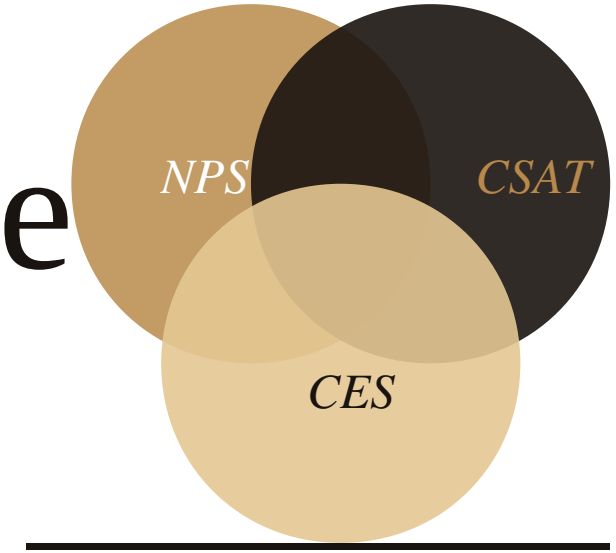
The complete guide to CX metrics.

NPS, CSAT, CES — and when to use which.



THE BIG THREE OF CX MEASUREMENT

A practical handbook for CX directors, managers and analysts who need a confident answer to the question every executive eventually asks — "are we measuring the right thing?" This paper compares the three metrics that run 90 % of CX programmes today, with formulas, worked examples, decision rules and the pitfalls we see most often in the wild.

Contents

INTRODUCTION

Why a single number rarely tells the truth.

Ask ten CX leaders which metric they trust the most and you will get three answers — NPS, CSAT, or CES — defended with the conviction of religious doctrine.

The truth is duller and more useful. Each of the three metrics measures a different thing. None is universally right. They are tools, and like any tool, they fail when used outside their intended job. A team that picks the wrong metric for the wrong moment ends up with numbers that move, charts that look credible, and customers who quietly leave anyway.

This paper exists because we have watched roughly forty client programmes pick the wrong metric, calculate it badly, or report it on the wrong cadence — and then conclude that "CX measurement doesn't work." It works. You just have to use the right one for the question you are asking.

We cover all three in depth: what they actually measure, the formula with a worked example, where they shine, where they break, and the industry benchmarks you can compare against. Then a side-by-side matrix, a decision framework, six pitfalls, and a self-assessment scorecard.

WHAT YOU'LL GET

A clear definition, formula and example for NPS, CSAT and CES.

A side-by-side comparison across eight practical dimensions.

A four-question decision framework for picking the right one.

Six pitfalls we see in nearly every CX programme.

A self-assessment scorecard for your current measurement setup.

WHO THIS IS FOR

CX directors and managers responsible for the measurement programme.

Voice-of-customer leads designing or rebuilding their survey stack.

Heads of operations, service or product who consume CX metrics and want to know what the numbers actually mean.

Analysts and data leads who own the reporting layer.

WHAT'S NOT HERE

This paper is metric-agnostic on vendor choice and does not benchmark survey platforms. The aim is to make you fluent in the metrics themselves — so that whichever platform you use, you use it well.

CHAPTER 1

The anatomy of a CX metric.

Five questions before any number means anything.

1

A metric is not just a number. It is a design choice.

Before reaching for NPS, CSAT or CES, separate the five design choices that determine what the resulting number actually means. Two teams measuring "satisfaction" can produce numbers that look comparable and are not, because these five choices were configured differently. Get them right and the metric becomes load-bearing. Get them wrong and you are decorating dashboards.

01

What is being measured?

Loyalty? Satisfaction with one moment? Effort to get a job done? The metric must match the underlying construct, not the other way around.

02

When is it asked?

Right after a touchpoint (transactional) or periodically about the relationship as a whole (relational)? The same question produces different numbers at different moments.

03

Who is asked?

Every customer, a sample, only buyers, only churners? Sampling bias is the most common reason CX scores look better than reality.

04

How is it asked?

Email after 24 hours, SMS after one hour, a kiosk at the exit, a pop-up in-app? Channel and timing change response rate and bias.

05

What is done with the answer?

If the answer triggers no action, the metric is decorative. A good metric is wired to a decision — a routing rule, a recovery workflow, a quarterly review.

