

# QA Program Workshop

Building and optimizing a successful QA  
program for customer support and call centers



# Group Presentations Summary



**Group 1**

**CX**   
**SUMMIT**  
NYC | 2023

# Agent Performance

**Luxury Resale Customer Win:** Made QA part of their 5 agent KPIs weighted at 20%. Positioned as the white night because so much of CSAT was not support related

**Financial Company Shared their Challenge:** No longer has it as a KPI given hyper focus on CSAT

# BPO Management

**Win:** *Luxury Fashion Retailer* – built a foundational in-house program over the past year. Before – was fully owned by the BPO with very low trust in results

**Their Challenge:** is consistently misaligned with BPO, and so there is still trust issues with the results

**Challenge of an Industry-Leading Optical Company:** BPO is self-reporting their QA, always asking: “where is this data coming from?”

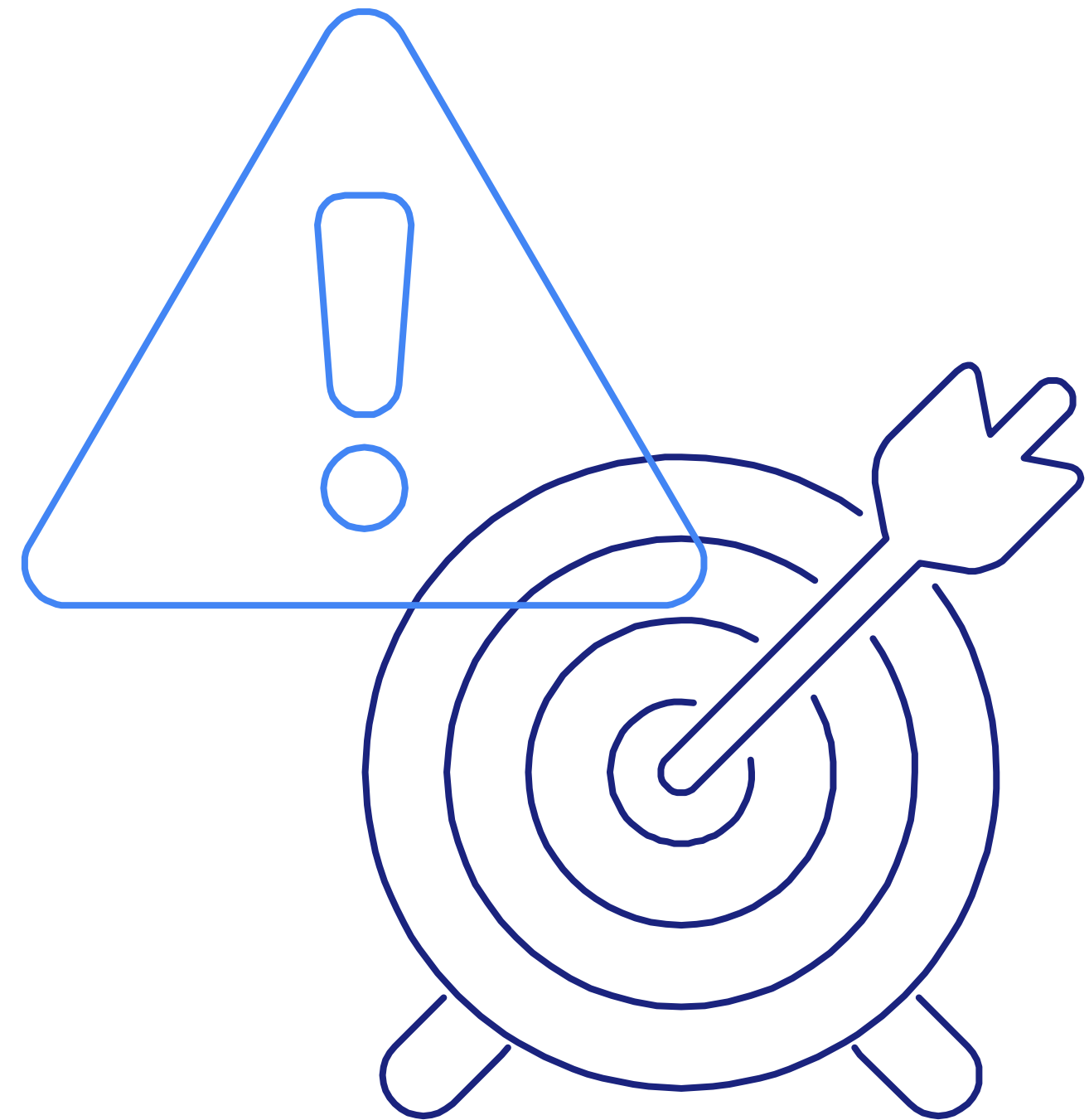
# Root Cause Analysis

**Financial Company Win:** Cares highly about CSAT, so QA is Targeted to reviewing DSAT. This allows the QA team to produce CSAT insights important for the BPO, internal team, and senior leadership

**A Pioneer in TV Steaming Win:** Wanted to reduce Cost per Contact and AHT was a contributor to cost per contact. ID'd through QA that a long Verification process was a primary driver in AHT. Roku simplified this process, and reduced AHT

# Takeaway

Large variance across our group on which of the Jobs were most critical and impactful



# AI Research

*Lightly researching AI but so far, but success in the below use cases*

## **Luxury Retailer:**

- Conversation summarization for QA and Escalations – taking an hour long contact into a few mins
- Call prediction based on customer history

## **Luxury Resale:**

- “Suggested Replies” for agents



# Group 2



# What Jobs Are We Focused on Solving?

## Improve agent performance / managing vendors

- Financial Company: **How do we demonstrate value?**
  - Do directors even know what these QA metrics are?
    - Maestro data is isolated from Team Leads using it to evaluate
- Luxury Retailer, Work Marketplace, Home Pro Services Provider: **QA data wasn't telling us enough**
  - Scores too high!
  - QA leads to having a sort of accountability for evaluating Performance (Connected Workplace Company)
