



CX as a Leading Indicator of Agent Performance

Executive Summary

It is critically important to use the right tool for the job – you wouldn't use a thermometer to measure your speed in running a mile and you wouldn't use a stopwatch to see if you have a fever. Net Promoter Score (NPS) is the gold standard in measuring enterprise growth by quantifying your customers' willingness to recommend your company to others; however, its use has expanded into individual channels – including the call center voice channel – as a metric for reporting on customer experience delivered by call center professionals. Cogito has analyzed the NPS scores for nine of its enterprise clients and found that NPS survey responses alone do not distinguish the business-relevant impact of agent performance on customer experience in an operationally relevant time frame. By contrast, Cogito's Customer Experience (CX) score enables call center managers to make agent performance evaluations in two days or less, a much shorter timeframe than the three to twelve months needed to make a similar conclusion with NPS. For enterprise call centers, Cogito estimates that at least 1,400 survey responses are needed to distinguish between a true NPS score shift relating to agent performance and statistical noise. This paper presents a quantitative analysis of NPS and CX scoring data from multiple Cogito clients, along with guidance for sample sizes needed to make data-driven decisions.

By using the right tool for the job, your evaluation of agents can take place in days, not months or years, leading to improved coaching and improved performance management in a much shorter time frame.

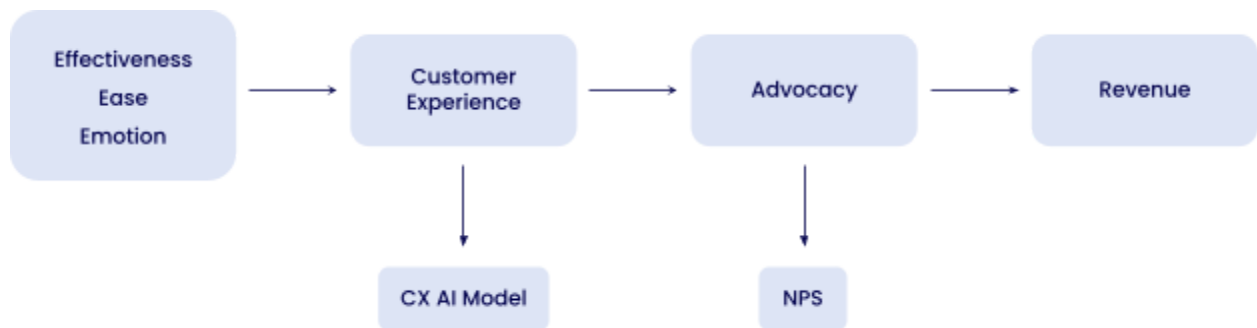
The Origins of NPS

Company executives have long desired a way to operationalize customer loyalty as a predictive measure of company success. In the search for an enterprise-level metric, [Fred](#)

[Reichheld and Bain & Co collaborated with Satmetics in 2003](#) to research ways to predict customer repurchase and referral behavior. They created twenty questions that were designed to categorize customers into groups - known today as Promoters, Passives/Neutral, and Detractors. Of these questions, one stood out as the most powerful, statistically significant predictor of desired customer behavior: "What is the likelihood that you would recommend Company X to a friend or colleague?"

Bain's NPS became the gold standard for measuring customer loyalty and, consequently, predicting company growth. However, business leaders began using NPS in other ways, some of which were not suited for the metric. In particular, the call center industry was eager to measure customer perception across all channels, including one-on-one customer service conversations. As a result, NPS is now used to gauge the occupational competency of individual agents. The NPS question is still frequently distributed to customers as a post-call phone or email survey in an attempt to understand the quality of their customer service interactions; despite NPS being a subjective, survey-based measurement. NPS is definitionally unsuited for evaluating individual performance of customer service interactions.

Complexities in Measuring Customer Experience as a Measure of Agent Performance



[Research from Forrester](#) has shown that customer experience is extremely important to organizations because of the significant impact it has on a company's revenue engine. Customers in most industries decide if they want to continue on with a company, if they want to give a company more or less of their wallet share, or whether or not they want to advocate on behalf of (or against) a company, based on their own customer experiences. These factors of retention, enrichment, and advocacy are explicitly linked with revenue and are positively correlated to customer experience: the better the customer experience, the more positive impact to your bottom line. These revenue dynamics vary by industry; more competitive industries showing more dramatic effects, whereas industries with lower

customer churn or high switching costs can show a muting effect on the impact of customer experience on revenue.

