Silent Suffering: Why Your Service Operation Looks Healthy and Isn't

By Rahul Jindal

There is a particular kind of operation that is failing and does not know it. The dashboard is green. Satisfaction sits at a respectable number. Cases close inside the service-level target. And yet, if you sit in the rooms where the people it serves actually talk, you hear something the dashboard never reports: skepticism, resignation, a joke at the function's expense, a quiet aside about how everyone just routes around the official channel now. The metrics say the operation is healthy. The people say it is not. Both are telling the truth. The gap between them is silent suffering, and most service organizations have no instrument that can see it.

This matters most in the operations nobody chooses. A consumer can switch banks. An employee cannot switch HR. When the user is captive, the usual market signal for a bad experience, the customer leaving, never fires. The dissatisfaction does not disappear. It changes form. It becomes a workaround, a reputation, a meme on a local alias, a slow erosion of the willingness to even try the front door. A captive audience is the easiest audience to fail quietly, because the one number that would force you to notice, attrition, is structurally unavailable to you.

The iceberg is real, and it has been measured for forty years

The foundational research here is old, deliberately cited, and has been replicated across hundreds of studies. In the 1970s and 1980s, TARP (the Technical Assistance Research Program, led by John Goodman) ran the work that became the seminal benchmark for how dissatisfied people actually behave. The headline finding has held up for four decades: the complaints a leadership team hears are a tiny, unrepresentative sliver of the dissatisfaction that exists.

In the TARP data, of the users who hit a problem, only roughly one to five percent ever escalate it to a manager or to headquarters. Around forty-five percent mention it to a

front-line representative and go no further. And roughly half never say a word to anyone in the organization at all. For business-to-business relationships, about a quarter raise nothing. The ratio between problems experienced and complaints that reach the people who could fix the system ran anywhere from six-to-one up to two thousand-to-one, depending on how many layers sat between the user and the decision-maker. The more structure you put between the suffering and the leadership, the more of it gets absorbed and silenced on the way up.

> ***"The complaints leadership hears are not the small end of the problem. They are a biased sample of the problem, selected for the people who still believe complaining is worth the effort."***

Then comes the finding that should worry anyone running an internal service function. Complaint rates are not uniform. They depend heavily on what kind of problem the user hit. When the problem is a clear monetary loss, money out of pocket, complaint rates run high, fifty to seventy-five percent. When the problem is mistreatment, poor quality, or plain incompetence, the kind of problem with no receipt attached, complaint rates collapse to somewhere between five and thirty percent.

Look at what an HR case, an IT ticket, or an internal-support interaction actually is. It is almost never a clean dollar figure. It is "I was treated like a number," "the answer was wrong and I had to figure it out myself," "nobody owned it," "I asked three times." These are precisely the soft, hard-to-itemize failures that the research says people are least likely to report. The categories of pain most characteristic of internal operations sit in the five-to-thirty-percent band. The structure of the work guarantees that most of the dissatisfaction it produces will be suffered in silence. This mapping is an inference, not a directly measured statistic for HR operations, but it follows cleanly from the problem-type data, and it should be the null hypothesis any internal-services leader starts from.

Silence is not neutral. It is the expensive failure mode.

The comfortable assumption is that a user who does not complain is a user who is basically fine. The data says the opposite. TARP measured the relationship between how a

problem was handled and whether the user intended to stay, and the order is consistent and counterintuitive.

For small problems, the loyalty ladder runs like this. Users whose issue was resolved quickly stayed loyal at over ninety percent. Users whose issue was resolved, but not fast, stayed at around seventy percent. Users who complained and were left unsatisfied stayed at forty-six percent. And the users who never complained at all, the silent ones, stayed lowest of the group, at thirty-seven percent. For high-stakes problems the same ordering holds and the gap widens: complained-and-resolved sits at fifty-four percent, complained-and-unresolved at nineteen percent, and the silent non-complainers at just nine percent.