    - Were getting metrics from BPOs that could not be fully trusting because there wasn't a way to confirm or deny
  - Rubric too long, so feedback was diluted (Luxury Retailer)
- **Need to tie QA outputs into coaching to truly improve performance**

# How Are Teams Conducting Targeted QA?

**Trade off: fairness of random QA vs. specifically digging into problem issues.**

- Focus on escalations (Home Pro Services Provider)
  - Why do things get escalated to leadership?
    - STOP selling Murphy beds! 😊
  - Do we need a separate team to focus on targeted QA?
- Focus on DSAT analysis (Work Marketplace Company)
  - Challenge is that when there's so much DSAT, it effectively becomes random
  - In a sea of DSAT, if a team is maintaining a high QA score, does the rubric need to be reshaped to focus?
    - Drive rubric criteria to focus on pain points

# How Are Companies Thinking About AI?

- Work Marketplace Company → AI and automation has been a part of our core
  - **Using AI to increase efficiencies**
    - First starting in allowing deflection of customer interactions reduce number of tickets that reach agents
    - Do you QA the chatbot interaction?
    - Can we use generative AI to have no distinction between the chatbot and the agent?
- Home Pro Services Provider → **Use AI to identify cases where you can upsell or cross-sell**
  - If high FCR... then what is the point of support agents?
  - Depends on industry, Luxury Retail, is supposed to provide a high end experience in store... but does this extend to online support?
- Luxury Retailer: Using AI to help **summarize support interactions**
  - Use AI generated summarize to efficiently make post-call notes or make handoffs between agents easier
  - Can AI be used to help assist an agent to better identify events that are happening in real time

# Group 3



# Why Do We Do QA?

- Fitness Service Provider : Agent Performance, RCA, Process Analysis
  - “Info is only as good as what we do with it” – Amy
- “Take the customer’s experience first, and let that guide QA” – Alison
- “Make our customers want to come back and talk to us again” – Alison
- “Part of IntelyCare is the quality of our people” – Brandon
- Evaluate the quality of trainings
- Find coachable moments
- Align on standards across CX

