught on Day One?

I've fallen for that one. An amazing DAC demo — open, fast, hyper-detailed. Six weeks in, I realized I was listening to a lab experiment, not music. I reached for reasons to keep it.

The 'attributes' that won the comparison had quietly become demanding. The effort to compensate didn't surface as a flaw I could easily point to. It silently crept in with low-grade fatigue: shorter sessions, restlessness, the slow realization that the system had stopped pulling me into my listening chair.

That is the Vigilance Tax settling its bill — not paid in dollars, but in attention.

Measurements and the quick A/B tests help answer the questions they were built to ask, but there's a gap between those answers and the whole story. They can score what's measurable on the bench and what's audible in a switch; they don't guarantee what's livable over weeks. The research of Olive and Toole at Harman has shown us that speaker measurements can reliably align with user preferences during blind testing in controlled sessions, with hundreds of listeners. That's solid science — for the question it was built to ask.

Food science has already shown us what happens next — *sometimes a sip tastes good and a whole bottle doesn't*. Does a tightly framed *snapshot in time* predict what it will feel like after months of use? That's a different question... and it's missing the long-haul evidence. If we treat the short-term answer as proof of long-term livability, we're not being more scientific. We're just skipping the hard experiment.

A/B testing doesn't just reveal differences — *it overweights whatever announces itself fastest*. Even if you could train your mind to stop chasing that spotlight... your body is already keeping its own score.

SECTION 4:

Different Inputs, Same Receiver

I've been observing the similarities between how we see and hear for years: composing a landscape photo vs. viewing one, performing music vs. listening to it. It made me curious — is this unique to audio, or does that same wiring reach into our *other senses*? It was time to shake some trees and see what would fall out.

What I found is — there's *much more* going on 'upstairs' than I ever anticipated.

The taste that wins the sip but loses the cup. There's actually a name for this: *Slowly Rising Aversion*. A food psychologist (E.P. Koster) gave it this name, after finding positive attributes tend to fade, while a minor defect — initially masked by novelty — gradually gets pulled forward. This goes beyond boredom. Across the studies his lab analyzed, fewer than half of participants kept their 'winner' in follow-up sessions.

The 'wow' normalizes, the flaw accumulates — entirely flipping the result.

The Spoon vs The Cup

In 1999, Zandstra and colleagues ran an elegant experiment using sweetened yogurt. They had people try it in different ways, ranging from a small sample to something they could continue eating. The result was blunt: optimal sweetness for a small sample was around 10% sugar. Once the question shifted to something they actually wanted to *keep eating*, it was only 5.9% sugar.

That is not a rounding error. The small sample winner had roughly two-thirds more sugar than what people ultimately enjoyed for the 'eat until satisfied' portion. The 10% yogurt won the sample... It lost *'the meal'*.

Short sessions reward intensity because *intensity yells first*.

While yogurt showed the sweetness contrast in a clean, compressed form, craft beer told the same story over multiple sips — and it shows us the mechanism in slow motion. In bold beers, the pleasant notes dominated the first sip, then faded as the taster's palate adapted; meanwhile the bitter edge came forward. With the Russian Imperial Stout, it more than doubled in dominance by mid-session (Simioni, 2018). The beer didn't change; the brain's processing of it did.

The pattern keeps reappearing when the sample leaves the test site and lands on your own table. Shi and colleagues compared central-lab tastings against home-use trials, then stacked their results against dozens of earlier product tests. On average, overall liking scores ran about half a point higher at home on a 9-point scale. Same products, different test, different verdict.

Translate the whole pattern back to the listening chair, and the trap cuts both ways:

On one side you might have exaggerated highs posing as ‘hyper-detail.’ In a quick switch, it wins because contrast has been exaggerated: more air, more bite, more ‘resolution.’

On the other side sits syrupy warmth: smoothed highs, thick lower midband, the kind of ‘velvety goodness’ that can feel seductive during a quick demo but turns into cotton candy once you live with it.

That ‘sparkle’ that won the quick switch? Give it a few weeks. That ‘feature’ becomes a glare ‘bug’ — that little tic you just can’t un-hear... the one that ultimately nudges the gear to the classifieds.

My main takeaway: every sensory domain I’ve checked is rhyming in the same zip code.

The brain adds *even more* to the stack: it quietly leans toward whatever tonal balance came before (dull or bright) and treats that as the new ‘neutral.’ In a 2021 study, Siedenburg and colleagues played the same recording after a dull or bright listening session. After the dull session, 66% called it ‘bright.’ After the bright session — *same exact file* — only 31%. The file didn’t change. The brain had dragged its internal anchor.

Ever review a photo and notice the color cast was too warm or too blue because auto white balance guessed wrong? Hearing does that too. Spend time with a bright, treble-hyped speaker and your sense of ‘neutral’ drags brighter. But drop a truly neutral component into that chain? Dull, lifeless, or even ‘broken.’ Not because it is — but because of the *auto white balance* for your ears. Your reference shifted.

When listening turns into sustained mental work, EEG research (Hunter, 2022) shows alpha and theta activity ramping up — a sign that your brain has started burning extra fuel. This is the biological footprint for the Vigilance Tax. You may not feel the fatigue in a 2-minute A/B test, but those rising brain measurements are the signal that your brain is working overtime to hold the picture together.

But the deeper I looked, the less ‘wrong winner’ covered what I was seeing.

The first impression wasn’t a lie — it just wasn’t the whole measurement.

A Toll Beneath

Ergonomics is a close cousin to what we experience in the listening chair. A plush office chair can feel incredible for a couple of minutes: soft padding, deep recline, that immediate “*ahhh... nice...*”

But eight hours into the workday, you’re shifting every ten minutes, your shoulders won’t unknot, and your back is a wreck.

I figured the explanation was obvious: the softness that felt like luxury is the same ‘feature’ that then trashes your back — same slider. But that’s not what Helander and Zhang found, or what De Looze et al. documented across multiple seating designs in 2003.

What they found: comfort and discomfort aren’t even in the same window — nothing about the chair’s comfort changed, but a different signal (on a different channel) had been building the whole time. The plush feel didn’t go away. The armrests had been quietly lifting the shoulders into a multi-hour shrug.

We experience this when listening to our systems too. We might be focusing on one aspect of the sound, while another sneaks up on us when we’re not looking for it.

This pattern kept showing up wherever I looked...

Smell has its own version. Ferdenzi and colleagues (2014) found the clearest case with chocolate: people who liked the aroma at first rated it less pleasant after repeated short passes. Same input, different readout.

Old fluorescent light ballasts more than doubled headaches and eyestrain — yet no one knew to blame the lights (Wilkins et al., 1989). An ERG later revealed why: the retina fires to invisible flicker (Berman et al., 1991). Body tallies the stress; mind lacks a target.

Keeping score:

Domain	Short-Term Winner	What Time Reveals
Taste	First spoonful (big punch)	Full serving (what you still want)
Touch	First minute in the chair (plush)	End of the workday (pressure, strain)
Smell	First pleasant aroma	Repeated passes can dull the pleasure
Vision	No flicker visible	Roughly double the headaches and eyestrain

For me, this pattern has been hard to ignore: quick A/B testing tells us what we can detect (assuming we can describe it). It does *not* guarantee what we can live with. After the first impression fades and you’re winding down on a random Tuesday night, it’s a different ballgame.

Science might call this a ‘longitudinal study’. We just call it ‘the hobby’.