Read that ordering twice, because it inverts the intuition that built your dashboard. The silent user is less loyal than the user who complained and whom you failed to satisfy. Voicing a complaint, even one you handle badly, keeps someone more engaged than suffering quietly. Silence is not the absence of a problem. It is the presence of a user who has already concluded that telling you is not worth it. These are repurchase-intention figures, not observed behavior, so treat them as direction rather than physics. The direction is unambiguous: the quieter the channel, the worse the underlying health, and the bigger the stakes, the more that is true.

> ***"The silent user is less loyal than the user who complained and whom you failed to satisfy. Silence is not the absence of a problem. It is a user who has decided you are not worth the breath."***

In an internal function, the word "loyalty" needs translating, because nobody is repurchasing. The captive-audience version of loyalty is trust, and trust spends itself in specific, observable ways: whether people use the official channel or build a shadow one, whether they escalate through your process or through a friend who knows someone, whether they speak about the function with respect or with a knowing eye-roll. When the silent-suffering rate climbs, none of your service metrics move. What moves is the function's standing, and that shows up everywhere except the place you are looking.

Why your satisfaction score cannot see any of this

The instrument most operations trust to tell them they are fine is the one least able to detect the problem. A satisfaction survey is answered by the people who answer surveys. That sounds tautological because it is, and the tautology is the whole issue.

Survey methodology has a name for it: non-response bias. When response rates are low, the people who respond are systematically different from the people who do not, so the average you compute is not a measurement of your population. It is a measurement of your respondents. A peer-reviewed study of the Press-Ganey patient satisfaction survey is a clean illustration. The response rate was sixteen and a half percent, well under the benchmark range for the field, and the people who responded differed measurably from those who did not on age, sex, and insurance type. The exact demographic skew there belongs to healthcare and should not be copied into an HR context. The principle is universal: a low-response score is not a noisy reading of the truth, it is a precise reading of a self-selected subset, and you have no idea which way it is bent.

Now layer the captive-audience problem on top. Who actually fills out the post-case survey in an internal operation? Disproportionately, the people at the two ends: the genuinely delighted and the incandescent. The vast middle, the people who got a mediocre answer, sighed, and got on with their day, are the ones the research says were already drifting, and they are exactly the ones who do not bother to rate you. The worse the everyday experience gets, the more people give up on the feedback channel along with everything else. So the survey can drift upward at the precise moment the operation is rotting, because the disaffected stop answering before they stop caring. A rising satisfaction score on a falling response rate is not good news. It is the single most dangerous reading on the board, and almost nobody flags it.

> ***"A rising satisfaction score on a falling response rate is not good news. It is the most dangerous reading on the board, because the disaffected quit the survey before they quit caring."***

Where the suffering goes when it does not go to you

Dissatisfaction is conserved. It does not vanish when it goes unspoken to you; it relocates. In an internal operation it tends to settle in four places, none of which appear on a service dashboard.

- **Workarounds and shadow process.** People stop using the official path and invent their own: the colleague who actually knows, the side spreadsheet, the direct message to someone three levels up. Every shadow process is a silent verdict on the real one. It is also invisible to you, because by definition it does not generate a ticket.
- **Reputation and the meme layer.** The frustration surfaces as culture: the joke on the local alias, the rant in a team channel, the shared knowing look when the function's name comes up. This is real data about your operation's health. It is simply held in a medium your measurement system was never pointed at.
- **Learned helplessness.** After enough low-value interactions, people stop expecting better and stop asking. This reads, on every chart you have, as improvement: fewer tickets, fewer escalations, fewer complaints. It is the opposite. It is the sound of an audience that has given up on you.
- **The credibility tax.** Once a function is quietly held in low regard, everything it does costs more. Its announcements are discounted. Its change efforts meet a wall of skepticism. Its good work is assumed to be an exception. None of this is on a balance sheet, and all of it is expensive.

The thread connecting all four is that they are real, consequential, and structurally invisible to the standard kit. The operation is accumulating a liability it has no account for. By the time it becomes legible, usually as a reorg, an engagement-survey cliff, or a new leader brought in to "fix the function," the suffering has been compounding for years.