# Where Did Your Program Start?

- Punitive — Alison
- Punitive + Manual — Brandon
- Random — Sandra
- Under-resourced — Amy
- Technical + Meticulous — Valeri

# QA Wins 😊

- **Fitness Service Provider:** Unified, structured process for training, promotions, and QA
- **Shipping Software Company:** Immediate CSAT + Employee Satisfaction jumps after creating a culture of ownership at
- **Mental Health Software Services:** Remote work can feel isolating. QA makes people feel supported and empowered



# Challenges

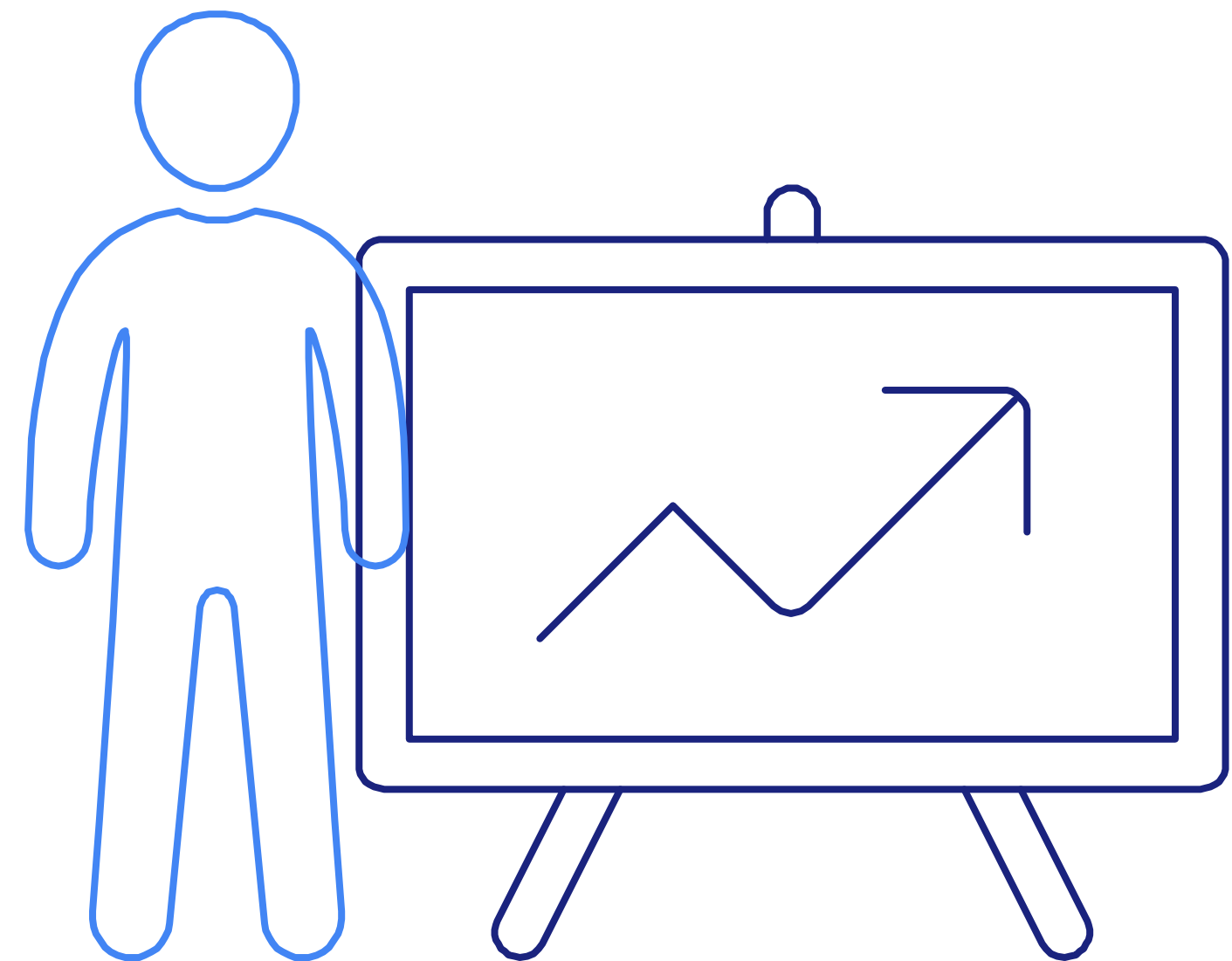
- Correlate with CSAT
- Limited sample size
  - Statistically Significant
  - Sampling the right things?
  - Give feedback to the people who need it most
- Limited resources to act on data
- Setting realistic expectations for the rest of the CX org
  - Need to plan cross-functionally