THE QUICK DIAGNOSTIC

Pick your most prominent CX metric. Write down each of these five answers in one line. If any of the five is "we never decided" — that is the place to start fixing measurement, not the metric itself.

Relational vs. transactional — the most useful split

RELATIONAL METRICS

Ask: "How do you feel about us as a company / brand / relationship overall?"

Cadence: quarterly or twice a year, to a sample of the active customer base.

Strength: tracks strategic direction. Good for board reporting and year-on-year comparison.

Weakness: low actionability. By the time the number drops, the cause is months old.

Typical fit: NPS (relational), annual CSAT.

TRANSACTIONAL METRICS

Ask: "How was this specific interaction?" — tied to a moment (purchase, support ticket, delivery, onboarding).

Cadence: continuous, fired by the event itself.

Strength: high actionability. Pinpoints which touchpoint is breaking and routes recovery in hours, not quarters.

Weakness: easy to drown in data without a clear hierarchy.

Typical fit: CSAT (per touchpoint), CES (per task).

CHAPTER 2

NPS
Net Promoter Score.

The loyalty metric — brilliant, abused, and frequently misread.

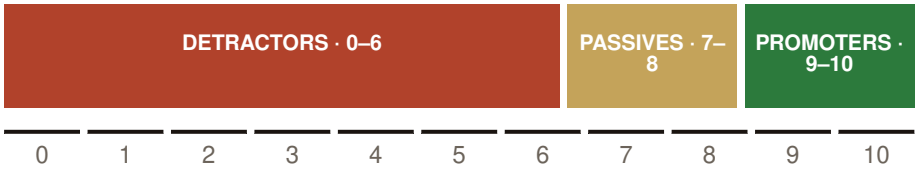
The single question that built a category.

NPS was introduced by Fred Reichheld in 2003 and has become the most widely used CX metric in the world — tracked by roughly two thirds of the Fortune 1000. It tries to measure **loyalty**: not how someone felt today, but whether they will stay, buy again, and recommend you.

THE QUESTION

"On a scale of 0–10, how likely are you to recommend [company / product] to a friend or colleague?"

HOW THE ANSWER IS BUCKETED



Note the asymmetry: 0–6 is seven buckets, 7–8 is two, and 9–10 is two. This is intentional — loyalty is rare, and the score punishes mediocrity heavily.

THE FORMULA

$$NPS = \%\ Promoters - \%\ Detractors$$

The result is a number between –100 and +100. Passives are deliberately ignored.

A WORKED EXAMPLE

1,000 responses from a relational survey.

Promoters (9–10)	540 — 54 %
Passives (7–8)	260 — 26 %
Detractors (0–6)	200 — 20 %

$$NPS = 54 - 20 = +34$$

READING THE NUMBER

The absolute value matters less than people think. What matters is the trend, the segment breakdown (B2B enterprise vs SMB will differ wildly), and the verbatim that explains the score. A trending NPS with no verbatim analysis is a vanity metric.

TWO FLAVOURS

Relational NPS — asked periodically about the brand or company as a whole. Strategic indicator.

Transactional NPS (tNPS) — asked after a specific touchpoint. Operational indicator. Same question, different framing in the second sentence ("How likely are you to recommend us after this support call?").

What NPS is genuinely good at — and what it isn't.

STRENGTHS

One question, one number. Executive teams understand it in 30 seconds. That is rarer than it sounds.

Industry benchmarks exist. Almost every sector has published NPS ranges, so you can position yourself relative to peers.

Predicts behaviour better than satisfaction. Promoters genuinely do refer more, churn less, and buy more across most B2C contexts.

Forces a follow-up. The natural pairing is "Why did you give that score?" which produces the verbatim you actually act on.

Cross-cultural-ish. Works across markets reasonably well, though scoring norms differ (see weakness below).

WEAKNESSES

Hides 26 % of your customers. Passives are mathematically invisible. A programme that moves passives to promoters and a programme that moves detractors to passives produce the same score change.

Cultural scoring bias. Northern Europeans rate harshly; Latin Americans rate generously. Cross-market NPS comparisons need normalisation.

Not actionable on its own. A score of +34 tells you nothing about why. Without verbatim and segment analysis it is a thermometer with no diagnostic.

Lagging. Relational NPS reflects months of cumulative experience. By the time it drops, the operational cause may be old or already fixed.

