

Evaluating the Quality of Customer Service: A Manager's Foundation



- Service quality = gap between expected and actual experience



- Evaluate before improving: you cannot fix what you haven't measured



- Avoid traps: single data points, efficiency vs. quality, sampling bias



- Course flow: gather evidence, benchmark, diagnose, act, repeat

The Five Dimensions of Service Quality

- Shared mental model:
SERVQUAL / RATER framework
- Reliability: deliver promised
service accurately and on time
- Responsiveness: readiness
to help promptly
- Assurance: competence,
courtesy, and trust-building
- Tangibles: visible signals
of professionalism



Choosing the Dimensions That Matter for Your Team

- Distinguish operational metrics (FCR, response time, resolution time) from human dimensions (tone, empathy)
- Separate quantitative evidence from qualitative insights
- Select 3–5 dimensions matched to your service type, not all metrics
- Weight dimensions by context: depth for complex support, speed for retail



Quality Dimensions	Retail	Healthcare	Financial
 First Contact Resolution (FCR)			
 Response Time			
 Resolution Time			
 Tone			
 Empathy			
 Clarity			
 Accuracy			

Reading Quantitative Signals



- CSAT, CES, and NPS measure distinct concepts: interaction satisfaction, effort, and loyalty



- Track operational metrics (FCR, response time, resolution time) alongside experience metrics



- Watch for score inflation, low response rates, and sampling bias

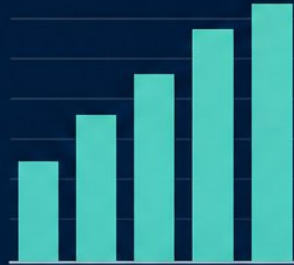


- Read trends over time; segment by channel, issue type, and agent

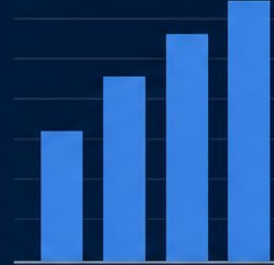


- Treat AHT as a capacity signal, never as a standalone performance target

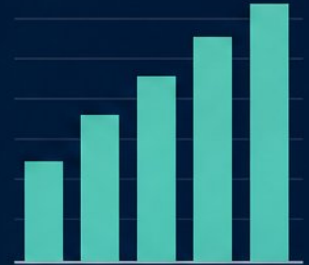
CSAT



CES



NPS



RESPONSE TIME TREND



Matching Metrics to Decisions

- CSAT: primary interaction-level metric, collected immediately after resolution
- CES: add when CSAT looks healthy but repeat contacts or refunds rise
- NPS: relationship-level, collect quarterly, not on daily scorecards
- FCR: strongest single metric; moves satisfaction and cost together
- Build a balanced scorecard of 3-4 core metrics, not 15



Qualitative Insights from Interaction Reviews



- Review tickets, chats, and calls against a written rubric



- Assess problem understanding, accuracy, tone, empathy, next steps



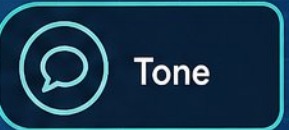
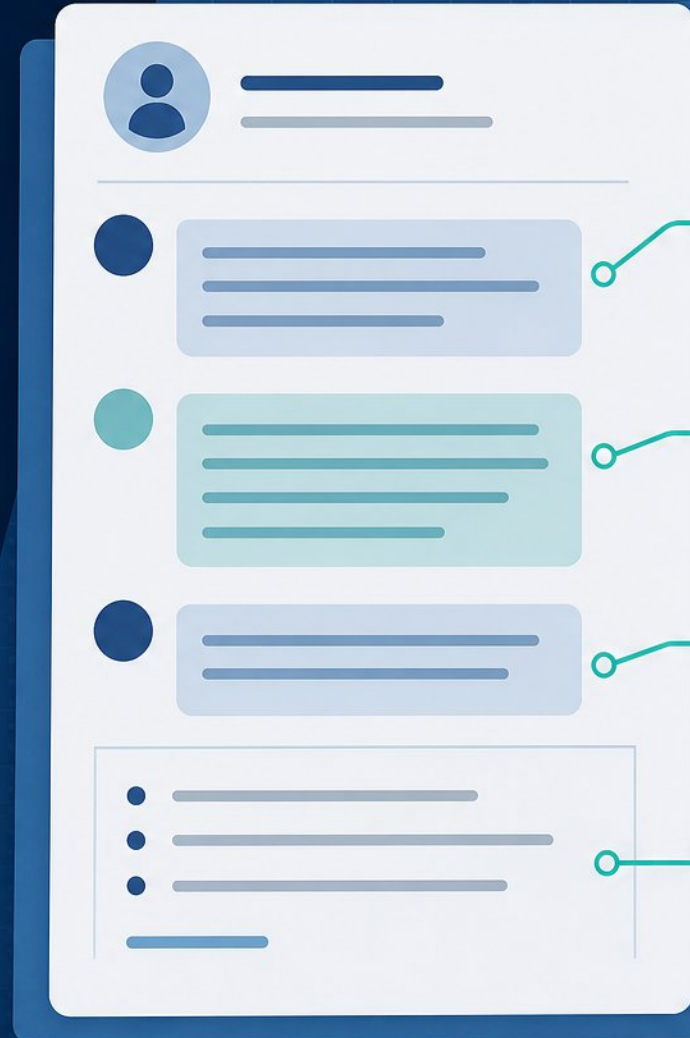
- Keep 8–12 weighted criteria with auto-fail compliance rules