# AI Research

- Mental Health Software Services: Starting with Chatbots
- All: Conversational analytics to point to pain points, use QA to dive deeper
- Use AI to help, not replace humans
- Deploy resources in a more targeted way

# Top Takeaway

- Our programs were focused across all 3 jobs to be done — really focused on agent performance and managing BPOs, and target based on OKRs or CX initiatives
- We want to use AI to deploy humans in the most valuable ways
  - Chatbots can deflect volume so agents work on most complex issues
  - Convo analytics point to problems so QA can dig deeper



**Group 4**



# Group 4 — QA

## **Jobs to be done**

- Win: BPO — validation of BPO scores
- Challenge: Coaching — closing the loop (and follow up)

## **QA impact**

- Using targeted QA to measure success of preparation for launches and uncover opportunities for product and marketing teams.

## **Targeted vs. random strategy**

- Over 70% random for us all

## **Use of AI**

- Contact deflection & reduced amount of back and forth

# Data, Metrics & Everything Else

## Current Metrics

- QA: Highly reliable / not looking at 100% of tickets
- CSAT CSAT CSAT!: True reflection of client experience / responses are either really high or really low
- Productivity (AHT, One Touch Resolution, FCR): Related to proficiency / unbalanced and don't always reflect the full customer experience
- % of AI Handled vs. Agent Handled: Attracts cost efficiency / must be paired with quality metric to reflect full experience

## Exploratory Metrics

- Triage Rate — easily measure how AI has performed when routing / can lead to bad CX and unneeded multi-touches
- Agent vs. Bot CSAT — easily compare how chatbot performs vs. agents / difficult to train chatbots
- Number of Public Responses — clear that more responses = DSAT / sometimes unavoidable given situation

**Data** — authoritarian approach, too many cooks in the kitchen & need to set clear expectations

**Group 5**



# TLDR

## Jobs to be done:

- Most groups want to leverage more Targeted QA

## QA impact

- Leadership wasn't aware of how much QA is going into this
  - It's not just "turn on Chat GPT"
  - Don't know what scorecards entail or the "how" to impact metrics business cares about
- Sometimes driving toward "Business level metrics" is not in alignment with agent level experiences
- Need to blend metrics leadership cares about (NPS, CSAT, AHT) with metrics QA teams or CX management needs to coach to (FRT, hold times, sentiment, tone, empathy) in order to impact NPS, CSAT etc.

