The Ultimate Guide
to CX QA Scorecards

A MaestroQA guide

Table of Contents

Foreword

Consumers expect more from support teams than ever before.

Research from American Express and Harris Interactive shows that customers tell *fifteen other people about* bad support experiences, and that *89% of consumers* have switched brands because of bad support.

But the way brands measure support is outdated. Metrics like AHT and FCR measure efficiency, but a speedy interaction isn't always a good one.

Similarly, CSAT can capture sentiment, but it's static. None of these metrics give teams actionable insights into how to improve support operations or the quality of interactions.

Teams are now turning to QA scores to get the information they need—putting their QA scorecards front-and-center.

With that in mind, we put together this guide to any questions you might have had about QA Scorecards.

In this guide, we draw on our experience helping *hundreds of support teams* build quality programs—you'll find guidance on building your first QA scorecard, how to know when it's time to modify an existing scorecard, and tips and tricks from our customers who have built successful QA programs from the ground up.

Let's get started →

Here's Why You Need QA:

Agents are expected to do more than ever.

Apart from juggling multiple channels of support, agents have to contend with knowledge bases and internal communication platforms, all while delivering a consistent customer experience.

It's hard for agents and managers alike to keep up without a clear way to judge performance and share learnings. QA software helps your agents track their performance across these different channels, while identifying opportunities for improvement and growth.

Management is expected to drive insights, not just performance.

While the agent's role has gotten more complex, CX managers have seen their roles evolve as well.

CX management used to concern themselves mostly with driving agent performance, but today's CX leader wears many hats—including seeking out insights about how to further optimize the processes that help their CX program run.

Additionally, CX teams are increasingly being seen as *the* Voice of the Customer, driving organic customer feedback directly to their product and marketing teams.

Consistent QA at scale helps deliver these insights to teams.

You can't automate the human touch.

It's simple: people build trust with people. In an era where people are buying more online than ever, an agent is often the only human touchpoint a customer has with a brand. And while chatbots and online self-help has grown in popularity recently, nothing beats a genuine conversation with another human being.

CX leaders have also found that agents thrive when given the agency to make decisions about how to deliver the best experience for the customer, instead of just following a script. As Etsy has discovered, the best way to drive agent performance is through direct manager coaching, not an AI-generated score.

QA allows teams to scale up their audit volumes, generating the process insights needed to make improvements, like monday.com, or market insights that result in massive cost savings, like ClassPass.

QA empowers CX teams at 3 different levels

Product and Market Insights

Identify market trends and product insights that build customer loyalty

classpass

Identified Market Insights Leading to 30% Higher Customer Retention

Process Insights

QA data illuminates gaps in your CX processes, allowing you to streamline processes and provide a better customer experience.

monday.com

Identified Process Improvements That Reduced AHT by 30%

Agent Performance

Understand how agents perform on customer interactions and identify areas for individual improvement.