- Calibrate reviewers so identical conversations score identically



- Target reopened tickets, low CSAT, and escalations alongside random samples



Running Calibration Sessions



- Reviewers score same 2-3 conversations blind



- Compare scores criterion by criterion



- Resolve contested criteria with one interpretation



- Measure agreement with Cohen's kappa



- Run sessions biweekly; monthly once stable



Listening to the Customer



- Keep surveys short: one rating, one open-ended follow-up



- Send immediately after resolution; match channel to interaction



- Cap survey frequency to avoid fatigue



- Route low scores to same-day or next-day follow-up



- Read unstructured feedback for recurring themes



- Pair every score with the open-text reason



Self-Assessment and Internal Observation



- Agents self-score against the same rubric to build ownership



- Peer review as a learning tool, not surveillance



- Calibrate peer reviewers more tightly than QA analysts



- Side-by-side coaching turns findings into actionable notes



- Publish the rubric in advance to reduce defensiveness



Benchmarking and Context



- Set internal baselines first; track month-over-month movement before chasing industry averages



- Use 2025–2026 ranges as context, not targets: CSAT averages 75–85% cross-industry



- Telecom and ISPs score structurally lower; healthcare and financial services higher



- Adjust expectations by channel, issue complexity, and customer segment



- Channel matters: live chat and phone score higher than email and social media

CSAT BENCHMARK BY CHANNEL



Current Internal Baseline



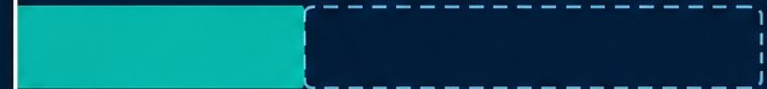
Industry Benchmark Range (2025–2026)
Context, Not Target



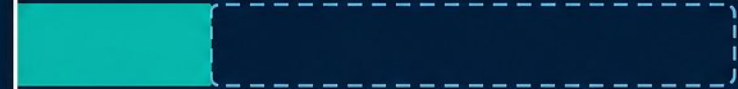
Live Chat



Phone



Email



Social Media



Use these ranges as context to interpret performance and set realistic expectations.