Customer experience is made up of customers' holistic perceptions of their experiences with a business or brand. [Forrester's research](#) shows that customer experience is driven by the 3 E's:

- **Effectiveness** – how much value are you providing your customers in their interactions with you.
- **Ease** – how much work are you making your customer do to get the value provided.
- **Emotion** – how have you made your customers feel during their interactions with you.

Since Cogito's CX score measures the overarching customer experience, some thoughtful experimental design must be exercised when using the CX score as a measure of agent performance. We must ensure that there is no inherent bias in routing more emotionally charged customers to certain agents. We must also be careful in comparing across sites to ensure that they have the same systems, processes, and policies that often impact the ease of the interaction with the customer. Once we can normalize the groups being compared to ensure that consistent systems, processes, and policies are being used along with no inherent bias in call routing, we can proceed to use CX as a guide to determine agent performance. Individual performance can be assessed by comparing to peers that answer calls from similar customers, similar call contexts, and use the same systems, processes, and policies. As Cogito continues to innovate, we're working on even more fine-tuned measures of agent effectiveness that would eliminate the need for such considerate experimental design in the future.

Traditional Performance Measurement in the Call Center

Call center executives and site leaders rely on their agents and supervisors to deliver high-quality customer service in every customer interaction. Therefore, it is imperative that their frontline workers are high-performing, emotionally intelligent employees. However, agents are subjected to repetitive yet complicated work and highly variant customer emotions. As a result, agents are at risk of experiencing long-term occupational stress, which can result in cognitive overload, compassion fatigue, and burnout. These conditions are not only a detriment to the work environment, but can have serious impacts on agent personal lives and psychological states. Supervisors are tasked with not only supporting their teams, but also keeping tabs on performance for hiring, firing, and compensation decisions. Unsurprisingly, these conditions lead to call centers with [turnover rates upwards of 30%](#).

Traditionally, call centers are difficult places to manage effectively; most quality measurement systems are delayed, costly, and incomplete. Most call centers must rely on isolated call listening, delayed feedback, and enterprise-level measures. NPS has become a commonly-used performance metric, despite its origins as a broader predictor of enterprise growth and company success. The danger lies in the extrapolation of customer perception to employee performance - a customer who is upset about a company policy, for example, is unlikely to recommend the business to others. This negative score falls on the agent who last spoke with the customer, who may be fired as a result.

As call centers adopt new technologies and move towards data-driven decisions, executives and site leads should reassess the use of NPS as the single measure of agent performance.

How Can We Analyze NPS?

NPS is based on a post-call survey to customers. Typically, the question asked is a variation of "How likely are you to recommend this company to others?" and responses can be made on a scale of 1-5 or 1-10. In each of these cases, a score at the top of the range (either a score of 5 or a score of 9/10) is considered to be a "Promoter", a score at the bottom of the response range (either 1-3 or 0-6) is considered to be a "Detractor", and the remaining scores are considered to be a "Neutral" (either a 4 or a 7/8). The Net Promoter Score is simply the percentage of Promoters minus the percentage of Detractors.

[Company NPS scores](#) are reported on a range from -1 (when every respondent is a detractor) to +1 (when every respondent is a promoter), and an industry standard benchmark for call center transactional performance is considered to be any value greater than 50%.

$$\text{NPS} = \frac{\text{Total Promoters} - \text{Total Detractors}}{\text{Total Responses}} = \frac{\text{Sum of Promoter Scores} + \text{Sum of Neutral Scores*} + \text{Sum of Detractor Scores}}{\text{Total Responses}}$$

The customer survey responses take on one of three values: Promoter, Neutral, or Detractor. Re-coding these values as +1, 0, and -1 (respectively) enables the attributes of this score to be analyzed quantitatively. After this re-coding, NPS is simply an average of the scores:

**Note that in this convention, the "Sum of Neutral Scores" is zero and the "Sum of Detractor Scores" is a negative number.*

Re-coding NPS scores numerically allows a standard deviation to be computed for the distribution:

$$\sigma = \sqrt{\frac{\text{Promoters}(1-\text{NPS})^2 + \text{Neutrals}(0-\text{NPS})^2 + \text{Detractors}(-1-\text{NPS})^2}{\text{Total Responses}}}$$

These measurements of mean and standard deviation can be computed at an agent or a population level and allow business analysts to identify differences between agent populations or changes over time using standard statistical significance tests.