## AI

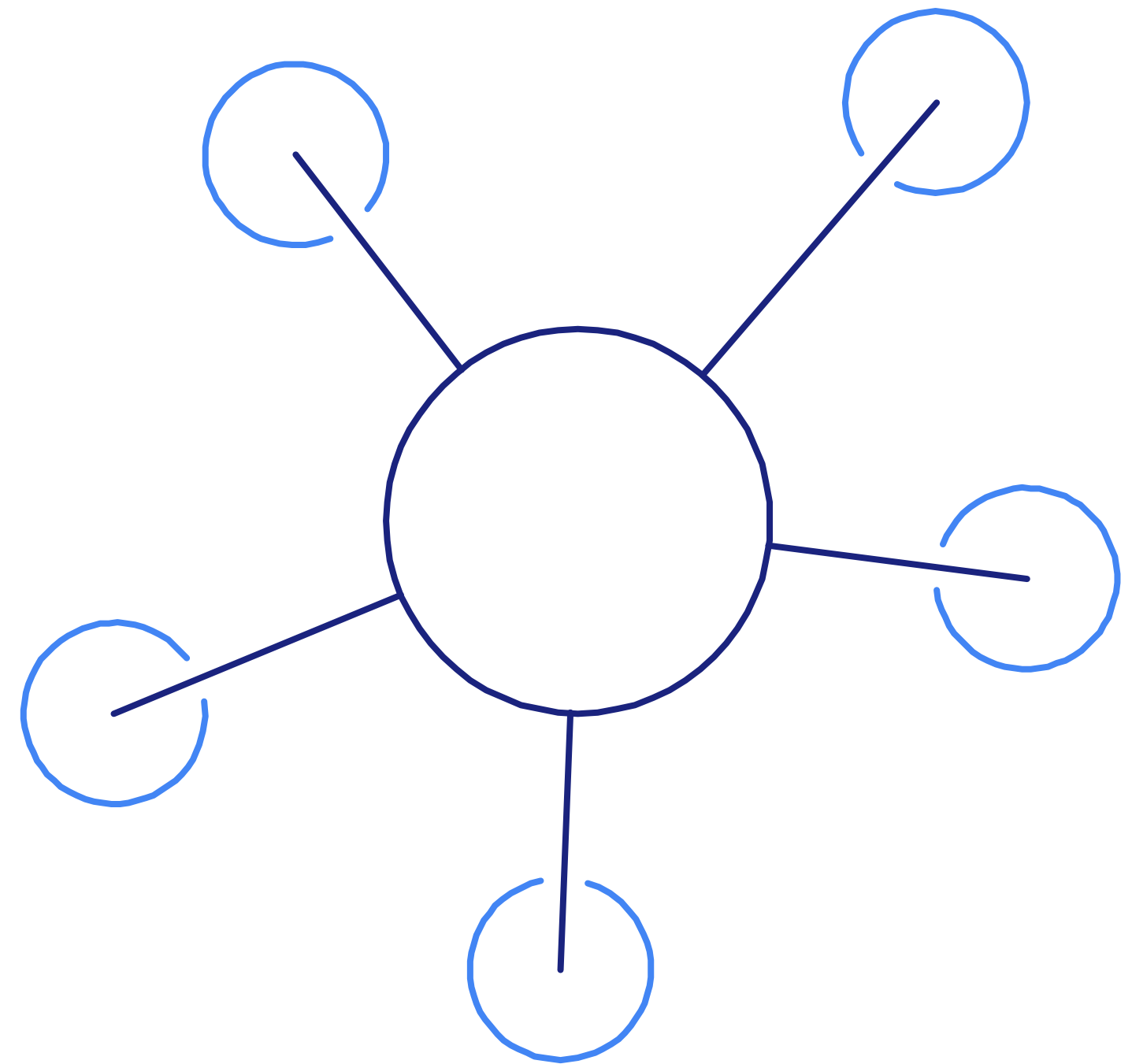
- Seeing a high use of AI in real-time assist/chatbots
  - QA on chatbots
- Want future AI to be directional, show me where to QA
  - Overall we have a low degree of confidence in AI, so QA will still require human intervention



# Opportunities

## Targeted QA

- How to implement?
  - Specific scorecards
  - Sprints
  - Impact to QA score/culture of QA score
- Use AI to help surface tickets to target



# Learnings

## What Metrics To Look At?

- FRT > AHT
  - AHT is not indicative of great service
  - Tickets closed prematurely is worse than open longer
  - Found customers care they are acknowledged and provided updates vs. issued resolved right away
  - Done right = longer handle time
- CSAT – QA correlation
- Leadership vs. Management / QA team
- Inflated QA scores
  - Leadership wants high QA score, but QA and CX know high scores are inflated