Etsy

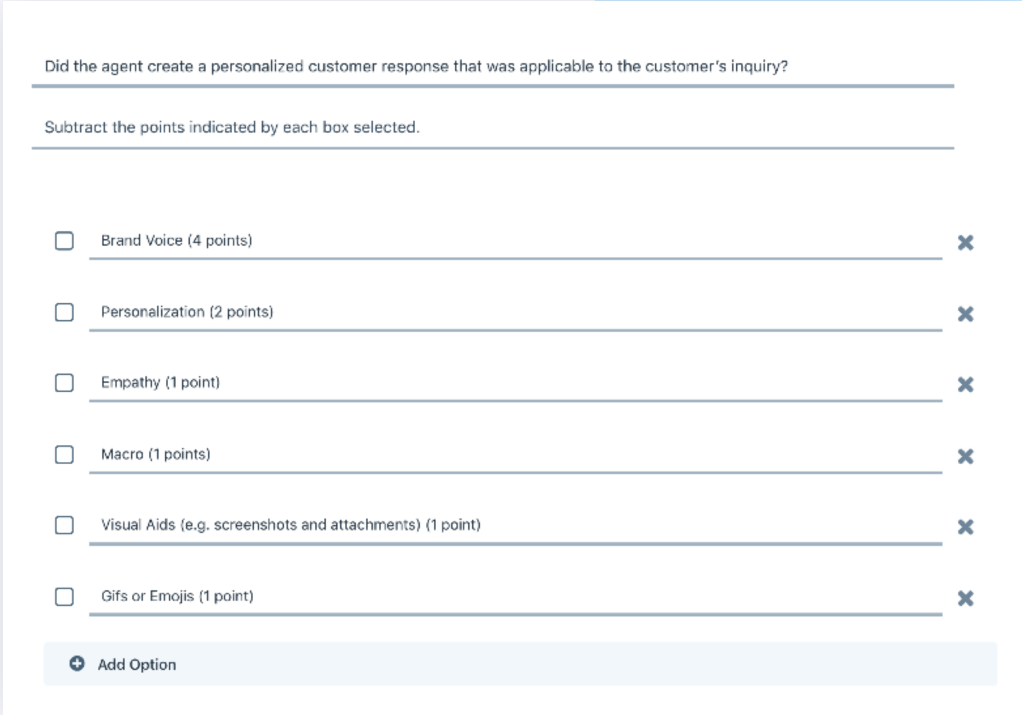
Improved Agent Performance by 14%

What is a QA Scorecard?

A quality assurance scorecard is a collection of criteria against which an agent's customer interactions are graded to ensure they are in alignment with their brand's standards, rules and regulations, and more.

Most quality assurance scorecards are a collection of customer service values and procedures that are presented as questions to the grader.

Here's an example of what a QA scorecard section looks like:



Real-World Example

MeUndies

Company: MeUndies, makers of comfortable underwear, loungewear, and apparel

Industry: Clothing & Apparel

MeUndies asks its graders to assess how well agents personalize their responses to customers—something that sets them apart from competitors on the CX front.

Tip: Linear point scales, yes/no dropdowns, and short answer fields are also commonly used in QA scorecards. These help your graders move faster and provide consistent feedback.

How to Build Your QA Scorecard

There are **four key steps** to build out a QA scorecard that works for your unique business and brand needs.

Step 1: Identify the Values and Goals that Drive Your Customer Interactions

Step 2: Create and Categorize Your Scorecard Questions

Step 3: Pick A QA Scorecard Tool

Step 4: Test Your Scorecard

1. Identify the Values and Goals that Drive Your Customer Interactions

Your customer-facing team is the most frequent human point of contact that a customer has with your brand—it's essential that they embody your brand's values.

To begin, grab a piece of paper or whiteboard and write "Our Values" on the left side. Fill it up with the values that drive your customer interactions. A good place to start would be your company and brand values, as well as any extra values that are specific to the support team.

On the right side, write: "Our Operational Goals."