The instrument: stop asking if people are happy

Site reliability engineering solved a version of this problem a decade ago, and the borrow is exact. You do not assess whether a service is healthy by surveying users about their feelings. You measure latency, error rate, saturation, and traffic directly, in the system, continuously, whether or not anyone files a bug. User reports are a lagging, biased supplement to that telemetry, never the basis of it. Service operations have the telemetry

available and mostly refuse to treat it as a health signal. The fix is to build a layered measurement model where the load is carried by data that does not depend on anyone choosing to speak.

Four layers, in ascending order of how much silent suffering they can catch.

Layer 1: Voiced signals, weighted by who is missing

Surveys and complaints stay in the kit. They are necessary and insufficient. The discipline is to stop reading the score in isolation and always read it against its response rate. A score is only as trustworthy as the share of the population that produced it. Track the response rate as a first-class metric in its own right, watch its trend, and treat any divergence between a rising score and a falling response rate as an alarm, not a footnote. Voiced signals tell you about the people who still believe you are listening. That is worth knowing, as long as you never mistake it for the whole.

Layer 2: Operational telemetry, independent of feedback

This is the layer most operations already have the data for and do not read as health. These signals fire whether or not the user says anything, which is exactly why they cut through the silence.

- **Reopen rate.** The share of resolved cases that bounce back open, calculated as reopened divided by total resolved. It is the shadow metric that first-contact-resolution and closed-ticket counts hide: a case marked solved that the user had to reopen was never solved, and naive resolution stats happily count it as a win. As a rule of thumb the ITSM field treats under five percent as strong and over ten percent as worth investigating. A reopen is technically a voiced re-contact rather than pure silence, which is what makes it so useful: it is a hard, countable trace of an inadequate first resolution that no survey was needed to surface.
- **Transfers, escalations, and bounce.** How many hands a case passes through before it is resolved. Every transfer is effort the user is absorbing on your behalf and a point where ownership was ambiguous. High transfer counts are a friction signal that no satisfaction number will show you.

- **Cycle time and its distribution.** Read the tail, not the average. The average hides everything. The ninety-fifth percentile of resolution time is where suffering lives. An operation can hit its average target while a long, quiet tail of cases rots for weeks.
- **Backlog aging.** The age profile of open cases, watched as a moving distribution. A growing old-case tail is dissatisfaction being manufactured in real time, fully visible, before a single survey goes out.
- **SLA breach patterns.** The breach count is the start. The clustering is the signal. Breaches that concentrate by case type, team, region, or policy area are a map of where the system is structurally unable to deliver, pointing you at the cause rather than the symptom.

Layer 3: Effort, the feeling the user actually has

The single most predictive question you can ask is not whether the user was satisfied but how hard they had to work. The CEB research, now Gartner, behind Customer Effort Score is decisive on this. A user who had a high-effort interaction is far more likely to become disloyal, and effort predicts loyalty markedly better than satisfaction does. In the original study, high-effort interactions left users dramatically more likely to defect and to speak badly of the service, while low-effort ones barely moved the needle. For a captive internal audience, "disloyal" reads as "will route around you next time and tell their team to as well." Effort is the early-warning version of the loyalty erosion the TARP data measured after the fact. Measure it, and weight it above raw satisfaction.

Layer 4: Textual and dark signals

The richest, hardest, and most neglected layer. The pain that never becomes a survey response is sitting in plain text already: in the free-text notes of the cases themselves, in the resolution comments, in the reopen reasons. Sentiment and theme analysis over case notes is a voice-of-the-user instrument that does not require the user to fill anything out, because they already told you inside the case. Beyond the case system, the dark signals are the leak points from earlier: the prevalence of workarounds, the recurring themes on internal aliases, the markers of learned helplessness such as falling contact rates paired with flat or worsening operational telemetry. These are harder to instrument and easier to dismiss, and they are where the truest reading lives. The honest caveat is that the efficacy

of text and sentiment analytics in this setting is more asserted than independently quantified, so build it as an experiment with a feedback loop, not a finished oracle.