# AI

- Highest use of AI today
  - Real time – chat bot / chat containment and routing (lang.ai)
    - QAing chatbot
- Future use of AI
  - Want AI to help surface where we should look/QA
  - Look at 100% of tickets, autoQA
  - *“bad AI is worse than no AI”*
- Don't expect AI to fully replace QA, but compliment it – **AI should be directional**



# Group 6



# Part 1

## Top Metrics – General Themes

### Pro:

### Con:

CSAT

Trends for bad behavior (ie: agent consistently has low CSAT)

Subjective and does not tell the whole story

SLA

Great for the company to hold the vendors or internal agents accountable for response times

This metric only tells how quick the response to the customer is and does not result in resolution. Lead to repeat.

AHT

Understand which agents/BPO's can handle more volume for positive cost analysis

Fight between leadership to see reduced times vs. CX desire for agents to personalize conversations and spend more time making sure all issues are resolved

# Part 2

## First contact resolution

- Pros:
  - Insight into repeat calls
- Cons:
  - You tell me?

## Data experience

- General theme is that multiple program, tools, stakeholders lead to data being all over the place

## How are we connecting data?

- “People are reporting on the weather vs. providing true insights”
  - Because there is no central access to the data

## How are you democratizing metrics to your CX org?

- “Business is making the decision while the CX team is making the mess”
  - Involving them in the conversations earlier

# Jobs To Be Done — Wins

## Agent Performance General Themes

- **Problem statement:**
  - Agents handling the same type of interaction have different outcomes
  - Agents receiving 100% resulting in targeted QA (Clothing Retail Company)
- **Job to Done:** Deep dive into these processes to understand why agents resolving differentiating
- **Outcome:**
  - Policy changes
    - Resulting in consistent agent / customer outcomes
  - Rubric changes (Clothing Retail Company)

# Common Challenges

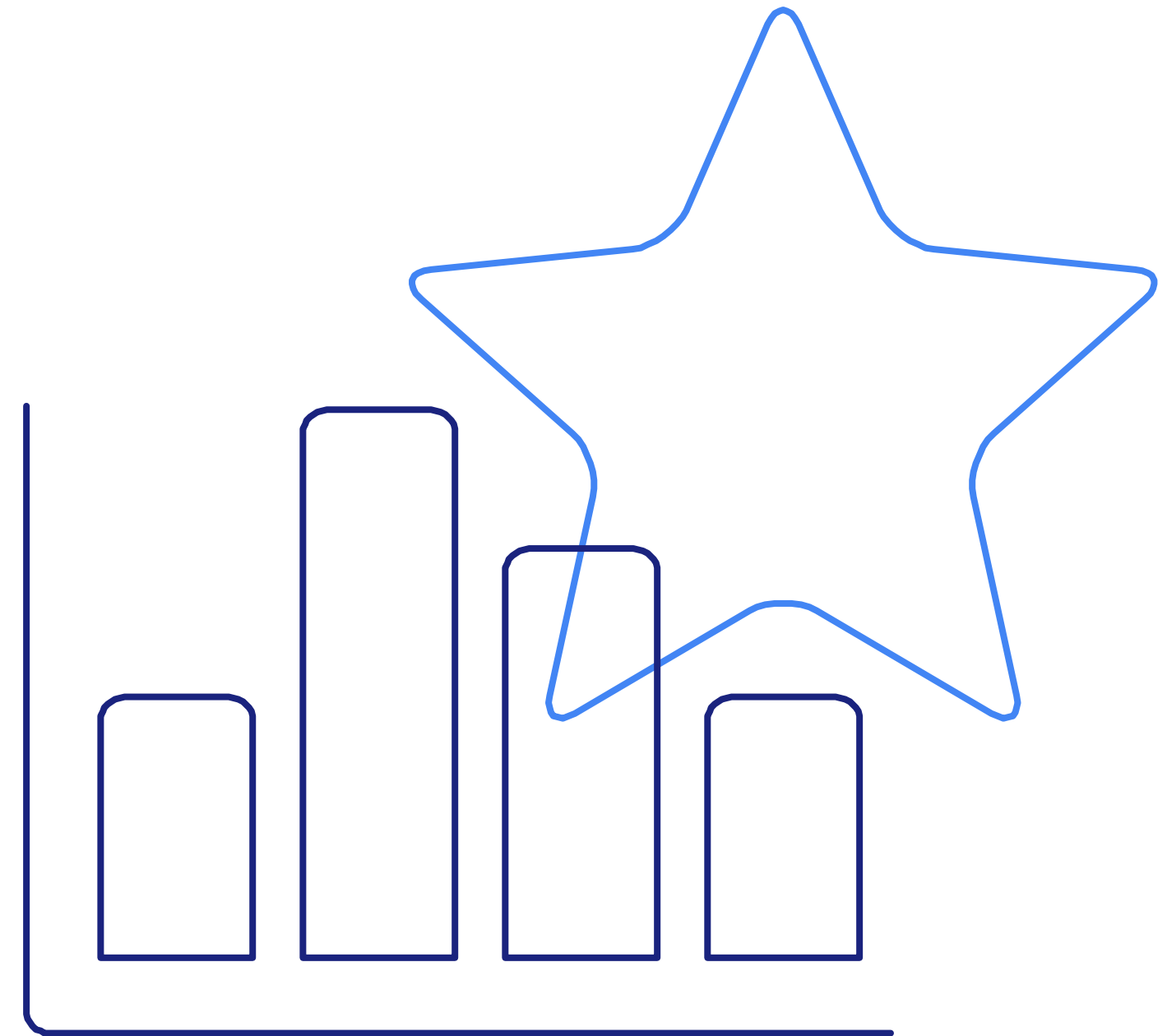
## Agent Performance General Theme of Challenges