Many of us have already lived this — the component that initially impressed but then made an unceremonious exit, or the one component ‘outdated’ by newer tech that keeps finding its way back into our rigs. This verdict first showed up in *what* you did, well before you knew *why* you did it.

SECTION 5:

The Silent Ledger

The audiophile community has been running this experiment for 70+ years. We've been living it... decade after decade, format after format.

Some components appear to be mayflies: they flare up in the spotlight, look unbeatable in a quick shootout... and six months later they're back in a box. Others get dragged through three houses, two marriages, and a dozen 'next big things' — and somehow they're still in the rack. The heirlooms get modded, restored, handed down — vintage turntables, original Quads, tube amps that quietly refuse to be buried. One group stays in the churn, while the others find homes. That gap between first impression and staying power has been showing up in racks for decades. Nobody coordinated that, because nobody had to. Stack enough anecdotes over time, and what you have is... data.

But the classifieds tell a second story. There are many reasons gear ends up there: *"Just not for me"* does not mean *"this is bad."* The amp that exhausted one listener's patience may become another's revelation — same hardware, different ears, different room, different life. The gear didn't change; the relationship did. The churn is not noise — *it's signal*.

Your body has been casting these votes all along. The readout is in your habits — patterns so quiet you might not have noticed them keeping score.

Once I stopped proofreading the audio and started paying attention to my own behavior, the tells became hard to miss...

The Volume Knob

With some setups the volume level changed, notch by notch. I wasn't thinking *"this is not good"*; my brain was just reducing the load. Your hand might know something your conscious mind hasn't flagged. If you're not much of a volume-tweaker, the same fatigue often shows up a different way: you get anxious, chase different tracks, or just wander away convinced it's an 'off' night. This steady pattern over weeks is a strong fatigue signal.

On keepers I caught myself nudging the volume up... because I wanted more. When the foot starts tapping on its own, that's the on-ramp — the balance starting to tip. For me, the tell is when I forget the foot is moving. The system invites you to 'play' — it doesn't put you to work.

The Playlist (Surfing vs Swimming)

Our playlists keep a similar log. System 'show-off' tracks... no longer a safe zone I was confined to. With the long-haul keepers I stopped surfing and started swimming at the 'deep end' of my music

library — I'd cue up one song and then realize it was already 2 AM... lost in tracks I hadn't touched in years, music I'd forgotten could hit that hard.

The Walkaway and the Tuesday Night Test

Then there's the walk-away. I notice it most on a Tuesday night: not a Saturday afternoon with fresh ears and strong coffee, but a Tuesday night after a long workday, just wanting to unwind. In that state, I don't want the music to make me work... Your system either pulls you into "*play time!*" or... you just go do something else. If powering down feels like taking off a pair of uncomfortable shoes, Threshold 3 was just lost.

The processing of what you heard doesn't just stop when you leave the listening chair. It's like waking up with the 'answer' to a question or problem you were chewing on the night before... or the 'improvement' a musician experiences between the end of the last practice and the beginning of a new practice session. Work is being done in the background, without us even knowing it.

The Reversal: Time as Truth Serum

If feasible (component-level change), after I've lived with a change long enough for the new sound to feel normal, I drop the old piece back in for a while — no stopwatch, no scorecard; just 'old roommate comes back for the weekend.'

That swap itself is T2 mechanics — a contrast that only tells me something changed. But what that contrast reveals after weeks of 'new normal' is pure T3: whether my brain was adapting up... or compensating down.

When I pull the new component out, if it feels like 'relief', it's a signal that something had been leaning on me while listening. If I feel a sense of loss, then it's a sign that the 'old' component earned its spot. *It goes back in.*

My rule is to keep everything else stable — same gear, familiar tracks, and the same starting level — then let my habits and the old piece's return tell me what the quick pass couldn't.

Demo Darling, Tuesday Night Keeper

A *Switch-Winner* is the system or tweak that grabs you in a quick comparison. Under Threshold 2 conditions — level matched, rapid swaps, brain in full vigilance mode — it rewards whatever flavor announces itself fastest.

It wins because it *demand*s your attention.

A *Session-Winner* is the setup that keeps pulling you back — the one you don't want to turn off. Over time, it runs a low Vigilance Tax and high immersion: you begin with a few tracks, but stay until the wee hours of the morning.

It wins because it *sustains your engagement*.

The see-saw tipped toward detection for the Switch-Winner. For the Session-Winner, it found its swing.

Some systems win both contests; when they diverge, the format hints at which one you're really scoring.

I still lean on quick A/B checks for Threshold 2 questions — troubleshooting a hum, verifying a gross mismatch, confirming whether a specific factor is readily audible. For that job, I want the spotlight. I just don't expect it to be a crystal ball for Tuesday night success.

A/B tells me what pops. *Long listening* shows me what soaks in. And *short tests* keep hitting reset before that process even starts.

“You're Just Adapting”

We've all heard some version of “*It's all in your head — you're just adapting to it...*”

In the listening chair, ‘just adapting’ falls short of describing the *two very different outcomes* we experience. Sometimes you end up hearing more than you did on day one. Other times you just stop noticing what was wrong. The question is... which one?

In a 2008 study, Nelson and Meyvis found that when participants heard vacuum-cleaner noise continuously, the irritation faded within minutes — adaptation quietly did its work. But when a pause interrupted the exposure and the noise returned, annoyance reset to at least full strength.

Applied to our own systems: if a flaw still grates after sustained, unbroken listening, *adaptation* isn't the diagnosis — especially after weeks in our rigs.

It's a flaw your brain refused to normalize.

Sometimes time doesn't dull the system — it sharpens you. The first night, a dense mix can feel like a wall. A week later, you're following the second guitar line, hearing the reverb tails, catching ghost notes you didn't even know were there... all seemingly without effort. And the tells show up in your behavior: the volume inches up, you stop track-hopping, and a ‘quick sit’ turns into a long session because you just can't leave.

That's acclimatization — adapting up. It's something your brain learned to hear, not learned to ignore.

An additive process.

Other times, you're not hearing more... you're just learning to ignore something. A hot edge on cymbals. A 'sparkle' that keeps turning into glare. The kind of irritation you can tune out if you try — vinyl clicks and pops, a refrigerator hum — but it still costs you energy. The volume drifts down. Your playlist turns into a search for something that'll connect. Sessions get shorter. Then, when you swap the old piece back in, the first feeling isn't "*oh wow, I lost detail*"... it's relief.

That's normalization — compensating down. Your brain learned to ignore, not absorb. **A *subtractive process*.**

The Accumulation Clock

My new listening room kept leaving me drained. A calibrated mic and raw Room EQ Wizard plots said I was good to go. My body said otherwise. Days later, I traced it to dimmed LED can lights making faint mechanical noise in the ceiling. I'd already installed the right dimmer switches to avoid the usual RF trouble, but the problem wasn't the switches. It was the bulbs. The moment I swapped the commodity LEDs for CREEs, the room relaxed... and so did I.

This wasn't some goofy one-off, either. I dug around and found a study where listeners heard music with inaudible high-frequency content above 22 kHz — content that only had an effect when the audible signal played alongside it. They couldn't consciously distinguish the two versions. Their brain activity and regional blood flow shifted anyway — documented by Oohashi and colleagues (2000).

At a different lab, with different instruments, Kuribayashi and Nittono (2017) found the same pattern using Bach: a measurably different brain state — one the listeners couldn't report. In Kuribayashi's data, that shift took roughly 200 seconds of continuous listening to emerge — like dark-adapted eyes that need unbroken minutes before they see what was always there.