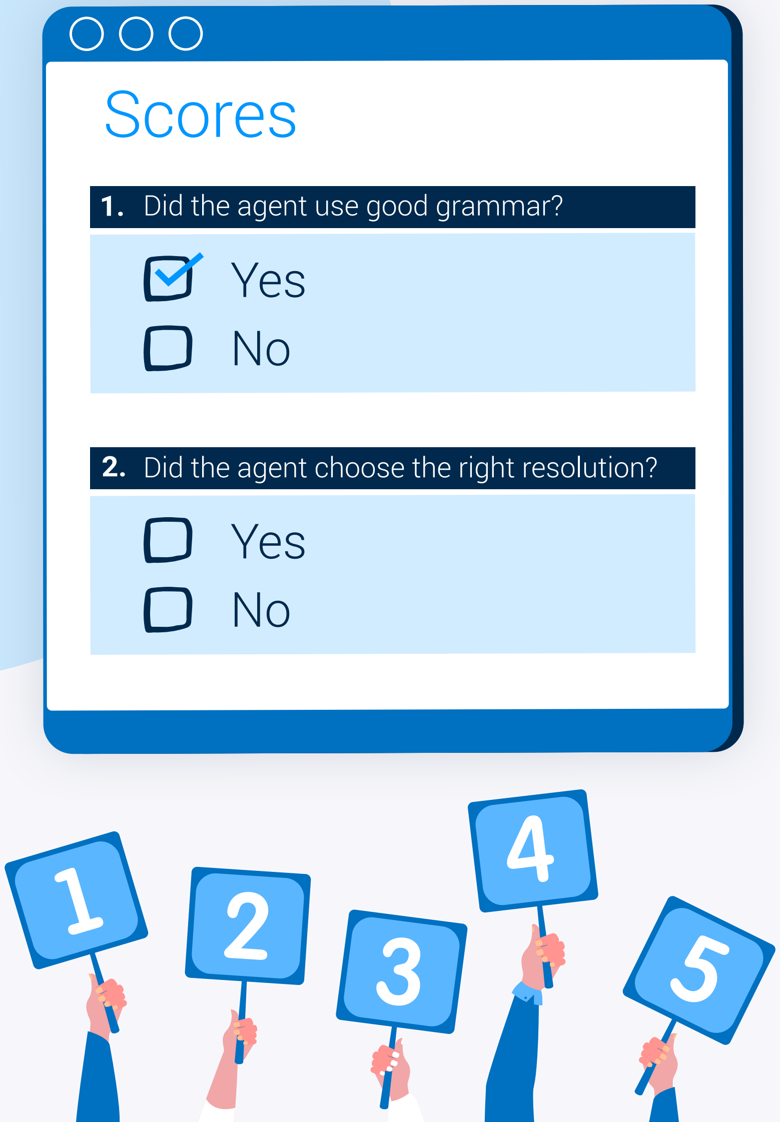
For example: Is ensuring that customers have their issues resolved on the first call important to your team's CX philosophy? Add "FCR" to the list. Are you aiming for your agents to deliver a certain number of interactions per hour? Pencil it in.

To round it off, spend time listening in on customer calls and speaking with agents. These interactions will help you to identify the pain points customers and agents face as well as potential opportunities to improve.

Our Values	Our Operational Goals
effortless experiences	aht < 10 mins
customer delight	85% customer retention rate

Tip: Try to find common themes across your Goals and Values—these will be key to formulating your scorecard questions in the next step.

2. Create and Categorize Your Scorecard Questions



Most scorecard criteria fall into one of four categories—Communication, Customer Connection, Compliance, and Clear and Complete Content.

When creating your scorecard, organize around the Four Cs to make it easier for graders to grade, and for agents to review their results.

Here are some examples of common scorecard questions you'll find in each of the Four Cs.

Communication Skills

Common examples: Did the agent speak at an easy-to-understand pace, and a friendly tone of voice?

For phone conversations, you might want to evaluate tone of voice, pace of communication (e.g. talking too fast or slow), or the excessive use of filler words (e.g. um, ah, and mmmmkay).

For text-based channels, grammar and spelling are critical for communicating clearly, concisely and in a way that's on brand.

Customer Connection

Common examples: Did the agent make a real connection with your customers? Did the interactions provide an experience that differentiates your brand from other companies?

Some common questions that make it into scorecards are:

- Greeting customers warmly and using their names wherever possible
- Listening carefully, acknowledging issues, and responding empathetically to the customer's mood and tone
- Communicating a willingness to help and taking ownership until the situation is resolved

Compliance and Security

Common examples: Did the agent follow all essential policies and procedures to keep the customer and the company safe? Did they handle customer information in the right way?

Security is a critical component of quality review. At a minimum, security means properly authenticating customers before disclosing or changing information on their accounts.

Depending on the industry, your team might also be required to follow protocols like PII, PCI, and HIPAA. Failure to comply with these regulations reduces trust with customers and could result in significant legal problems for your company, so it's wise to make sure agents are regularly graded on compliance and security.

Correct and Complete Content

Common examples: Did the agent provide correct and complete answers and use the right tools to arrive at those answers? Were all internal processes followed?