- **Problem statement:** Sample size is small due to limited QA resources and technology. We know certain behaviors are happening but we can't see the scale at which this is occurring
- **Job to be done:** Root cause analysis
- **Outcome:** Lead to agent performance improvements, policy changes. High level of effort to understand and see trends at scale



# Key Takeaways on Jobs

- **Agent performance trends:** Use targeted QA and coaching to improve behaviors + improve policies
- **BPO:** Measure BPO's using GraderQA / calibrations to ensure they are being held to consistent standard
- **AI research:** Ability to see trends at scale to know where to target QA (agent performance, policy, etc.)



# Where Have Impact; Priorities and Jobs To Be Done (cont.)

1. Measuring agent performance; insights into agent performance
  - a. Success Stories
    - i. Online Fantasy Sports/Sportsbook Provider - new product. Want to direct customers there than reach out to get their referring information. Find out if the agents are referring customers. Reached out to George to build out a report to find out how often mentioned
      1. Got # calls and # of mentions
      2. Coach agents to refer to the referring hub to lessen reach outs
  - b. Challenges
    - i. Work Marketplace Provider; are you grading for this or grading for that. Work with CX directors to make sure grading things that they are looking for. Non-gradable section for the analysts to raise things that may result in changes to be made.
      1. They're team is "customer experience improvement" rather than "QA"
        - a. Look beyond just the interaction to see if larger process change to make
      2. QA process leads to process change but leadership often doesn't see that because process changes don't call out being made based on learnings through the qa process, so leadership mostly just looks at qa as a police force to make sure agents haven't done something wrong rather than what value the qa program really delivers
      3. So much information and feels like just barely touching the surface of what's in there. Share a weekly business review, but she mostly doesn't know what to do with all of the data. Can see something is trending in a certain direction but often doesn't have the sample size she wants and isn't sure what to do with it anyways – seems like always training as the outcome