Analysis: NPS as a Measure of Employee Performance

Given this quantitative background on the NPS scoring distribution, it is possible to measure the robustness of this metric for use in agent performance measurement. Or, in other words, how many post-call survey data points are needed to distinguish between a high-performing (or a low-performing) agent & an average agent using NPS scores?

The Student's t-test for a statistically significant difference between an agent's NPS score & a population baseline is:

$$t = \frac{NPS_{\text{Agent}} - NPS_{\text{Population}}}{\frac{\sigma}{\sqrt{n}}}$$

Where:

NPS_{Agent} is the NPS of the agent of interest

$NPS_{\text{Population}}$ is the NPS of the entire agent population

σ is the standard deviation of the post-call survey response distribution

n is the number of survey responses collected

t is the parameter which determines the likelihood that the difference in an agent's NPS at least as large as the one observed could have occurred if the agent's NPS matched the population NPS

Rearranging this equation to focus on the sample size that is needed to determine if an individual agent's performance differs from the average:

$$n = \left(\frac{t \cdot \sigma}{NPS_{Agent} - NPS_{Population}} \right)^2$$

For $n \geq 30$, the t-distribution is approximated by the standard normal distribution ($=0, =1$). For $n < 30$, n is determined iteratively, since t is a function of n .

Analysis: NPS as a Measure of Employee Performance (Cont.)

For the subsequent analyses, a “high-performing” agent is an agent with an NPS score at the 90th percentile of the agent population and a “low-performing” agent is one with an NPS score at the 10th percentile. These percentile thresholds were used because the performance difference between an average agent and an agent at the 90th (or 10th) percentile is considered to be business-relevant. The table below provides anonymized data from nine of Cogito’s clients:

	Client A	Client B	Client C	Client D	Client E	Client F	Client G	Client H	Client I
Industry	Insurance	Finance	Insurance	Insurance	Insurance	Pharmacy	Telco	Consumer	Telco
Average NPS	47.2%	54.1%	53.9%	52.6%	67.5%	87.2%	62.0%	65.0%	84.7%
Low-Performing Agent NPS	25.4%	32.1%	27.0%	28.0%	35.7%	80.6%	36.4%	47.8%	74.0%
High-Performing Agent NPS	65.3%	64.7%	77.5%	76.2%	86.7%	92.6%	79.5%	80.6%	93.1%
NPS Standard Deviation	81.4%	77.5%	75.9%	76.1%	65.2%	45.8%	78.4%	69.9%	51.4%

For the scenario in which a call center manager has to identify the highest and lowest performing agents from a pool of average agents, a high level of statistical confidence is required: a low confidence level when conducting hypothesis tests of many agents will lead to multiple false-positives. Falsely identifying an average agent as a low-performer can have a significant impact on agent well-being (especially if compensation is tied to performance data), so a high confidence level of 99.5% is chosen for this context.

Analysis: Time Needed to Detect Performance Level

Applying the sample size equation to the client data listed in the table above generates the following survey requirements needed to identify top and bottom performers at:

	Client A	Client B	Client C	Client D	Client E	Client F	Client G	Client H	Client I
Top Performer	135	359	69	69	77	487	133	134	249
Worst performer	92	82	53	64	28	316	62	109	154

As a next step in this analysis, each client's survey response rate and average call volume is displayed in the table below:

	Client A	Client B	Client C	Client D	Client E	Client F	Client G	Client H	Client I
Survey response rate	3.50%	4.86%	3.53%	0.76%	0.24%	11.58%	3.15%	8.07%	8.15%
Average calls per agent per day	21.3	32.0	31.4	26.8	16.4	31.1	33.2	14.2	19.5

The final step of this analysis is to determine the number of months of data that must be collected to distinguish agent performance based on post-call NPS survey data. This analysis assumes that each month is 4.3 weeks and that agents will handle calls five days per week.