Diagnosing Root Causes



- Map low scores to customer journey stages, not isolated agent failures



- Apply Five Whys to recurring failures, grounding each answer in evidence



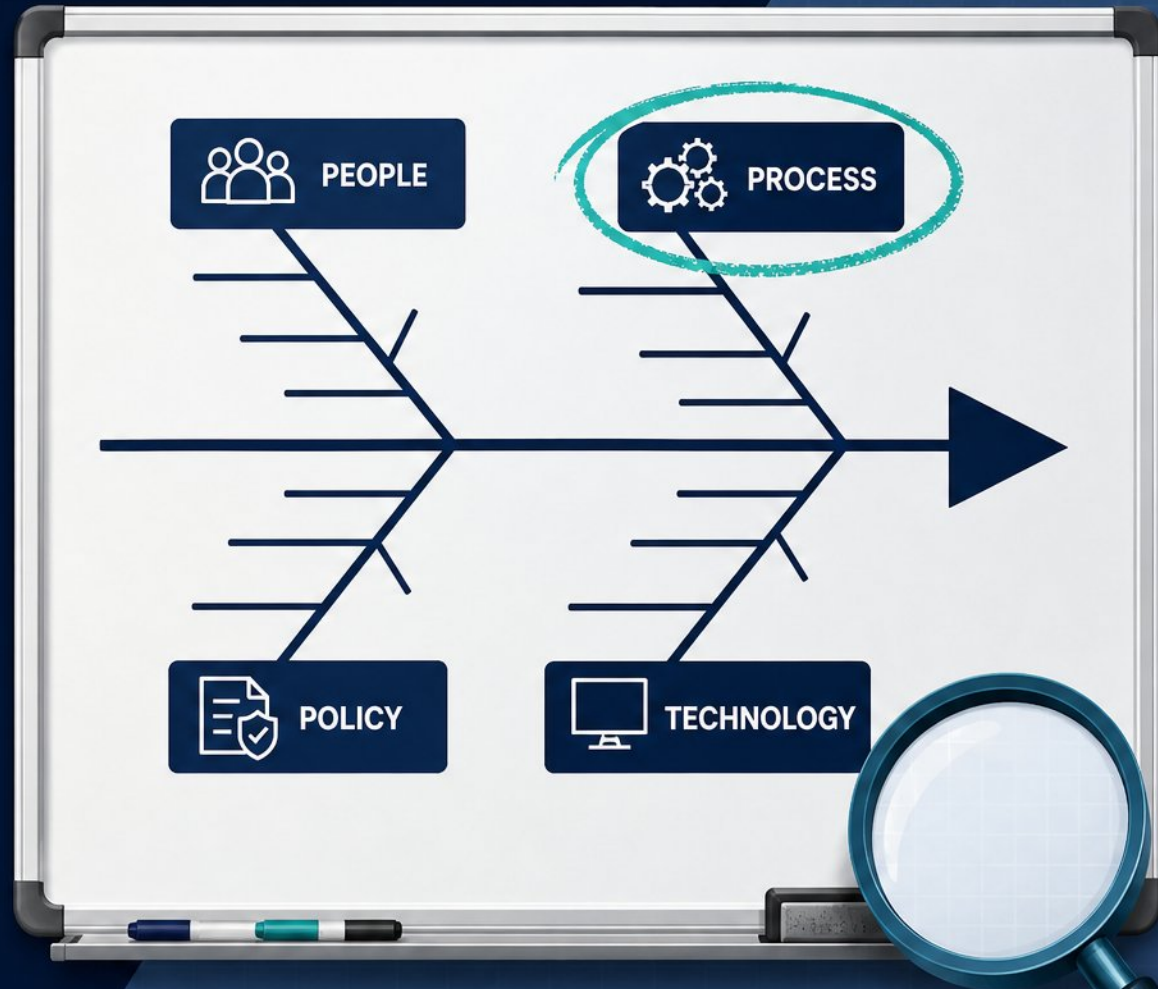
- Many agents failing the same criterion usually signals a system problem



- Use fishbone diagrams to map people, process, policy, and technology causes



- Prioritize fixes by customer impact and feasibility; fix the root cause, not the symptom



Building an Ongoing Evaluation Rhythm



- **Weekly:** review leading quality signals and trending criteria



- **Monthly:** assess outcome metrics (CSAT, NPS, resolution rate)