> ***"You do not measure whether a system is healthy by asking its users how they feel. You measure the system, continuously, whether or not anyone files a complaint. Service operations have this telemetry and refuse to read it as health."***

What to start tracking, by function

The four layers are the model. Here are the concrete metrics they translate into, because a framework you cannot instrument on Monday is just a nicer way to stay blind. Start with a universal core that applies to any service operation, then add the few signals specific to your function. Almost all of these are computed from data you already hold in your case system. None of them require asking a single user how they feel.

The universal seven (any service operation)

If you track nothing else, track these. They fire whether or not anyone speaks, which is the whole point.

1. **Reopen rate.** The share of resolved cases that bounce back open. The single best detector of a closed case that was never actually solved, and the metric that naive first-contact-resolution hides.
2. **Cycle-time tail, not the average.** The ninety-fifth percentile of resolution time. The average is where suffering hides; the tail is where it lives.
3. **Backlog aging.** The age profile of open cases as a moving distribution. A growing old-case tail is dissatisfaction being manufactured in real time.
4. **Reassignments per case.** How many hands a case passes through before it resolves. Every transfer is effort the user is absorbing and a point where ownership was ambiguous.
5. **Escalation rate and SLA-breach clustering.** The clustering matters more than the raw count: which case types, teams, or regions concentrate the failures, which

points you at the cause.

6. **Effort score plus repeat-contact rate.** A Customer or Employee Effort Score at case close, and how often the same person comes back about the same thing. Effort predicts disengagement earlier than satisfaction does.
7. **Survey response rate, tracked as its own metric.** Put it on the same chart as the satisfaction score, always. A rising score on a falling response rate is your loudest alarm.

HR and People Operations

The case mix is the trap here. A payroll error, a benefits question, a leave request, and an accommodation case run on completely different clocks and stakes, so an aggregate resolution-time number tells you almost nothing.

- **Resolution-time tail and backlog aging, broken out by case type.** Watch the sensitive categories most closely: employee relations, investigations, accommodations. Those are the highest-stakes silent-suffering cases, and a long quiet tail there is the most expensive thing your dashboard is not showing you.
- **Reassignments per case and Employee Effort Score.** Count how many HR people an employee had to re-explain a sensitive situation to. Weight that effort above HR-CSAT.
- **The dark signal that matters most here: back-channel resolution.** The share of questions answered by a friendly HRBP or a manager instead of the official case system, plus any quiet drop in cases-per-employee in a given org. A falling case count in a frustrated population is learned helplessness, not a happier workforce. Add sentiment on case free-text and on internal-alias chatter about HR.

IT and the service desk

- **Reopen rate and first-contact resolution read together, by category.** A high apparent FCR with a high reopen rate means tickets are being closed, not fixed.
- **Ticket bounce.** The number of reassignment hops between teams before resolution. The clearest friction signal a service desk produces.

- **Failed self-service.** Knowledge-base searches that return nothing or get abandoned. This is unmet need that never becomes a ticket, so it is invisible to every ticket-based metric you have.
- **The dark signal that matters most here: shadow IT.** The prevalence of unsanctioned tools and the colleague who knows the workaround, plus a drop in ticket volume on a service everyone privately considers broken.

Customer support

The external sibling, and the one case where the user can actually leave. That gives you a churn signal the internal functions do not have. Use it, but do not lean on it, because most dissatisfied customers leave without a word, and by the time churn moves the damage is done.

- **Customer Effort Score, weighted above CSAT.** This is the metric's home ground. A high-effort interaction predicts defection far better than a low satisfaction rating.
- **Repeat-contact rate and channel-switching.** How often a customer had to come back, and how often they were forced from chat to phone to email to get one issue resolved.
- **Silent churn, watched against engagement.** Accounts that quietly reduce usage or fail to renew without ever filing a complaint. Read a drop in contact volume against product usage, so you can tell a happy customer from one who gave up.
- **The dark signals here reach outside your walls.** Sentiment on chat and ticket transcripts, and public review and social sentiment, which is where customers vent the complaints they never send you directly.

Finance, procurement, facilities, and other shared services

The same universal seven apply, plus one fingerprint specific to process-heavy operations: rework. Track invoice and purchase-order exception and re-work rates, the tail of approval-cycle time, and repeat work orders raised against the same asset. Rework is the operational signature of a process the user had to fight, and it shows up in your own data long before anyone files a complaint about it.