Every A/B switch is like a flashbulb for the ears.

Here's the part that makes Threshold 3 so powerful: *it exposes what short tests miss*. T1 misses what you didn't think to measure. T2 can over-reward whatever the format spotlights. But T3 has a built-in solvent — the tincture of time.

Some shifts show up fast, while others take weeks before the verdict becomes trustworthy. This is the gap where the craft of listening lives.

Give it enough nights, and the solution presents itself: the playlist tells you, the clock tells you... and the gear either stays — or it goes.

Generations of listeners have been running this test — we just didn't have a shared framework for it.

SECTION 6:

The Voyage in the Chair

The music just ended. I pause for a few seconds, and realize I'm 'back'. My eyes open. It's 2 AM on a Wednesday morning, and I don't know where the time went. Perfection... at least for tonight...

That moment isn't false. It doesn't have a clean data entry on a plot, but it's data nonetheless. If you've been in this hobby long enough, you know exactly what it feels like, and it's a driving force for us.

We've all heard it: some version of "*You paid a lot, so you convinced yourself it sounds better.*" On the surface, it sounds like a completely rational argument... until the opposite happens.

The Inverse Placebo

Years ago, I engineered what I thought would be the 'ultimate' speaker cable. I sourced high purity copper foil, researched exotic materials and processes, and found an aerospace company willing to perform plasma vapor deposition of parylene onto it. I expected this cable to blow away anything else I'd ever heard in my system. I *really* wanted this to work...

When I hooked those cables up? The system sounded dark, closed-in... lifeless. I thought maybe, with more time... I'd warm up to them. I kept them in for a few weeks, dipping-in to compare against my steady references... and quietly hoping they — or I — would eventually come around.

It never happened.

I finally cut my losses and went back to what sounded 'right'. A \$2,000+ experiment now collecting dust in a closet, despite measurements that hit my targets... and despite all the hope, pride and sunk costs that should have favored these as 'the best'.

This is the *Inverse Placebo*. I should have loved that cable, and maybe paraded my new invention to my audiophile buddies. It didn't happen. That my ears overruled my investment is what ultimately matters. I have to own the failures alongside my wins. The hallmark of a healthy T3 process is all the gear that didn't make the cut. It may have won on paper and anticipation... but it *lost the chair*.

It's not just my story... If you've been in this hobby for any length of time, you have your own version — the power conditioner that took the grunge out of the AC, but also sucked the life out of the music, or the DAC that won all the online shootouts, but quietly exited your rack.

People sell gear for many reasons, but it's easy to see evidence in the wild: page after page of Audiogon listings that read "*pristine condition*", "*barely six months old*", etc... Much of this gear was bought with high expectations and many reasons to keep.

Different rigs, high expectations on the line, no one comparing notes — and the gear keeps walking out anyway.

The Inverse Placebo is the final score — the receipt issued when our ears say “*no*,” even when our plans, ego, or wallets are insisting “*yes*.”

Letting the Ghost Notes Breathe

So how does this all play out between the bench and the listening chair?

The ‘big stuff’ requires Threshold 1 to really hold. Speakers and the room dominate the equation. If you have a nasty 15 dB room mode, don’t expect a newer DAC, tube roll, or cable to bail you out. The *macro* is much higher up the food chain.

Threshold 1 is a baseline, not a crystal ball. I’ve experienced too many real-world interactions that evade that first checkpoint. I’ve learned measurements are best used to survey the landscape and lay some foundation for a good general read. But from there, I build an intuition, correlating bench data with long-haul experience and satisfaction.

Threshold 2 isolates the variables — troubleshooting a hum or confirming a change is readily audible. But as we’ve seen, its format *tilts the verdict*.

Threshold 3 is where we actually live with the result. It’s what I care about most. Livability often hides in the small nuances that evade even the most rigorous application of T1 and T2. These are the ghost notes of our playback chains. Bury them under the mental weight of a quick swap, or rush them through without enough time for your brain to fully absorb and sort the scene — technically ‘correct’ can turn music into a mechanical, sterile exercise.

Take these ghost notes away... and the soul gets sucked out of the music.

When your system continues to lure you for ‘play time’: That’s the Tuesday Night Test. You aren’t auditing. You just want to hit play and *disappear into the music*. When a component keeps surviving those nights, you know whether something just ‘interviewed well’ or truly earned its place...

The Lumley Resolution

This brings us back to where we started: those glorious Lumley monoblocs in Cross Plains, WI.

In the showroom, they were juicy and seductive. They were the ultimate T2 Switch-Winners. They dazzled my vigilant brain and satisfied my ‘*quick sip*’ ear.

But at home, with my complex, multi-layered instrumental music... they tangled everything up. *They failed the T3 livability test in my rig*. The solution wasn’t to re-home them.

Coming from solid state, I loved what tubes could offer — but the GEs’ lush midrange was like bathing in honey when my music needed just that little ‘chef’s kiss’ of richness instead. Those Svetlanas gave me the tube magic without going overboard.

That’s part of the audiophile journey: the swap didn’t make the amps measure meaningfully better in my room (T1). In a quick female-vocal demo, they might’ve even been less ‘seductive’ (T2). But they were the right elixir for livability (T3) — the Svetlanas detangled the dense harmonic layering and let my music breathe.

I still own those Lumleys, and they get action in my rig 35+ years later. They’ve survived decades of upgrades, comparisons, and ‘next big things’ because they passed the test that matters: they proved it’s possible to win the session without needing to win the switch. Same for my Lavry DA2002 DAC. It’s ‘only’ 24/96, but it’s run the gauntlet of newer DACs and still rotates in and out of my systems.

After decades of swapping and comparing, attending audio shows and local group sessions... living with my own gear has taught me that the finish line isn’t a ‘universal best’ for everyone. If that component doesn’t move me from a random Tuesday night into the wee hours of Wednesday morning, time will have its say.

I’ve come to realize that I’m not building a reference rig for anyone else. I’m building the best possible system for *me* — in *my room*, for the music I love.

The Release

We spend our lives in the *evidence room*, hunting for the difference. We train our ears to be precision instruments, analyzing the treble, measuring the bass, judging the imaging... *like an accountant reconciling a ledger*.

But eventually, the notebook closes. The lens cap goes on. The detecting brain finally clocks out and... silence.

You open your eyes, and check the time: It’s 2 AM and you wonder where the night went. You’ve successfully connected with the music you love, not the specs of the gear.

Threshold 2 asks, “Can I hear the gear?” Threshold 3 asks, “Can I forget it?” The ‘holy grail’ isn’t about finding the component that wins the argument — it’s about finding the one that lets you forget the argument ever existed at all. When you stop auditing the equipment and start living in the music, your gear vanishes... and *the music remains*.

> Sources used for this article <

Published May 12, 2026. © 2026 Chris VenHaus. All rights reserved.

Fidelity to What?

BY CHRIS VENHAUS | VH AUDIO



SECTION 1:

Interpreting the Scene

I was meticulous (borderline neurotic) about dialing-in my setup: proper placement of speakers, multiple subwoofers, bass traps, acoustic treatments on walls and ceiling. *Check*. I tamed the room modes with parametric EQ on my Xilica XD-4080 and got the blend spot-on for subs and mains. *Check*. Shot one last plot with Room EQ Wizard and everything looked awesome. I was super thrilled and stayed happy.