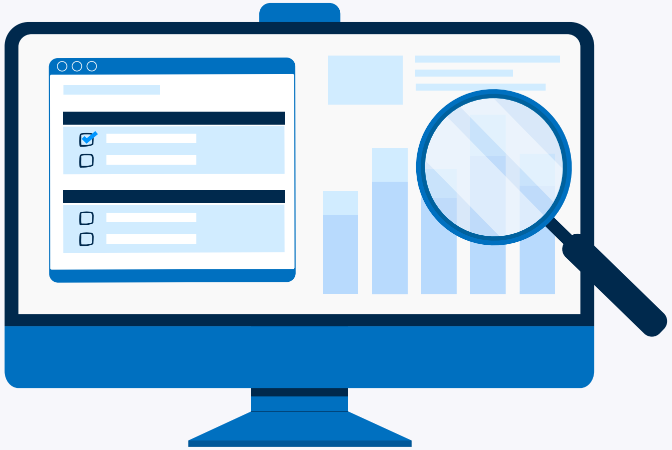
Providing wrong or incomplete answers defeats the purpose of having a customer service team. After all, customers reach out to gain access to information that will solve their specific challenges. When customers receive bad information, they'll either lose confidence in your team or call back again in the future—leading to unnecessary spikes in call volume.

3. Pick a QA Scorecard Tool

There are two basic tools you can use to create a QA scorecard:

Spreadsheets

VS.



QA software

Spreadsheets are often the starting point for smaller teams or teams building out their first QA program. While they may work in the short term, they don't scale with teams as they grow.

We built MaestroQA to help CX teams go beyond spreadsheets. In fact, many of our customers who switched from spreadsheets have experienced significant increases in efficiency and CSAT scores. Here's Clever's story:

Real-World Example

Clever

Company: Clever

Industry: Education

Clever's QA program was built on Google Sheets and Google Docs, which created a variety of administrative bottlenecks. "Between matching the Google Sheets and Google Docs, servicing that to managers, sharing it with individuals, it was just such a mess," said Natalie Mercuri, QA + Learning and Development Lead at Clever.

Grading five tickets took at least ninety minutes of work—not including the feedback sessions. Switching to MaestroQA provided one platform for all things QA and accelerated the grading process.

Grading five tickets with MaestroQA only takes 30 minutes—a 67% improvement compared to spreadsheets.

"Having a tool power your QA program will pay for itself tenfold within the first week"

Mercuri said.

4. Test your Scorecard

Your QA scorecard needs to work for all of your graders, managers, and agents. That's why properly testing your QA scorecard is a key step for ensuring a successful rollout. Consider these best practices when testing:

Assemble a Small User Group

Form a team that will test the new scorecard. When selecting team members, look for stakeholders who are familiar with your CX goals and willing to provide constructive feedback.

Then, agree on a methodology for using the new scorecard. If this is your company's first QA scorecard, graders could immediately begin using it to assess all tickets during the test period.

However, if you're replacing an existing QA scorecard, you will need to decide on how many tickets should be graded with the new rubric vs. the old rubric.

Remember to keep agents in the loop, too. Reassure them that any tickets graded with the new rubric are for testing purposes only and will not influence their KPIs.

Make It Easy for Team Members to Share Feedback

Create a dedicated Slack channel to collect real-time feedback from graders. Encourage graders to post anything about the new scorecard that seems unclear, time-consuming, or misaligned with your values and policies.

Calibration sessions—meetings that get all of your graders talking about the new scorecard—are another great way to solicit feedback and identify potential issues. To maximize the impact of each calibration session, use your new scorecard to grade one or more tickets prior to the meeting. Compare your grades to those of other graders and then use the meeting as a forum for overcoming misalignment.

Put Your Scorecard to the Test

Three to four weeks of focused testing should be an adequate amount of time. That said, wrapping up testing in the least amount of time is not the primary objective. What's more important is ensuring that graders understand each question and that your new scorecard is an effective communicator of feedback.

Tip: Workflow automations in MaestroQA can accelerate the testing process by simplifying the assignment and grading of tickets.

Using Your QA Scorecards

Creating your first QA scorecard is a major milestone—but there's more work to be done. Before your get started with using your QA scorecards to grade, here are a couple of things to consider:

Decision #1: How Many QA Specialists Do You Need?

Most teams hire dedicated QA specialists to run their QA programs and perform grading. This allows managers to focus on other key tasks, like hiring, coaching, and onboarding.

Dedicated QA specialists develop deep specializations in grading agents, interpreting the data produced, and providing trusted CX insights for improving CSAT, reducing churn, and maintaining a positive brand reputation.