	Client A	Client B	Client C	Client D	Client E	Client F	Client G	Client H	Client I
Months needed to identify a top performer	8.4	10.7	2.9	15.8	90.7	6.3	5.9	5.4	7.3
Months needed to identify a worst performer	5.8	2.5	2.2	14.5	32.9	4.1	2.8	4.4	4.5

For some of Cogito's clients, over **one year of post-call survey data** is needed to distinguish top and bottom performers.

Analysis: Use of NPS for Enterprise Performance

Clearly, NPS is not useful as a tool for distinguishing top and bottom performers on an agent level. How useful is this measurement system for detecting shifts at an enterprise level call center? Given 300 or 1,000 or 3,000 post-call survey data points, what is the size of the shift from the long-term average that can be distinguished from noise?

Equation 1 can be rearranged to focus on this question:

$$NPS_{\text{Sample Period}} - NPS_{\text{Long - Run Average}} = \frac{z * \sigma}{\sqrt{n}}$$

A z-score of 1.96, corresponding to a “confidence level” of 95% is a reasonable value for this analysis. The table below gives the sample sizes needed to distinguish a true shift in NPS for Cogito’s clients’ NPS score ranges:

NPS Score Range	40% - 50%	50% - 60%	60% - 70%	70% - 80%	80% - 90%
NPS Standard Deviation	81.4%	76.5%	71.1%	<i>No Cogito clients have an NPS score in this range</i>	48.6%
Responses needed to detect 6-point shift	707	625	540		252
Responses needed to detect 4-point shift	1,592	1,405	1,215		566
Responses needed to detect 2-point shift	6,366	5,621	4,861		2,265

Given that most of Cogito’s clients have an NPS standard deviation between 0.75 and 0.80, a useful heuristic for detecting NPS shifts over time is that 950 responses are needed to detect a shift of five percentage points, **1,450 responses are needed to detect a shift of four percentage points, and 2,600 responses are needed to detect a shift of three percentage points.**

The CX Score Advantage

Given this long lag in NPS as a measurement system for identifying agent performance issues or shifts in call center or enterprise service levels, Cogito has developed a real-time measurement of the conversational behaviors on every call. The Cogito CX score is an objective measure of the customer’s experience. Unlike NPS, the CX score distribution is continuous and normal, and a data point on the customer’s experience is generated on every call. These properties enable CX-based analyses to detect smaller performance shifts using fewer data points.

Cogito CX score data for the same seven clients previously analyzed is presented in the table below:

	Client A	Client B	Client C	Client D	Client E	Client F	Client G	Client H	Client I
Population Average CX	4.817	4.816	4.142	5.235	4.913	5.217	4.715	5.225	5.342
10th percentile Agent CX	3.940	4.007	3.528	4.614	4.055	4.482	4.054	4.343	4.671
90th percentile Agent CX	5.805	5.463	4.756	6.008	5.686	6.113	5.355	5.840	6.020
Population CX Standard Deviation	1.568	1.597	1.671	1.318	1.524	1.412	1.758	1.572	1.561

CX Score as a Measure of Employee Performance

In the same way this data was analyzed for NPS, an estimate of the number of CX score data points needed to identify good and bad agents can be determined:

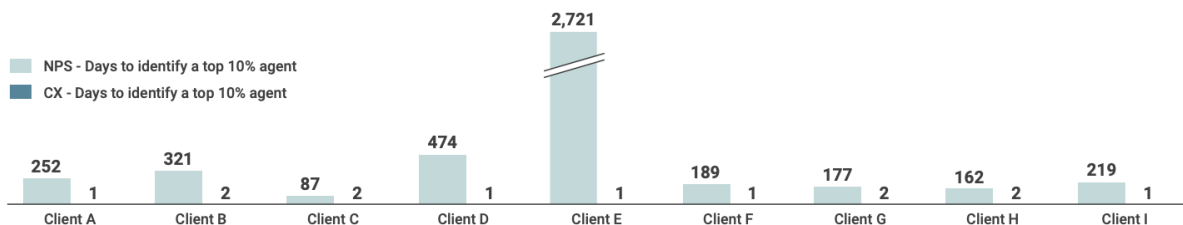
	Client A	Client B	Client C	Client D	Client E	Client F	Client G	Client H	Client I
CX scores needed to detect a top performer	21	44	53	23	30	21	54	47	39
CX scores needed to detect a worst performer	25	30	53	34	25	28	51	25	40

In every instance, the number of CX scores needed to distinguish agent performance is less than the number of survey responses, and for some clients, this difference is greater than a factor of 10.