That's precisely what the tools are designed to do: reveal real problems, show whether a tweak is moving things in a useful direction, and get the system into a zone where the real listening can begin.

But I just couldn't leave well enough alone...

Like most of us, I tried to move the bar higher. That 1/6th octave 'bump and hump' I saw in the REW plot was the little devil on my shoulder, just daring me to 'fix' it. I eventually succumbed to the ~~darkside~~ hobby, pulled out the calibrated mic and dialed-in a couple narrow parametric EQ adjustments to make everything look 'right' on my computer screen. That hump and bump was electronically beaten into submission by science, and the sound was excellent.

For a while...

After a week or so, I felt uneasy while listening... something was 'off'. I defaulted to my normal routine: go back and undo the last tweak. I played the first track and quickly knew something changed for the better, but there was no obvious 'thing' I could grab onto as to 'why'. Over weeks of listening, I'd go back and dial-in some EQ to make that plot look better and each time, I couldn't detect *exactly* what was wrong... it just didn't sound 'right'.

Measuring 'better' in this instance didn't mean it sounded better.

I shared this experience with a recording engineer friend, and we eventually turned to the subject of mics. With tongue firmly in cheek, I asked if he could simply EQ a 'flat' mic into performing like the

legendary Neumann U47. His response — after a bit of a groan and slow shake of his head — was that you can get a mic to *look* like the U47 frequency response, but that doesn't make it a U47. My friend's engineering wisdom on that topic ended with: "*measurements don't always track the mojo.*"

When our conversation turned into a geek session over music we both grew up on, my drumming heroes were on deck: John Bonham & Ian Paice, as well as the work of Glyn and Andy Johns — both pioneers in the recording world, especially as it relates to recording drums.

I brought up Glyn Johns, and he also knew the legendary story of how his brother Andy was able to get that *huge* sound on Led Zep's *When the Levee Breaks*. This wasn't just a few mics put in front of his kit in a studio. You're hearing the character of a stone-built 18th century mansion (Headley Grange) with Bonham and his kit positioned at the bottom of a three story stairwell. Two heavily compressed Beyerdynamic M160 mics were positioned high above 'Bonzo' in the stairwell to capture the reverb. Then a Helios console, a Binson Echorec, and tweaked tape speed helped frame the scene. The snare wasn't just an ordinary Luddy 402: it transformed into an instrument now worthy of being struck by a drumming God. The 26" bass drum became so thunderous, Thor himself might consider another line of work.

This sound wasn't simply 'captured' and documented. It was *crafted*.

Nobody in that room thought these choices were cheating. It was craftsmanship... later recognized as pure *genius*. Singers may choose one mic for more bite, or another for more air. Guitarists know when one mic is wrong and another one is better for that guitar, amp or song. Engineers choose ribbons, condensers, angles, preamps, echoes, tape speed, and overload points because the scene capture is not clerical work. It's craft under pressure... with a tired back, a ticking clock, and musicians waiting on the talkback.

What we ultimately get to hear are the cumulative decisions that are rendered onto tape or into digits. *The engineer frames the scene.*

The rub is when you try to pin down what exactly the 'accurate' sound was. What version of the scene is correct? The musicians were the only people with a direct window to the performance. I should say... *multiple windows*: the drummer was inside the boundary of his kit, the singer heard the room through chest and jaw, the guitarist turned toward the amp instead of the room, the bass player locked with the drummer and feeling the note arrive in the fingers before anyone else. The musicians in that room were not standing in one shared acoustic truth. Then there's the engineer: they get their own special 'window' of truth — their monitors, room, experience, mood on that unique day. It happened once. Then it was gone.

That was the human side. Each person collected memories that only they can own.

The window we get to hear is captured not by a human brain, but by a device. The device wasn't 7th row center at Royal Albert Hall. A vocal mic may be close enough to catch spit on a hard consonant; a kick drum mic may live a foot from the resonant head where no human listener would dare stand. This is not a privileged seat. It's a tool position, chosen precisely because it hears differently. Classical listeners will note that a Brahms symphony could have a live 'reference', but *which one* — Row 12? The conductor's? The first oboist's? The recording engineer's mic position, which no audience member occupied?

Even in the most stable case we can imagine, the 'ultimate reference' dissolves the moment you try to pin it down. Every listener in the seats heard a slightly different acoustic event, and every memory of it kept moving after the lights came up and the venue emptied.

The 'reference' doesn't hold still.

Of course, there is always the recording itself — fixed, final, and indifferent to who was in the room. This assumption deserves a closer look.

SECTION 2:

Translating the Artifact

Geoff Emerick didn't approach Beatles sessions like a flat-transfer exercise, and a Shure SM57 on a snare certainly isn't a faithful 'witness'. Bonham under M160s in that Headley Grange stairwell wasn't clerical, and neither was Sinatra through a U47 mic.

Recording engineers intentionally select ribbons, condensers, preamps, overload points, console paths, etc... as part of the craft. The sound is being shaped before anybody at home gets a vote. Source coloration isn't the failure case in professional recording. It's the craft.

Even inside one mic manufacturer's own family, the voicing can shift: Peus (1997) measured the same Neumann capsule across circuit generations and found the response shifted. Same capsule family. Different result.

Engineers who use two microphones that are spec'd similarly on paper but hear one clear the window, while the other casts a haze on the glass know the feeling. In Sennheiser's own investigation early in the digital era, engineers complained about exactly that kind of transparency loss. CD-level 'cleanliness' unmasked what tape and vinyl had been forgiving for decades. The usual THD number wasn't picking this up, and the mic capsule's own distortion figures weren't either.

When Sennheiser dug deeper with a two-tone test, the culprit was found in the mic capsule itself: the diaphragm didn't move symmetrically into and out of its air gap. The subsequent push-pull MKH mic

design attacked this problem by making the air-gap conditions symmetrical. German *engineering* for the ‘win’ — but the ‘mother’ of this win was the engineers who heard what the measurements missed.

One of the most consequential microphones in history is the Neumann M50. The three-microphone recording technique known as *The Decca Tree* grew around it and defined classical recording for decades. Engineers routinely refer to the M50’s depth, image stability, and spatial realism as if they’re describing a perspective, not a frequency curve.

Because they are.

EQ can copy the M50’s on-axis balance. It cannot copy how it weighted the source. That choice happened in air. It happened before electricity. The mojo the recording engineer hears is real. It just wasn’t found in the published spec sheet.

My dear friend Brian Kelly — a retired jazz guitar player who was doing session work at the age of 16 in 1960’s Cincinnati, and had the ear of major record label execs — recently told me what it was really like inside the recording industry... the part of the sausage getting made we don’t usually hear about. The engineers who choose the mic, the distance, the console, the compression, etc... get paid to deliver a record that works *outside of their control room*: Cars. Clubs. Headphones. Giant floorstanders in tuned rooms or cheap boomboxes in ugly ones.

This doesn’t cheapen the artifact... it helps define it. The engineer’s high-wire act means the record has to survive real world usage — not just our purpose-built listening rooms and five or six-figure systems. That’s the reality of the world Brian saw.

From the other side of the glass, my friend [Christos Tsantilis](#) (multi-platinum mixing/mastering engineer and acoustician) told me about the multiple layers of editing that are embedded in the process, starting with the *first* engineer who captures the audio:

“What they’re hearing during the mix isn’t even the final version that gets released. The mastering engineer can reshape the record — sometimes subtly, sometimes significantly. So the final master is faithful to the mix, but it’s also a refinement — and often a different experience.”