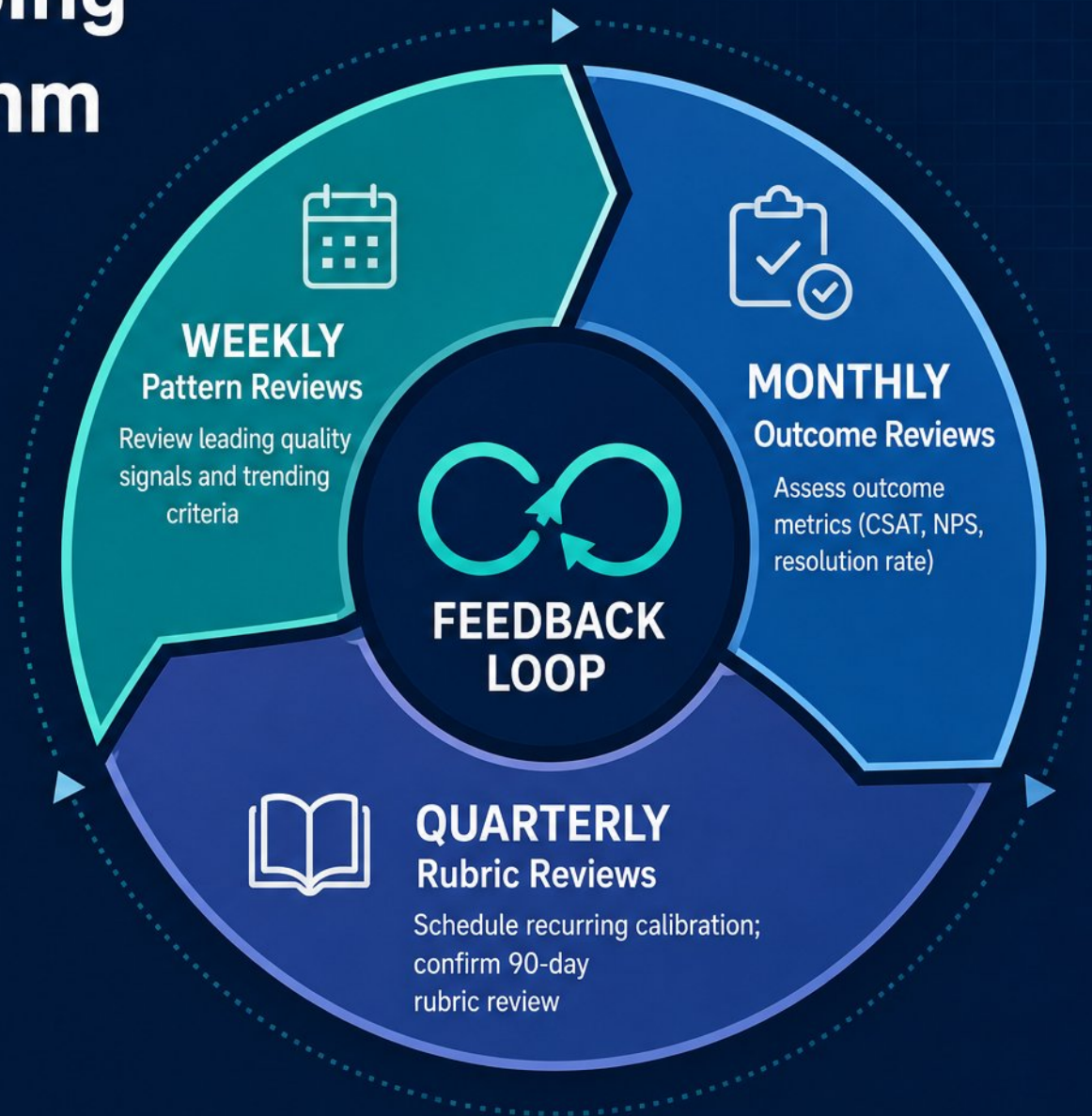
- Keep a lightweight scorecard, reviewed on a fixed schedule



- Schedule recurring calibration; confirm 90-day rubric review



- Balance experience, efficiency, and compliance in goals



From Evaluation to Action



- **Build a prioritized backlog:** one owner, one deadline, one success measure



- **Communicate constructively:** show trends, name one behavior, use real examples



- **Contain now, correct later:** protect the customer, then fix the process



- **Verify over 30–90 days:** confirm the change improved the targeted metric



- **Close the loop:** share wins and lessons to make evaluation visibly productive



Key Takeaways and Next Steps

01



- Evaluation precedes improvement: gather quantitative, qualitative, and observational evidence

02



- Use a small set of matched metrics and qualitative rubrics, not a sprawling dashboard

03



- Calibrate reviewers, benchmark against your own baseline, and diagnose root causes

04



- Build a recurring rhythm of review, coaching, calibration, and verification

05



- This week: define 3–5 quality dimensions and review 20–30 conversations to draft your rubric

THREE SOURCES OF EVIDENCE



Quantitative Evidence



Qualitative Evidence



Observational Evidence

THE EVALUATION LOOP



FIVE QUALITY DIMENSIONS



Accuracy



Empathy



Clarity



Resolution



Professionalism

PersonWise.

PersonWise • AI digital-human course platform



Scan the code or click anywhere to watch this course live, with voice Q&A

personwise.ai/t/61594caf/public