And, since they aren't directly in the queue, QA specialists provide a fresh, third-party perspective on issues that inevitably pop up during grading.

Not all teams are large enough to need a QA specialist, and some don't have them. The alternative is to have managers, team leads, or senior agents double up to do QA.

It's a great way to train senior agents on a new skillset and provide opportunities for career progression. This is where team maturity comes into play—you need to have some established veterans on the team to help grade.

That said, team leads, managers, and senior agents are *busy*. Between onboarding, hiring, coaching, reporting, and answering tickets, setting aside time for QA might fall off their radar.

When this happens, agents are the ones who ultimately miss out on coaching and training opportunities.

So, exactly how many QA specialists do you need?

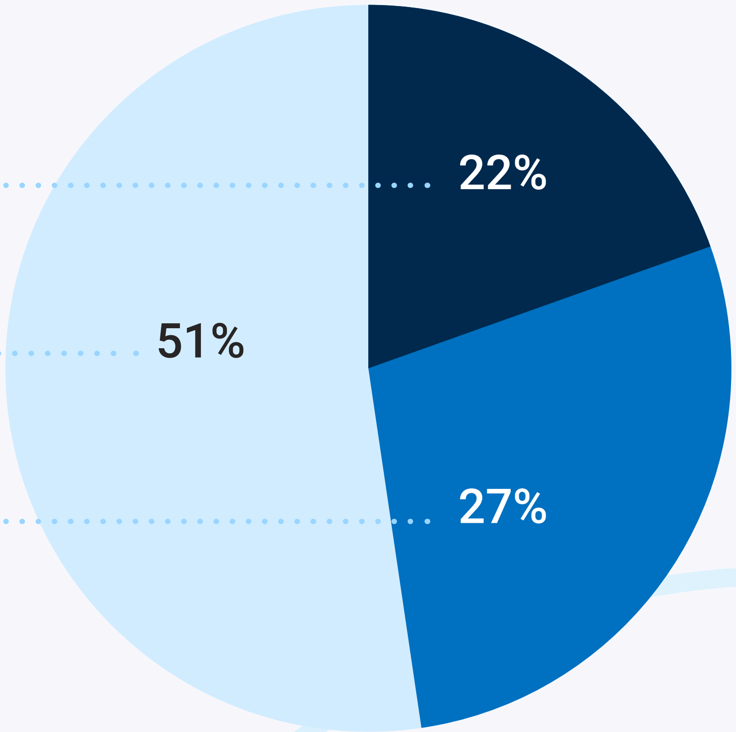
Finding the Right Mix

We asked our CX network about their grader-to-agent ratios and found that most teams have a ratio of roughly 1:20. Here's a more detailed breakdown:

22% of teams have a grader-to-agent ratio of 1:10

51% of teams have a grader-to-agent ratio of 1:20

27% of teams have a grader-to-agent ratio of 1:40 agents or more



In general, larger teams tend to have more agents per grader. As the team matures and builds a base of senior agents who require less grading, they shift their energy to newer agents.

Decision #2: How Often and How Much Will You Grade?

Next, you need to decide how many tickets need to be graded, and how often grading should happen to hit that goal.

60% of MaestroQA customers grade between 1% to 5% of all interactions. The other 40% grade at least 5% of all interactions, but these tend to be smaller companies with a dedicated QA specialist on the team.

There are 3 main factors that determine how frequently you should grade:

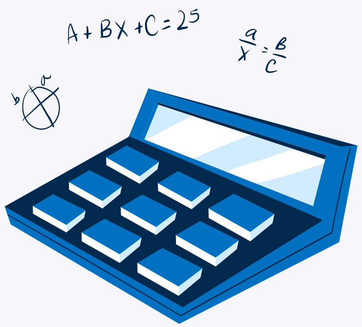
1. CX Team Resources

If you're relying on team leads to handle QA, you might want to lower your grading volume to accommodate their other responsibilities.

With dedicated QA specialists, it's easier to work out realistic QA goals. We like using a percentage of tickets to set a grading goal, because it works for any size team.

Example: let's imagine that your team consists of 10 agents and 1 grader. The team handles 5,000 tickets per week, and you'd like to grade 5% of tickets.