Mastering legend Bob Ludwig added another facet to the mastering process when he explained the evolution of Avenged Sevenfold’s *The Stage*. He cut a dynamic master aligned with his usual approach, and the band signed-off. A few days later, the band got nervous it wouldn’t hold up against the loudness of their peers — so Ludwig produced a compromise version. After some back-and-forth, the band circled back and chose the more dynamic original for release. Same engineer, same album, days apart — and the ‘right’ call shifted more than once.

A record is an artistic statement, but it is also a commercial product built to survive translation, and ultimately... sell. What leaves the building is not the untouched event. It is the strongest available artifact of the reference made at that moment despite the compromises of that era.

Time. That's another factor. Do you recall that original vinyl pressing you owned, or the 180 gram reissue, '87 CD, 2003 remaster, or the 2015 hi-res issue that quietly replaced it in your library? If the original master was a fixed reference, should sanctioned revisions be viewed as 'vandalism'? Should we recalibrate our systems to accommodate the new master?

A great example of the shifting 'reference' can be found during the making of the Eagles *Hotel California* for the DCC Gold Disc series. Steve Hoffman — widely regarded as one of the best mastering engineers *ever* — received the master tapes, but felt the sound was muffled and bass-heavy. He pulled the safety copies — identical. It wasn't bad tape. *It was the way the album had been originally mixed.*

He nearly walked away from the project.

Hoffman then did something that would reverberate through the decades: he tracked down a similar pair of the JBL monitors the album had originally been mixed with. He then set them up in the middle of the room without any bass reinforcement, and... *that same tape was now workable.* The mastering decisions he made followed entirely from which speakers he was listening through, which included a 14 dB cut at 100 cycles (Hz). A measurement didn't point to those JBLs — it was his ears, experience, and judgment.

As he put it, he had two choices: sell a pair of JBL monitors with every gold CD or... master it so it sounded 'right' on everyone's stereo. He chose the latter. The tape didn't change. He changed his perspective with the equipment. Had he never found those JBLs, a different *Hotel California* would have shipped.

What Hoffman ran into wasn't a defective tape. The tape held the sum of all the choices made to that point — every mic choice, level tweak, compromise, and decision made through the locked master. But what it didn't hold — *what it couldn't hold* — was how it would be heard once it left the studio. That requires equipment, speakers, and a space where someone actually listens.

I think of the master tape as a fixed layer at the center of a longer chain, similar to a 'digital negative'. The negative is simply bits on a hard drive or memory card... until it's processed.

The artifact is fixed. The rendering engine gives it 'life'.

The recording industry has known this for a long time. Standards bodies wrote specs for what a control room should sound like, acousticians designed the rooms, and studios spared no expense to build them. But there was a problem: rooms built to the *same spec* still sounded different to the engineers who

worked inside them. Despite the recording industry’s best efforts, the same recorded artifact leaned differently in different rooms...

SECTION 3:

The Renderer and the Body

The Room’s Lean

Let’s imagine the recording engineer at the final mix, totally unconstrained. No commercial motivation, nobody in their ear about ‘*less bass*’, or ‘*more cowbell*’... Just mixing for the *best possible sound* for their ears, using those monitors, amp and electronics on that specific day. The best possible scenario...

Now... change *only the room*. Different walls, different ceiling, different floor. Play exactly the same recording. Same speakers, same electronics, same listener... Does it sound *the same* or will it render differently?

Most of us know what an aggressively treated, over-damped room sounds like: the speakers can measure clean and technically ‘correct’, but the music veers toward ‘sterilized’ and mechanical. The soundstage can become narrower, with less depth. But this doesn’t necessarily mean a dry room is ‘wrong’. A space *built for analysis* can help recording engineers hear balances, edits, and low-level problems that matter to the work. But there is a limit: the recording industry has moved *away* from overly dead studios, after musicians routinely complained their instruments didn’t sound or *feel* right.

It wasn’t just musicians.

Sound reproduction expert Floyd Toole has highlighted work that recording engineers preferred a strong, direct sound field for mixing — but preferred more reflections when listening at home. *Inside the profession*, Tervo and colleagues (2014) found a related split: mixing engineers leaned toward drier, higher-clarity environments, while mastering engineers leaned toward more reverberant ones. Even the ‘dry’ rooms in Tervo’s study didn’t cross over into a full-blown ‘anechoic’ state.

The chair that helps you work on a record is not automatically the chair that helps you live with it.

Bringing this into a more familiar world: in addition to mixing and mastering for platinum level artists, my engineer friend **Christos** also designs state-of-the-art mix rooms/studios. His recent project was for Rock & Roll Hall of Fame artist *Steve Miller*. When comparing ‘studio sound’ to what is preferred by most listeners in their own rooms, Christos’ response was:

“Most Hi-Fi systems introduce a low-end lift — often somewhere between 1 to 6 dB, below 100 Hz — which adds weight and excitement. In contrast, mastering rooms are typically tuned to be flat, or at most have about a 1 dB rise in that same region.”

If the ‘perfect’ mix only sounds perfect in that specific room, with that specific equipment, by that specific engineer on that day... don’t we end up with an audiophile version of the *Texas Sharpshooter Fallacy* — drawing a ‘target’ around holes... *after the fact*?

Glenn Meadows — who’s mastered over 1,300 gold and platinum records at Masterfonics — has said:

“I think that mastering is, and always has been, the real bridge between the pro audio industry and the hi-fi industry. We’re the ones that have to take this stuff that sounds hopefully good or great on a big professional monitor system and make sure it also translates well to the home systems.”

Gear Pressure

If we acknowledge the room leans on the signal, what’s the gear’s fingerprint on it? I love measurements, and use them every day for work... but like an xray, blueprint, or dyno on a car... they rarely offer the complete picture.

Noise matters. Frequency response matters. Stability, clipping behavior, output capability, linearity into defined conditions... all of this absolutely matters. But is testing a simple signal through a single component into a static load going to foretell what happens when a real music signal with full harmonic complexity moves through a chain of reactive components?

Audio Precision said it cleanly in a 2010 technical note: if you neglect reactive-load testing, you’re left speculating about actual use. Renardson made a similar point decades earlier: swap out the simple 8 ohm resistor ‘load’ you test an amp with, and put in a real loudspeaker. He wrote that THD with loudspeaker loads is typically about 10 dB higher, citing earlier work.

Of course this made sense when I first read it, because a speaker’s impedance (AC’s version of resistance) can swing wildly with frequency. It stores and returns energy back into the amp (back-

EMF), and generally behaves nothing like the well-mannered fixed resistor the spec sheet was tested into.

I like to think of it like this: a manufacturer tests your car's 0–60 time on a perfectly flat, dry, climate-controlled track — no traffic, no hills, no weather, no potholes. The car behaves beautifully. Take it on your actual commute: stop-and-go traffic, a steep hill, a gravel patch, a tailwind on the highway, or even... *off-road*? The engine has to constantly adapt to changing conditions, and now you notice behavior the spec sheet never hinted at. The test track wasn't lying — the test just never challenged the car the way the real world does. A speaker is your commute. A resistor is the test track.