# Where Have Impact; Priorities and Jobs To Be Done

- ii. Sometimes trainings and processes change for the same thing and it can confuse the agents because they learn to do the same thing three different ways. Debating if should make changes less, e.g. only after a quarter or two, to really let the change set in
  - iii. BPO graders are bias and want to give their agents good scores. So, they want to either lean into grade the grader or bring it in house. Alternatively, automated insights and a computer grading may be a better experience than human graders and agents feeling that there is human bias... a computer and objective thing may be better.
  - iv. Connecting one metric to following things... long handle time vs. short .. also really matters if the customer is going to call back / does call back
  - v. BPOs always seem to be operating at 100% even though they aren't... will submit appeals for anything less. Important for their comp / contract.
    - 1. Found some success in bringing the BPOs to the table in creating rubrics rather than just using theirs or just using the internal one without them
    - 2. Also found success in redefining the range... that 75% is their 100% and 100% is the customer is ecstatic and way above and beyond
- 2. Silos; BPOs being siloed
  - 3. Targeted QA + RCA

**Group 7**

**CX**   
**SUMMIT**  
NYC | 2023

# Biggest QA Challenges

- Overcoming the bad guy issue
  - Making sure the agents know we're here to help and not to criticize
  - Ultimate goal is to offer the best customer experience possible
- Finding the perfect scorecard structure
  - How to better utilize point system and not be too nit picky
- Driving internal engagement with QA data

QA  
Program  
Wins



<input type="checkbox"/>	Front Line Customer Service Quality Scorecard Updated 03/29/2023	Edit	Clone	Edit Access	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Drop Ship & Delivery Support Quality Scorecard Updated 03/29/2023	Edit	Clone	Edit Access	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Fraud Quality Scorecard	Edit	Clone	Edit Access	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Print & Marketing Quality Scorecard V2	Edit	Clone	Edit Access	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Reso Quality Scorecard Updated 03/29/2023	Edit	Clone	Edit Access	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Social Media Scorecard	Edit	Clone	Edit Access	<input checked="" type="checkbox"/>

Created dedicated scorecards to each workflow to ensure feedback in most directly targeted for each specific process.

Soft Skills

Standard Section 50 % Weight

Did the agent exhibit empathy appropriately, when required?

\*Agent exhibited the right amount of empathy in the appropriate places throughout the conversation  
\*Agent exhibited willingness to assist the customer.

Multiple Choice | General Feedback |  Show / require feedback based on answer

Yes	is worth 5 point(s)	If selected	No Change
No	is worth 0 point(s)	If selected	Show and Re
N/A	is worth 5 point(s)	If selected	No Change

Simplified the scorecard structure to help increase engagement and streamline the grading process.

# Coaching sessions

Agent

All agents ▾

Filter by template

All templates ▾

Date field

Created Date ▾

Date range

past 30 days ▾

New Session

View Report

AGENT ▾	NAME OF SESSION ▾	TEMPLATE ▾	SESSION TYPE ▾	CREATED ▾	SESSION DATE ▾	COACH ▾	STATUS ?
Patrizia	P4 2023	Original Templat	In person	06/07/2023	06/07/2023	tanya@sta...	✓ ✓ ✓
Tiwana	P4 2023	Original Templat	In person	06/07/2023	06/07/2023	tanya@sta...	✓ ✓ ✓
Viji	P4 2023	Original Templat	In person	06/05/2023	06/05/2023	tanya@sta...	✓ ✓ ✓
Vishal	P4 2023	Original Templat	In person	06/05/2023	06/05/2023	tanya@sta...	✓ ✓ ✓

Having structured coaching program within Maestro has been a huge win for both internal and BPO agents and TLs. Increased engagement due to direct and targeted feedback.



## Top Agent Groups



Group	Avg QA Score	Count
Charmaine - Fraud	98.0%	32
Tanya - Print & Marketing	98.0%	94
Gina - RESO	95.0%	107
Karen - OPS: Delivery Support/Dropship	95.0%	143
Ronald - Tarlac	94.0%	190
Omar - Fes	94.0%	570
Bilingual - All Agents	94.0%	1.11k

Leadership and agents review data from QA performance dashboard monthly to understand broader team performance and to celebrate high performing agents.

Looking at both QA and CSAT metrics within these monthly reports.