5% of 5,000 tickets = 250 graded tickets per week



Using simple math, your grader needs to grade 50 tickets per day to hit the goal.

We encourage an iterative approach to determining grading volume—try this goal for one week and see if your grader can reasonably achieve it. If they can handle more tickets without compromising grading quality, increase the goal. But if 50 proves to be too much, consider setting the bar a bit lower than 5%.

2. Average CX Ticket Difficulty

If the majority of tickets in the queue are easy to handle, choosing random tickets to grade might not provide the insights that you truly need. Put simply, if a majority of your tickets are "easy", the more "difficult" (error-prone) tickets that require more agent knowledge will be graded less often based on a random sample of tickets.

Here's how MaestroQA customers, Handy and WP Engine, have moved beyond random ticket sampling to overcome this issue.

Real-World Example

Company: Handy

Industry: Home Services

The team at Handy only grades DSAT tickets—tickets that either have negative CSAT scores or those flagged by agents because they struggled with an interaction and want a review. This allows the tickets with the most opportunity for learning to float to the top, leaving Handy with the best possible insights to improve their training and CX programs.

Real-World Example

Company: WP Engine

Industry: Web Hosting

WP Engine has a hybrid program. They filter every DSAT ticket out for grading but maintain a separate QA instance that randomly selects tickets to grade. This allows the company to reap the same benefits as Handy's program, while ensuring that the general performance of the CX program is maintained.

3. Agent Seniority

Most companies that track QA data come to realize one common truth: agent performance increases with experience. The main way to account for this is to set a minimum threshold score for grading.

- When you set the minimum threshold score (for example, 85/100), consider lowering the grading volume and frequency for the senior agents who have consistently maintained an 85+ QA score. This will allow you to focus more time and effort on newcomers, while building trust with senior agents. QA software should be able to help you group agents based on their seniority or even select tickets for grading based on their CSAT score.
- Taking it one step further: involve senior agents in setting the minimum threshold score. This approach gives senior agents ownership over part of their coaching program and provides new opportunities for growth.

Feedback & Reporting

To maximize the impact of your QA scorecard, you need a scalable way to collect, aggregate, and distribute meaningful feedback and insights to agents, managers, graders, and the rest of your CX team.

Relaying QA Results to the CX team

There are three main ways to relay your hard-earned QA scores to your team: email notifications, coaching sessions, and team meetings.

Email Notifications

Some QA platforms allow agents to receive their scores via email immediately after the grade has been submitted by the grader. This enables real-time feedback that the agent can apply to the queue, while staying engaged with the QA program.

Agents shouldn't be left to interpret their QA scores on their own, however. This method should be paired with regularly scheduled coaching sessions.

Coaching Sessions

Coaching sessions provide managers with an opportunity to analyze each agent's long-term results and offer qualitative feedback.

After all, the numbers don't care about an agent's feelings—but managers do. Coaches can help reframe QA results and put them in context. For example, a below-average score could be due to a recent product launch or a one-off event rather than a long-term trend.

Most teams aim for one coaching interaction per agent per week. This number can be decreased over time for more senior agents.

Team Meetings

Team meetings are used to dig deeper into the QA data, identify trends, and proactively address localized issues before they become widespread.

Mailchimp uses team meetings combined with a QA newsletter for this purpose—and they always experience a positive spike in QA scores for the highlighted section of the rubric.

Reporting on Scorecard Data

Aggregating your QA data into reports and dashboards makes it easier to evaluate the effectiveness of your QA scorecard from several perspectives:

Grader Efficiency

How many tickets can your graders grade per hour, day, or week? Are your graders becoming more or less productive over time? What factors are influencing the trends that you see? If you've made adjustments to your QA scorecard, how does grader efficiency compare now vs. before the change?

Quality of Insights

How reliable is the information that your QA scorecard generates? Is there a scorecard-related issue that is causing graders to misgrade certain types of interactions?

For example, let's say that a QA report surfaces that agents consistently score low marks on refund-related interactions. Are agents being graded too harshly for relaying the refund policy itself (which is outside of their control)?