This dynamic doesn't require an active source > load chain to show up. It also lives inside individual, passive components. Bateman (2002) measured a signal capacitor under real-world operating conditions, with signal running and DC bias applied. *The twist*: he used real-time measurement hardware instead of the conventional averaged approach. The distortion was there, but it wasn't constant. It appeared and disappeared at irregular intervals that were intertwined with the signal itself. The conventional FFT average smoothed all of that into a nice smooth line.

The spec sheet wasn't wrong — it just couldn't properly weight what appeared between samples. The standard methods were still telling the truth about steady-state behavior. They just weren't screening the whole 'movie'. Different tool. Different slice of reality.

When these distortion artifacts are injected on top of the harmonics and decay, the 'mortar between the bricks' is corrupted. A cymbal can keep its top-end energy but still lose the harmonic richness that gives it a unique personality. A snare or tom can keep its attack, but also feel like the overtones were bleached-out. This is the sort of thing that made me wonder: how do we define what 'clean' and 'transparent' really mean?

I learned the answer was always right in front of me: *time*. Trained normal-hearing listeners are capable of discriminating left-right timing differences below 10 microseconds, as Thavam and Dietz (2019) measured in spatial hearing. This tiny scale helped me gain new perspective on our sensitivity to timing differences, and made me wonder just how this might map to what we try to replicate in our hobby: creating a sense of 3-dimensional space that sounds *real*. Stuart and Craven (2019) mapped the question further into the audio chain: they modeled eight stages chained together and showed the accumulated timing errors can blow past their own 10-microsecond target without any single stage taking the blame. *Time behavior stacks*.

This brings me back to that 'hump and bump' I chased with my Xilica's parametric EQ. Each PEQ move was supposed to clean up the frequency response, and the Room EQ Wizard trace beautifully confirmed the 'fix'. What *I didn't understand at the time* was that a filter is still a filter: magnitude, phase, and group delay travel together. I had elevated that frequency plot as if it were telling the whole

story. My ears found the trap before I had an explanation for it. Later, it was Radlović and Kennedy's work (2000) that pointed towards a possible 'why': a correction can reduce the visible magnitude error in the SPL graphs I was using, while also disrupting the timing behavior. After many rounds of beating my head against the wall chasing that 'beautiful' looking frequency graph, I learned to stop treating large, narrow corrections above the bass range as 'free money'. The plot indicated 'better', but listening told me 'worse'.

The tool wasn't broken. It was just the wrong one for this job.

My listening room echoed this experience, but at a different scale. In 2011 I spent a good chunk of my summer doing what any slightly 'off' audiophile might call 'fun': armed with a calibrated mic and Room EQ Wizard, I measured every conceivable speaker placement option, xover settings, slopes, etc... and then walked two 1,000 watt subwoofers around the perimeter of my listening room inches at a time.

For the sub placement, I paid special attention to pesky nulls where *no amount of EQ* would correct the problem. What I didn't expect was how differently my room presented these nulls. Left wall: big suck-out (null). *Same spot* on right wall: smooth as butter. After some brief head scratching, I figured it out... The left side wall and front wall are framed next to a poured concrete foundation. The right and back walls are freestanding. Each construction type produced measurably different ringing that interacted with my room's modes. The conventional wisdom of 'put your subs in the corners' or 'keep them symmetrical in the room' didn't square with what the measurements alerted me to. The optimal solution was asymmetric placement to work with the room and the materials it was made with, not fight against it. The measurement didn't fail. It flagged something was 'off', and it was up to me to figure it out.

Feel the Vibe?

After my new speakers arrived, Bill Dudleston (Legacy Audio) was over helping me get things set up. During this session, Bill pointed out something that stuck with me: too much bass in a specific band doesn't just color the sound — it can make the room *physically pump*. Many of us are familiar with the 'one-note bass' you can get with some subs or poor speaker placement...

This was different. I wasn't just hearing this problem. My body could actually *feel* the room pumping.

I dig a nice headphone rig. But... it's a very different experience from my main rig. For me, I've always felt like the body 'listens' too. But it wasn't until I had a bad case of sinusitis, and I was taking a shower, when I spotted something interesting. As the shower water hit my head, I could hear it — even when I plugged my ears with my fingers. "Hmmm... *this just might mean something.*" This question became another audio related 'itch' that I needed to scratch...

When you think about it, it makes sense: our ears aren't just hanging out there, isolated in free air like a microphone. Our ear's 'guts' are mechanically connected to the rest of our body via fluid, tissue and bones... all good mechanical conductors of vibration.

Body-surface vibration under airborne low-frequency sound from loudspeakers was measured by Takahashi and colleagues (1999) and they found that chest and abdomen vibration increased with SPL in the tested range. The body behaves as a mechanical transducer — with the chest shown to be affected more than one's abdomen. This response also varies between body types: different build, different response.

This explains how the tactile punch of a bass drum or vibe from the bass player changes how I feel the groove. In the very best systems I've heard — *you don't just hear bass... You wear it*. You feel it in the sternum, the floor through your feet, the chair back, the little tightening in the gut when kick drum and bass line lock together and the whole track starts pushing air with purpose.

That's bass you feel *and hear* at the same time. If you're a headphone aficionado, you're probably all too aware of the arguments that "*headphones don't let you feel the bass*", as well as the counterargument that it's not needed for musical content. I enjoy both, so I usually kept my head down during those debates. But I did wonder... is body-felt bass really just for movie booms and rap thumps? My body said otherwise... and I eventually found science backed me up. Cameron and colleagues (2022) hit a live show with *inaudible* sub-bass: dancing jumped 11.8%. *That's pure body reflex kicking in*. Siedenburg's team (2024) found body-coupled bass cranks *engagement* far more than 'liking'. *You can't name it, but it activates and engages you*. No gimmick. Your chest keeps score, before the adjectives arrive.

Off-Rack or Tailored?

I've attended audio shows and have listened to the exact same track, in the same seat as some well-known gear reviewers. The opinion they came away with was sometimes miles apart from my assessment. Many of us have also listened to the same track on the same system at audiophile buddies' homes and came away with a differing impression.

We often see this during the various audio shows, where the 'best room' or 'best of show' never seems to settle onto a single, or even consensus 'best'. Wade into the audio forums and the spread of opinions is even wider than polling what the best pizza joint is in town.

This isn't just mood, training, or stubbornness. Reviewers usually know their stuff. Our audiophile buddies usually have good ears too. So what's at play here? Part of the answer was closer than I thought — literally attached to my head...

Our outer ear shapes every sound before it reaches our eardrums — and that shape is as unique as your fingerprint. Middlebrooks (1999) measured it: the sound shaping is mainly in the upper frequencies where your brain gets its directional and spatial cues — and it's not something you can just dial in or out with a tone knob or EQ. Wenzel and colleagues (1993) found the practical fallout when researchers simulated three-dimensional sound using a 'generic ear' model. Listeners actually *lost track* of whether sounds were in front or behind, above or below — and each person got confused in their own way. The 'generic ear' might be an interesting lab creation, but it's just not how people hear...

Going another layer deeper, the ear continues to show it's not 'off-the-rack'. There is a small opening deep in our ear's 'plumbing' (the helicotrema) that acts like a pressure relief valve, and affects how we perceive low frequency sound. The interesting part is the *size of this opening* varies from person to person. This means one person's "*Prodigious bass!*" assessment can be another's "*Crank that sub up, dude...*"

Cherri and colleagues (2024) found that listeners who all passed a 'standard' hearing test could vary widely when it came to finer details like separating a voice from background noise, or locating a sound in space. A normal audiogram is like passing a driver's vision test: you read the eye chart fine, but it says nothing about spotting a deer alongside the road or judging gaps in traffic. It's a 'normal' threshold label, not a perceptual fingerprint.

