



# Evaluating a Customer Service Workshop

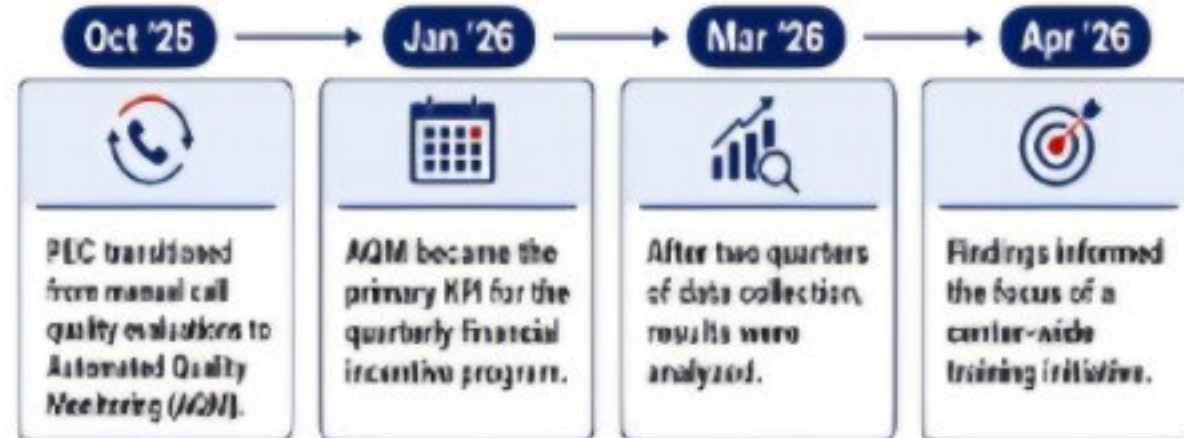
A Difference-in-Differences Analysis of Contact Center Agent Performance

Cole A. Lyons MBA, Kristi M. Madden BA, William R. Stewart M.Ed

The University of Pennsylvania Health System, Patient Engagement Center



## Background



**Aim:** Use Automated Quality Monitoring (AQM) results to inform the subject of a center-wide training initiative and improve agent AQM performance.

### Objectives:

- Support the development of a center-wide training to address the identified AQM KPIs.
- Monitor performance in the prioritized AQM KPIs to assess the impact of the training initiative.

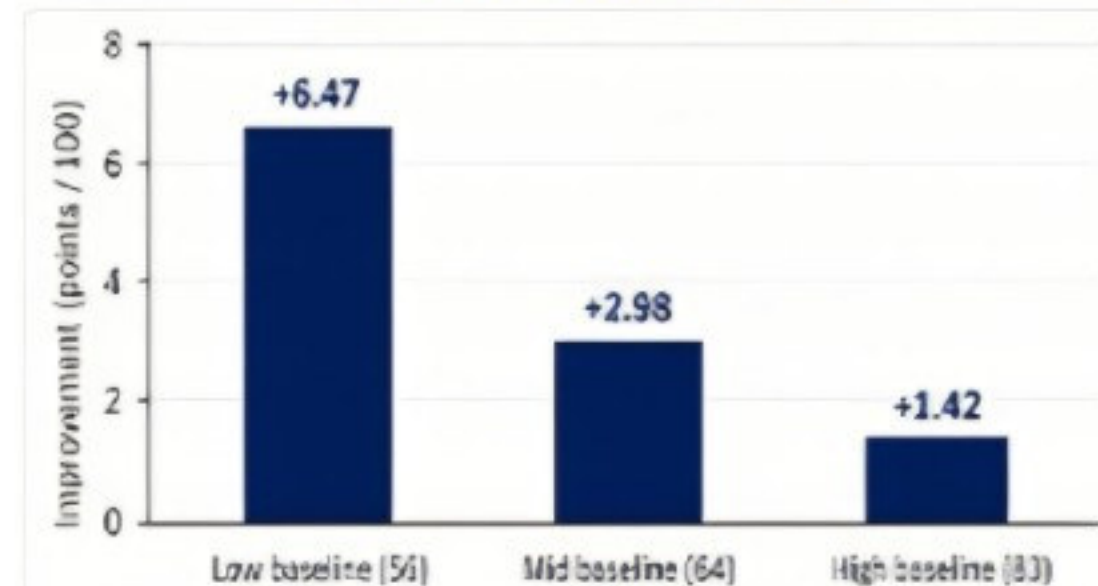
## Approach



## Results



### Effect by Pre-Training Baseline



**Results analysis.** Live workshop attendance was associated with statistically significant improvement in AQM Standards. Courtesy Phrases showed the larger lift while Positive Language improved more modestly.

**Opportunity gradient.** The strongest Courtesy gains occurred among agents with the lowest pre-training scores. No statistically significant ACD penalty was detected (95% CI -0.25s to +5.79s;  $p = 0.072$ ).



**Training reach and completion.** For the first time we had 100% completion by formulating a partnership with workforce analysts & trainers for real-time monitoring of agent training attendance and participation. Only one makeup session was needed out of the six available.

## Methods

- 1. Design.** Quasi-experimental difference-in-differences comparing workshop attendees with contemporaneous matched untrained agents across the same pre/post period.
- 2. Cohort & Matching.** Agents were included with sufficient Verint AQM volume; comparison agents were matched using baseline performance, queue/workgroup context, and evaluation volume.
- 3. Outcomes.** Primary KPIs were Courtesy Phrases and Positive Language; ACD served as a secondary balancing metric for efficiency trade-off.
- 4. Analysis.** Two-way DID estimated treatment effects, corroborated by Mann-Whitney U testing and OLS regression adjusted for baseline score and evaluation volume.
- 5. Data.** Verint Automated Quality Monitoring evaluations, workshop attendance records, and queue metadata were aggregated to the agent-month to stabilize estimates.

## Conclusion & Next Steps

### Takeaways:

- Training produced a statistically and practically significant lift in Courtesy Phrases, with a smaller but detectable effect on Positive Language.
- Low-baseline agents improved most; the intervention delivers its largest returns where they are most needed.

### Lessons learned:

- Create a communication channel for trainers and workforce analysts to provide real time monitoring and support for agents throughout the training.
- Centralize training schedules to the Workforce analysts by using read-only scheduling documents.
- Re-examine the Positive Language criteria to determine whether performance gaps reflect a need for additional agent coaching or refinement of the evaluation method.
- Obtain additional months of pre-intervention data to strengthen analyses by providing a more reliable baseline for comparison.

### Next steps:

- Conduct long-term tracking of KPI performance improvement to determine if there is a sustained effect.
- Utilize AQM data insights to targeted one on one coaching sessions lead by trainers for agents who continue to fall short of the established performance standard.