If so, a scorecard adjustment, procedure adjustment, and/or additional training could be warranted to remedy the artificially low QA metric.

Scalability

Standardized reports can only tell part of the story. To gain meaningful QA insights at scale, you need the ability to slice and dice your QA data across a variety of agents, teams, helpdesk tags, and scorecard criteria.

That's why MaestroQA makes it easy to build custom reports that can be filtered by group, rubric, question, and other variables.

Updating Your Existing Scorecards

Because QA scorecards represent a snapshot of your company's values, CX processes, and communication style, they must be updated frequently to ensure they're still aligned with your ever-evolving brand and CX goals.

When to Re-evaluate Your Scorecard

At a minimum, you should re-evaluate your scorecards every 6 months to ensure constant customer service quality. That said, there are several key events that should prompt an automatic re-evaluation.

1. Updated Brand Values

QA scorecards represent your company's values. When your values change, your scorecards should change, too.

For example, the communication section of your scorecard may regulate the use of certain words, such as referring to customers as "members," or check for a friendly tone of voice. These practices may change as your brand matures and evolves. As such, a brand refresh or the launch of a revised brand values handbook is a good opportunity to re-evaluate your scorecard and confirm alignment.

2. A Push for Increased Efficiency

It could be a new C-suite hire or a company-wide call to do more with the same resources. Whatever the trigger, a push for increased efficiency is something that most teams go through at some point.

As the backbone of your QA program, your scorecards are a good starting point for increasing your grader and agent efficiency.

3. Fewer CX Insights Observed

Trustworthy CX insights are the primary goal of any strategic QA program. When these insights start to dry up or are less impactful, it's time to re-evaluate your QA scorecard.

The "insightfulness" of a QA program is not easily measured, but experienced CX managers know when their QA programs are not delivering the same level of insights and when it's time to switch things up.

4. Signals from QA Data Indicate Scorecard Staleness

Your QA data can also provide powerful signals for determining when to re-evaluate your scorecards.

A few indicators to keep your eye on:

- **Stagnant QA Scores:** Handy's CX team also noticed that their agents consistently scored in the high 90s. While some might see this as cause for celebration, they decided to dig deeper into their scorecards and settled on grading only tickets with negative CSAT scores.

  If you're a team that's constantly scoring high and wondering how to gain more QA insights, this approach might be worth exploring.

  But, if your team's average QA score is in the mid-70s and showing no signs of improvement, shifting the goalposts is less likely to be beneficial.

- **Increase in Agent Appeals:** When an increasing number of agents appeal their grades, it can mean that your scorecard is misaligned with agent training and procedures.

- **Graders Missing their Grading Targets:** If graders are missing their grading targets, there might be an efficiency issue with your scorecards. They might be structured in a way that's difficult to interpret or have too many questions, which leads to long grading times.

Real-World Example

Company: Handy

Industry: On-demand Home Services

Handy's CX leadership team realized that their QA program was no longer delivering insights that the team could use to improve their program. Further investigation revealed that the sheer volume of simple, easy-to-handle tickets was crowding more complex, error-prone tickets out of the grading queue, causing Handy to miss out on opportunities for improvement.

Handy's CX team decided to only grade tickets with negative customer satisfaction scores. As a result, they were able to pinpoint opportunities for improvement on complex tickets and drive performance improvements.

Tip: Revamping your scorecard based on stagnant QA scores is essentially shifting the goalposts. New data can't be directly compared to the historical, pre-shift data.

Better Scorecards = Better Outcomes

Your scorecards can make or break your QA program—there's no better time to build your first one or iterate on one that already exists.

While there's a lot to consider in the process, we boiled down into three handy checklists to help you get started.

Checklist: How to Build My First Scorecard

- **Identify the Values that Drive Your Customer Interactions**

  Make a list of both brand values and operational goals that your team needs to hit.

- **Create and Categorize Your Scorecard Questions**

  Draft questions and categorize them to make sure graders don't have to jump around too much when grading.

- **Pick A Scorecard Tool**

  Decide where you're going to host your scorecard for grading—a tool like MaestroQA can help!

- **Test Your Scorecard**

  Test your scorecard with a small group to ensure it helps you achieve your goals.