I was reminded of this when my local audiophile friend 'Dave' commented that he listened to drums differently after I walked through what I heard (as a drummer) in a track. These were cues that most drummers would reflexively listen for, but won't hit the radar of most non-drummers. I'm sure a pianist or cellist would absorb instruments differently than I do, as well.

Mixing engineer Jack Joseph Puig made a keen observation that really nails it:

"I feel everybody has what I call a 'preset' — something they default to. The preset is developed by the synthesis of all your experiences as you've grown up listening to music."

This explains a scene most of us know very well: experienced listeners sitting in the same room, same system, same sweet spot, same music — one is looking for an off-ramp, while the other gets lost in the music. Part of the rendering chain now lives in a chunk of cartilage, bone, and a few centimeters of very non-generic geometry. Our ear is the last piece of hardware — and it's truly unique for each of us.

SECTION 4:

Push Me... Pull Me

By the time I realized what was happening, my body was already there. Eyes closed, foot tapping, and lost in the music...

I'd spent days *in* 'engineer mode' for the listening session with some local audiophile buddies. Subwoofer tweaks, crossover tweaks, a new battery-operated network gadget... all the arrows in my quiver. Knowing the crew's music preferences, I scripted the tracks like I was calling plays for the first series of a Packers game. The session started, and as we moved through tracks, the comments started coming about soundstaging, bass, etc...

Then something *special* happened: Stevie Ray Vaughan's *Tin Pan Alley* came on and the 'demo' posture dissolved. I didn't *decide* to close my eyes. Nobody in the room did. Smiles appeared before words could be formed. After the track ended, Dave turned around to look at me and we all just *knew*. It was a moment. I bet most of the guys still remember it.

That's the state we love: time disappears. Our inner critic goes quiet. We stop thinking about the gear and dissecting every aspect of the sound. It's like being a kid on the playground. *Pure enjoyment*. This isn't the absence of detection; it's *what the detection looks like* when the body gets there before the inner critic.

It was at that point in the night that we also took a break and reloaded on some snacks, before getting back to the session. But that was it. The SRV moment was gone, and we didn't get back to that place on that night. It wasn't the system's fault, and the other recordings were actually quite good as well.

That's the fragile part. *Tin Pan Alley* took us all in, but during the break we drifted into tech talk about bits, sampling rates, tubes, comparisons... Once we were back in evaluation mode, we didn't recapture that absorbed state later in the same session. Other tracks still sounded good. Other moments still got through. But that high-water mark was behind us.

My practical takeaway after this experience and many others is: don't mix business with pleasure. If I'm in 'tweak' or comparison mode, it's an evaluation session with my propeller hat and laser level in hand. If I want to disappear into records, I protect a different night for that. Once the scorecard opens, I find it tends to stay open for the rest of the session. I was curious if there was a parallel to this in the recording engineer world, so I asked Christos. He confirmed it:

"I have to step away from a project before I can hear it as a music fan, instead of an engineer."

The analytical tools do their work — we’re being diligent. But this also confirms our need to ‘clock-out’ to fully enjoy the fruit of our labor. Our mental ‘see-saw’ leans on this movement...

The Scenic Route

In taste expertise, researchers found a pattern that might sound familiar. Beginners respond holistically. Then comes the enthusiast phase: more vocabulary, more analysis, more confidence. But also... a greater tendency to miss the whole while staring directly at it. With deeper expertise, the best tasters don’t become less analytical. They become more flexible. They can analyze when it serves its purpose, but can then return to the holistic gist when it’s time to just experience the thing in front of them (LaTour & Deighton, 2019). These studies were not run on audiophiles, but we’ve been running that experiment on ourselves for years. We know the score.

Most of us have lived some version of this. Early on, you just loved music. You may have had no language yet for image density, soundstaging, timing, timbre, tonal balance, or why one system felt more believable than another. You only knew that some nights a record pulled you in and some nights it didn’t. Then the ~~affliction~~ hobby enters your bloodstream. You learn the audiophile jargon. You start hearing deeper into recordings. You also start isolating traits, comparing components, lurking in forums, chasing A/B impressions, building mental checklists, and discovering that the same sharpened attention that helps you hear more can also box you in. *This stage matters*. It builds the ear. It teaches discrimination. It is how you learn that a system can be spectacular in pieces and still fail to satisfy in whole. But it also creates a recognizable trap: the food critic at the barbecue, the photographer who can’t look at a landscape photo without reverse-engineering the lens used, the audiophile who cannot unhear the squeak in Bonham’s bass drum pedal, the tube hiss in the noise floor, or the last post they read on a forum an hour earlier.

Then, for many serious listeners, something surprising happens. They circle back toward something that may look like the beginning... but it’s not. It’s now familiar ground that’s been earned. The analytical tools are still there, but are now internalized and deployed strategically. They no longer block the entrance. Immersion never went anywhere — the analytical phase just pushed it aside. Evolved experience lets you use both pathways more strategically (Melcher & Schooler, 1996).

This is where Bob Ludwig’s comments resonated for me: he colloquially mentioned that the ‘left-brain-technical’ and ‘right-brain-creative and musical’ are both real demands, but they are not supposed to run full-tilt at the same time. Saturday morning can be for tweaking or taking a new component for a test ride. Another evening can be for wine, a stack of albums or tracks and no agenda at all. Both are legitimate. Both are part of the craft.

If you’re carrying a hammer, you’re always looking for nails. *Analysis is a tool*, and sometimes you need to learn when to put the tool away. Analytical mode is the chisel. Immersive mode is where you

step back to see what you made. What changed is authority. The tools stopped inserting themselves into every session by default. Rigor isn't rejected... it's just parked after serving its purpose. It is the last step in becoming serious enough to enjoy what your work uncovered.

This explains the listener side of the ledger. But it doesn't explain the night Stevie Ray Vaughan entered the room. Everyone in my listening room had heard SRV's *Tin Pan Alley* before... we all know it's a great recording. On a different system it might have been merely 'really good'. That extra effort I put into tweaking my rig gave that recording exactly what it needed to fully breathe. While the system was still delivering an hour later, that special moment was in the rearview mirror.

Same rig. Different part of the session. We didn't decide to hear differently; we just did.

Some nights I settle in and the room opens; some nights I can't quite get there — I get anxious, or restless, and I call it a night early. It is not the gear. Sometimes... it's just me. By '*me*', I'm not referring to just mood or mindset — it's the whole package, including what's in my blood...

Hippies in the 60's confirmed this 50+ years before researchers did... but for the sake of science, Ferreri et al. (2019) showed how body chemistry can affect people in a double-blind controlled study: they gave listeners two different drugs that affected dopamine. One dampened musical pleasure (risperidone), the other amplified it (levodopa). Our body chemistry affects how we listen.

SECTION 5:

The Final Polish

Of course, the whole '*Sometimes... it's just me*' is more than body chemistry. I bring other variables to the game. Maybe I'm set to test a new component or tweak, maybe it's to dig around for some new music, or maybe I'm just there to go for a 'ride'. My approach for that listening session can be deliberate, but it can also get hijacked by wildcards I am unaware of. Either way, how we frame a listening session affects how we listen, which aligns with Enstroem & Schmaltz (2022).