But this fact “undersells” the relative power of CX compared to NPS for distinguishing agent performance. A CX score is generated on every call, compared to a post-call survey response, which may be generated on fewer than one in a hundred calls. The table below incorporates this response rate and shows, for each client, how much faster top-performing and a low-performing agents can be distinguished from the average based on CX score data compared to NPS data:

	Client A	Client B	Client C	Client D	Client E	Client F	Client G	Client H	Client I
Relative speed to identify top performer using CX vs. NPS	187x	166x	37x	389x	1079x	204x	78x	35x	78x
Relative speed to identify worst performer using CX vs. NPS	105x	57x	28x	248x	467x	96x	39x	54x	47x

Days Needed to Collect Sufficient Data By Which to Categorize Agents As Top / Bottom 10% Performers with 99.5% Confidence



CX Score For Detecting Enterprise Growth

The analysis above supports Cogito's CX score as a faster measure of agent performance, but how does it compare to NPS in measuring enterprise performance? Because CX is measured throughout every phone call, there are substantially more data points than a randomized survey. This amount of data ensures a robust measurement that is resistant to the skew of outliers and inaccurate assumptions. CX score analyses are also protected against the subjectivity and bias of customer survey data. Additionally, CX scores provide real-time insight into the quality of customer experiences, whereas NPS responses are often delayed by a number of weeks.

Based on these advantages, it is safe to conclude that the CX score is a comparatively stable measurement, and that statistically significant shifts in the score are more useful for making strategic decisions about call center initiatives and strategy.

Conclusion

Although NPS has been the cornerstone of quantifying customer engagement for many years, our analyses indicate that it is better suited as a quick predictor of overall enterprise success. The subjectivity and infrequency of NPS results make it ill-suited for use as an agent performance measurement system.

In regards to a sole measure of individual call center professional performance, NPS may be disadvantageous for agents, supervisors, and site leads. These scores are a reflection of the customer's opinion of the enterprise at a specific point in time - typically after an issue has occurred that required a call in the first place - rather than the quality and competency of the agent in which they spoke to. Supervisors, who are held responsible for the success of their team, are forced to use NPS results to justify key hiring and firing decisions. Lastly, site leads and executives strive to design data-driven strategy but are ill-equipped to do so with NPS as their key indicator.

We believe that Cogito's CX score is a superior estimate of employee performance due to its ability to measure customer experience constantly and in real-time. The CX score detects behavioral signals of both the agent and the customer in every phone call. Scores are calculated with a massive amount of data and provide unparalleled insight into the emotional intelligence of call center professionals. Employees of all levels are able to track CX throughout the day, month, or year and can alter their behavior accordingly. Agents are empowered to improve their own performance, supervisors are enabled to coach their teams, and the decision-makers have the data needed to inform their strategy.

However, NPS has been shown to be highly correlated with company success. Therefore, it would be best to pair NPS with other customer experience metrics. We believe that NPS is a worthwhile supplement to the CX score – just as qualitative and quantitative measures compliment each other in other realms. Call center executives can use this data to have a robust understanding of their customers and employees. For example, the following questions can be analyzed with a combination of NPS and CX score data:

- **How do customers feel about our latest policy change?**
- **Which call center team is most effective?**
- **How were agents and customers impacted by COVID-19?**
- **Who is worthy of a promotion and who requires additional coaching?**
- **How effective are in-office supervisors versus those who work from home?**
- **Was last month's employee training session worthwhile?**
- **How long does it take an agent to be fully onboarded?**

These are just a few of many possibilities. As call centers are adopting new technologies, it is imperative that the robustness of their measures evolve as well. As was mentioned, it is important to use the correct tool for the job. NPS is the right tool for quickly and conveniently predicting enterprise growth; Cogito's CX model is the right tool for objectively measuring agent performance through the lens of the impact on customer experience in real-time. The next generation of call center executives must be equipped to draw faster and more reliable insights on employee performance and customer engagement.