Checklist: How to Redo My Scorecard

- **Get Insights from Stakeholders**

  Consult leadership, graders, and agents to ensure you're aligned on the insights needed from scorecards and make note of existing issues.

- **Look for Ways to Simplify Your QA Scorecard**

  Depending on the issue you're trying to fix, you may need to eliminate or reformat questions to provide more structured data, or to make your grading process more efficient.

- **Re-examine Company Values and Policies**

  Make sure your scorecard aligns with company values, policies, and brand identity.

- **Ensure that Your Scorecard + QA data is easy to interpret**

  Check that your scorecard's output is clear and easy to understand.

- **A/B Test Your Scorecards**

  Test your changes against your old scorecards, making one change at a time, to determine the impact of these changes.

How to Manage a QA Scorecard Revamp?

Phase 1: Collect Insights from Stakeholders

Survey agents, graders, and managers. Ask open-ended questions about the ease of use, the insights delivered, and for any other feedback they might have.

Phase 2: Re-examine Your CX Goals and Policies

Have these changed recently? Remember to update your scorecard to match any updates to CX policy or goals.

Phase 3: Build Your New Scorecard and Collect Feedback

With these insights in hand, it's time to build your new scorecard! Remember, your MaestroQA Customer Success Managers are always on hand to help you strategize and lay out your scorecard.

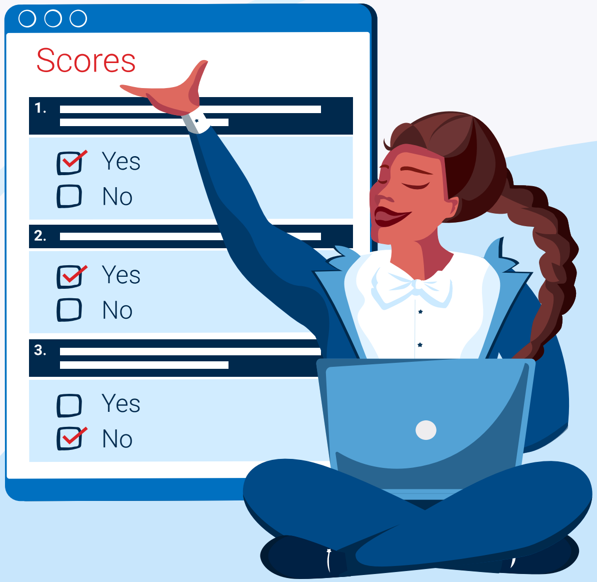
Phase 4: Test the New Scorecard

Do a soft launch. Have graders grade tickets on both the old and new scorecards, and collect feedback from them over two weeks. This allows you to compare scores across your old and new scorecards, to ensure you're seeing the results you expect.

Also remember to introduce the new scorecard to agents, and use your LMS or Knowledge Base to ensure that information about the new scorecard is easily available to them.

Phase 5: Launch!

Consolidate feedback from the previous phase and apply it to your scorecards. You're now ready for launch! Send out an email announcing the new scorecard, and reward agents who perform well with it. Our favorite? A reward for the first agent to hit a 100% score on the new scorecard.



A Sample Timeline for Your Next Scorecard Revamp

Typical scorecard revamps take about 6 weeks from start to finish. Of course, not all of this is active time on your part. Most of the time in this sample timeline is set aside for you to collect data on your old and new scorecard.

CX/QA Managers		CX Agents
Re-examine CX goals and policies	**Week 1**	Participate in surveys, interviews about current state of QA
Analyze results, build new QA scorecard	**Week 2**	
Test QA scorecard amongst graders	**Week 3**	
Soft launch: test QA scorecard amongst graders	**Week 4**	Soft launch: test QA scorecard and provide feedback to management
Edit scorecard with feedback	**Week 5**	

Week 6 Launch!

MaestroQA

Need more scorecard help? We're always here for you.

MaestroQA makes omnichannel quality assurance software for modern support teams. Etsy, Mailchimp, Peloton, Zendesk, and more use MaestroQA to improve agent performance, optimize CX processes, unlock business-level insights, and enable amazing customer experiences—all while improving the metrics that matter like retention, revenue, and CSAT.