From "we have a satisfaction score" to "we can measure operational health"

Most service organizations sit on the first rung of a ladder they do not know they are on. It is worth naming the rungs, because the gap between them is the whole argument.

1. **Survey-only.** Health is a satisfaction number. Response rate is unmonitored. The operation is effectively blind to everyone who did not answer, which is almost everyone.
2. **Survey plus volume.** Satisfaction alongside ticket counts and closure rates. Better, but every metric here can improve while the operation gets worse, because falling volume can mean learned helplessness and fast closure can mean premature resolution.
3. **Operational telemetry as health.** Reopen rate, transfers, cycle-time tails, backlog aging, and SLA clustering are read as the primary health signal, independent of who chose to speak. This is the rung where silent suffering starts to become visible.
4. **Effort plus text.** Effort is measured and weighted above satisfaction, and case-note sentiment is mined as a standing voice-of-the-user stream. The operation now hears the people who never filled out a survey.
5. **Closed-loop and dark-signal.** Every detected silent failure routes to an owner and a fix, and the function actively listens to its own reputation, on aliases, in workarounds, in the gap between contact rate and operational health. The operation can find the suffering before the suffering finds a reorg.

The cheapest first move, and the most uncomfortable, is to put the response rate next to the satisfaction score on the same chart and sit with what it implies. Almost every operation that does this discovers that its reassuring number rests on a single-digit slice of its population. That discomfort is the start of measuring honestly. The second move is to pull two pieces of telemetry you almost certainly already collect, reopen rate and the cycle-time tail, and watch them as health signals for a quarter. They will disagree with your satisfaction score. The disagreement is the point. It is the silent suffering, finally showing up on an instrument.

If you want to place your own operation on this ladder, the Silent Suffering Scorecard scores the four layers in about four minutes and tells you which one to build next.

The function that measures its own silence

The deepest reason silent suffering persists is that the measurement system and the suffering are pointed in opposite directions. The dashboard rewards closure, speed, and the goodwill of the people still willing to answer. The suffering accumulates among the people who stopped answering, in channels the dashboard cannot read. An operation can optimize its visible numbers for years while its real standing decays, and nothing in its instrumentation will object.

The way out is not a better survey. It is a decision to treat a service operation the way an engineering team treats a production system: as something whose health is measured directly and continuously, where the absence of complaints is never read as the presence of health, and where the most important signals are the ones that fire when no one is speaking. The functions that learn to hear their own silence will be the ones still trusted in ten years. The rest will keep reporting green, right up until the day someone finally asks the room, and the room answers.

Take it with you

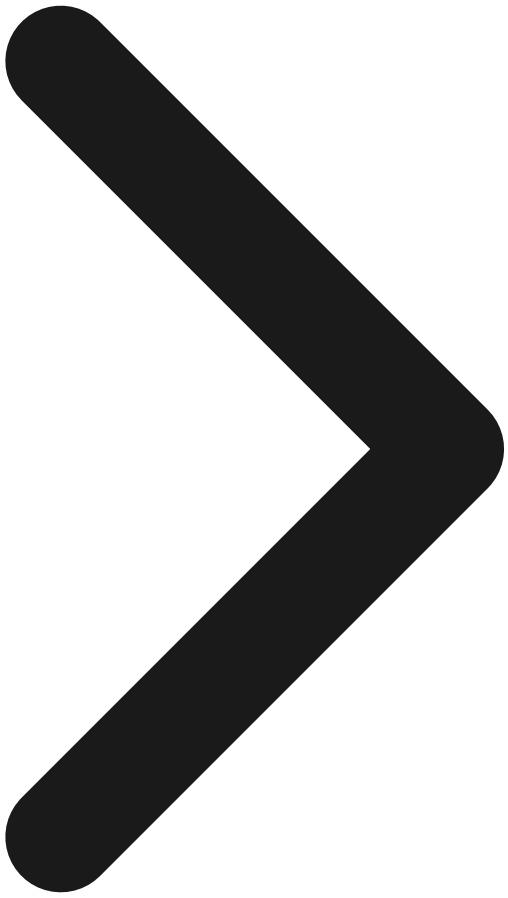
Email this as a LinkedIn pack

Get a feed-ready LinkedIn post (under the 3,000-character cap), a long-form LinkedIn article version, the hero image, and an editable document version of the full essay, delivered to your inbox. Ready to post.