We always bring something into the room with us before the first note even plays.

Many of us already know this — it's why we build routines that might look fussy (maybe even neurotic) to those outside the hobby, but *perfectly reasonable* from our viewpoint. We usually listen for a few sessions and let it soak in before we flip back to our familiar baseline after a while. We usually know not to put too much weight in that first hour... good or bad. The science behind this made perfect sense to me when I dug-in: when you judge a change against a system you're already intimately familiar with, you're anchoring — and with that foundation, the comparison stays more stable than bouncing between two unknowns (Mathias et al., 2020). It's like comparing a test drive to the car I already own, versus comparing two different models back-to-back that aren't my daily driver.

Similar to a test drive or reading an article, when we listen to music, the peak and the ending are sorted and weighted differently in our memory (Rozin, Rozin, & Goldberg, 2004). It's prioritization. It's efficient. Our brains even start this assembly *while we listen* — what we hear in the first half is already being updated by what comes after (Mercer & McKeown, 2010). The version of the session we carry forward gets shaped going in, updated as it unfolds, and compressed into a mental 'zip' file we reach for later. Each of us takes away a different, *very personal* sketch of that same scene.

The Chase... and the Hobby

Spec sheets, searching for user reviews, the late-night forum scroll. We've all done it. It's part of the hunt. It's fun. Then comes waiting for the tracking number, checking it multiple times until it hits the porch, and finally the unboxing and the install preamble for the first power-up and 'liftoff'. I'm guessing astronauts about to fly into space might be able to relate to how we feel... This is a different excitement than the actual listening. The distinction between 'wanting' (the pursuit) and 'liking' (enjoyment) gives these elements of our hobby a shared ecosystem, without minimizing either side of it (Berridge & Robinson, 2016). In practice, these two elements feed off one another.

The 'chase' is the acquisition phase of the listener's education. The quiet night is where that education gets tested against actual satisfaction. The components or tweaks that don't survive are just part of the tuition, and the insight we gain from this makes us better listeners. The chase informs the listening. The listening sharpens the next chase. Both are important.

The Anchor

Every new listening session anchors to something I've already lived with long enough to know by feel. That's the reference — not a number on a bench, not someone else's benchmark, but the one I built in my room over years of calibrating to reference tracks, my personal library, and new music finds. This means '*my faithful*' was never someone else's reference... and yours may not be mine.

I believe the concept of 'fidelity' has been doing two jobs for us:

Signal fidelity describes how accurately a system preserves the electrical information it is given. Measurement helps us detect a drift, and acts as a benchmark.

Performance fidelity creates both upstream and downstream questions: upstream, when the engineer applies their craft, and downstream — when we tweak our systems.

A familiar example most of us already know: turn the volume up and the *sound is re-shaped*, it's not simply a uniform lift across the full spectrum. This reality of how we perceive sound was documented all the way back in the 1930's when the *Fletcher-Munson equal loudness curve* was established.

While looking into how the updated loudness curves were built, I learned they were aggregates pulled from young adult listeners with ‘normal’ hearing. The surprise hidden behind these smoothed curves: the reported numbers implied a ~20 dB spread in listener judgments. It would seem that if a microphone is to be the ultimate witness, it requires a ‘jury’ of other microphones...

Sean Olive expanded beyond equal loudness research when he studied listeners’ *preference* for bass loudness. His numbers showed a 17 dB spread in the level of preferred bass across his test group. At least *some* of this spread is likely bound to each listener’s unique anatomy: known structures inside the ear proven to shape perception. What struck me about this research is how similar these numbers are to the scatter noted in the equal loudness curves...

Generic ears and averaged plots are approximations... they aren’t individuals.

The auditory measurement world has acknowledged the limitations of its tools, and where real listeners are required to finish the job. In fact, the field’s benchmark textbook by Bech and Zacharov cuts right to the chase: there is no direct measurement for what we actually perceive.

Of course, we all have our own biases, too... But it’s not the whole story. To neatly wrap ‘bias’ around the differences, it would require a stable, fixed target for the ear to be ‘wrong’ about. If there’s one common thread I’ve found in all my years of study and experience on this subject, it’s shown just how *slippery* that target truly is...

Beyond the measurements, and validation... replicating an *experience* is harder to pin down. I wondered how this translates to the real-world of mastering, so I went back to my ‘well’ of sound engineering expertise (Christos) and asked what mastering engineers listen to at home, for enjoyment. I suspected they’d want to replicate their studio sound with the same equipment. His response: “*Most engineers have a different rig at home because they need something to represent the home environment, not the studio.*”

This came as a bit of a surprise to me, so I poked around and found that while *some* engineers try to take ‘work’ home with them to more closely resemble their studio sound, most use different equipment at home. The bigger surprise to me was many use vacuum-tubed equipment in the chain: Bernie Grundman (McIntosh), Steve Hoffman (Audio Note Ongaku), Bob Ludwig (Manley Steelhead), Dave McNair (VAC Master). The same engineers who are known for crafting some of the best-known recordings ever made, also craft their home systems differently than their studio rigs...

An engineer selects a specific vacuum tube microphone, positions it carefully to get the sound that’s just ‘right’. Decades later, an audiophile fires-up their vacuum tube monoblocs and tweaks the rake angle on their speakers to get locked-in. One of these acts we call ‘craft’. The other?

SECTION 6:

A System for One

Dennis arrived and the rest of our audio group started to get settled for the evening. My rig was in that rare place where I'd finally stopped fiddling... Tubes re-biased, subs tweaked... everything buttoned down. Dennis was still early to audiophile gear, but he was as big a music nerd as any of us, and still regularly attended live shows. At some point the subject of 'best' bass players came up. I mentioned John Entwistle and Dennis totally lit up... "*The Ox!!!*". I verbally nudged him into taking the sweet spot, while I quietly queued-up The Who's *Eminence Front* while the group continued to debate the best bass players. I then hit play... let the song run, and the room changed. The background conversation stopped, fingers started tapping, and we were all pulled into the layered Glyn Johns' mix. The intensity of Entwistle's bass by mid-track was a force of nature. When the track ended... a brief pause... and then Dennis turned back to me and nodded with a huge grin.

Later, at other group meetups and in other rooms, Dennis would often mention how great that track sounded at my place that evening. His place, somebody else's, a driveway after the music stopped — he'd circle back to that experience of hearing *Eminence Front*. There was never a mention of any specific gear.

None of us can recall what tubes were in the amp, what DAC was used, or the last filter settings I spent that summer dialing-in. What's remembered is how the room went quiet when that familiar track arrived with far more weight than we were ready for, and the stillness afterward...

Most nights don't become stories. A playlist queues up... and somewhere a few tracks in, I've stopped managing the sound. Different nights, different tracks, different people settling into chairs. After a song has filled the room that way... we all know before anyone has to say a word.

Eventually, we queue-up the next track and see where that one takes us...

> Sources used for this article <

Published May 12, 2026. © 2026 Chris VenHaus. All rights reserved.

About the Author

Chris VenHaus founded **VH Audio** in 2003. He's played drums for over 40 years, and loves to build electric guitars and make landscape photos. Chris still uses bench tests daily, but Tuesday nights in the listening chair make the final call.

Read the trilogy online: vhaudio.com/three-thresholds-trilogy

News, system tweaks & DIY tips: [sign up for the VH Audio newsletter](#)

Questions or comments: vhaudio.com/contact-us



VHAUDIO.COM · V-CAP.COM