Email this

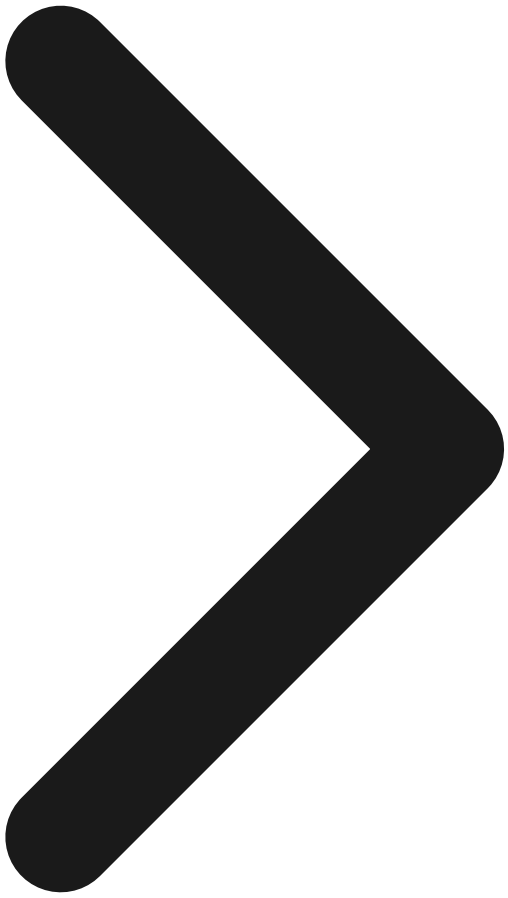
Related Insights

Operating Model 8 min
The Operating Model Every Enterprise Services Org Is Missing



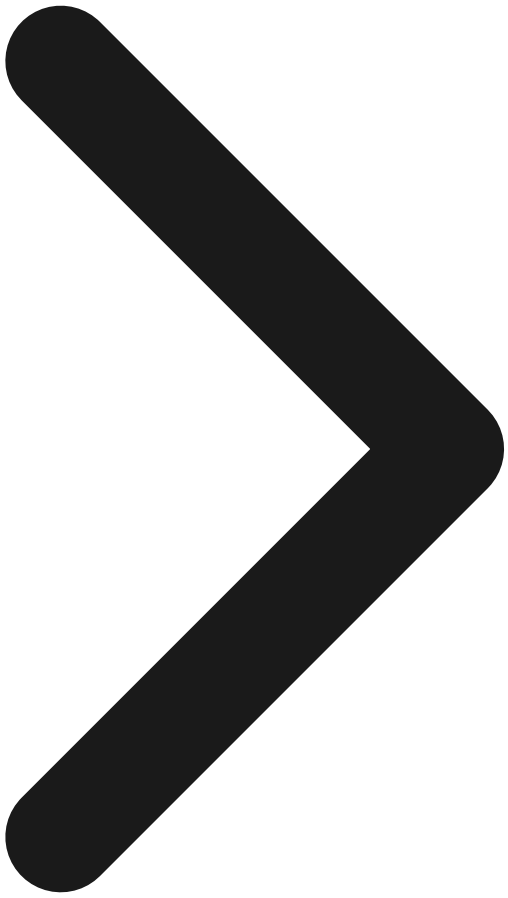
Framework7 min

Why Your Product Roadmap Looks Healthy but Users Keep Churning



Operating Model12 min

The Orchestrator Is a Delivery Manager



Can your operation see its own silent suffering?

The Silent Suffering Scorecard turns this essay into a four-minute instrument. Twelve questions across the four health-signal layers, and it places your operation on the five-rung ladder from flying blind to self-aware, with the one layer to build next.

Take the Silent Suffering Scorecard