Easily gamed. Tie bonuses to NPS and you get score gaming — rounded-up surveys, complaints suppressed, awkward customers excluded from sampling.

USE NPS WHEN

You need a single executive-friendly KPI for board reporting.

You operate in a market with published benchmarks you can compare against.

Loyalty and word-of-mouth materially drive your growth (most B2C, subscription, marketplaces).

You have the operational capacity to analyse verbatim — the number alone is not enough.

TYPICAL INDUSTRY RANGES (RELATIONAL NPS)

Tech / SaaS	+30 to +50
Retail / e-commerce	+30 to +55
Banking / financial services	+10 to +35
Insurance	+5 to +25
Telco	–10 to +25
Healthcare	+10 to +40
B2B enterprise services	+25 to +50

Ranges are indicative, not statistical — treat them as a sanity check, not a benchmark.

CHAPTER 3

CSAT
Customer Satisfaction.

The workhorse — simple, immediate, perfect for a single moment.

The oldest CX metric — and still the most useful at the touchpoint.

CSAT predates NPS by decades. It asks one direct question: **how satisfied were you?** — usually right after a specific interaction. Because it is tied to a moment, it is high-signal for operational improvement: it tells you not just whether things are good or bad overall, but exactly which touchpoint is hurting.

THE QUESTION (TYPICAL)

"How satisfied were you with [this interaction / product / service]?"

THE SCALE — COMMON VARIANTS

CSAT does not have one canonical scale. The three most used are:

1–5 (most common)

1–7 (research-grade)

Adds resolution between "okay" and "good." Used in academic and serious research contexts.

1–10 (when pairing with NPS)

Allows comparison or combination with NPS. Heavier cognitive load on the respondent.

THE FORMULA

$$\text{CSAT \%} = \frac{\text{"satisfied" responses}}{\text{total responses}} \times 100$$

"Satisfied" usually means the top 2 boxes (4 & 5 on a 1–5 scale).

A WORKED EXAMPLE

500 responses from a post-purchase survey on a 1–5 scale.

5 — Very satisfied	280 (56 %)
4 — Satisfied	120 (24 %)
3 — Neutral	50 (10 %)
2 — Dissatisfied	30 (6 %)
1 — Very dissatisfied	20 (4 %)

$$\text{CSAT} = (280 + 120) / 500 \times 100 = \mathbf{80\ \%}$$

A FOOTNOTE ON "TOP-BOX ONLY"

Some teams report only the top box (5 only) — a stricter definition that pushes for excellence rather than acceptability. Top-box CSAT for the example above would be 56 %. Decide once and stick with the definition; comparing top-2-box numbers to top-box-only numbers is a classic reporting error.

WHERE IT FITS

CSAT is almost always **transactional**. Common firing points:

· Right after a support ticket closes

· After delivery or product receipt

· After onboarding milestones

High-signal at the moment — weak as a strategic indicator.

STRENGTHS

Intuitive. Customers understand it immediately. No mental conversion needed — "how satisfied?" is the most natural question in the entire CX vocabulary.

High response rates. Typically 15–40 % on transactional surveys, the highest of any CX metric — because the question feels relevant in the moment.

Operationally precise. Tied directly to a touchpoint, so a drop in CSAT after a process change is unambiguous: the change broke something specific.

Cheap to run. One question, one scale, one number. Works across email, SMS, kiosk, IVR, in-app — any channel.

Pairs naturally with verbatim. "What could we have done better?" produces immediately actionable comments tied to a real moment.

WEAKNESSES

Ceiling effect. Customers who are mildly happy still rate 4 or 5. Mature CSAT programmes routinely sit at 85–90 % and stop moving — you can't distinguish "great" from "just okay."

Doesn't predict loyalty. "Satisfied" customers churn at almost the same rate as "very satisfied" ones in most studies. Satisfaction is necessary but not sufficient for retention.

No standard scale. 1–5, 1–7, 1–10, top-box, top-2-box — benchmarking is treacherous because everyone calculates it slightly differently.

Recency bias. Heavily influenced by the most recent moment, not the relationship. A customer with three years of friction will still score 5 if the last interaction was good.

Self-selection bias. Voluntary CSAT responses skew toward extremes — the indifferent middle rarely answers.

USE CSAT WHEN

You need to measure a specific moment (support call, delivery, onboarding step).

You want a metric that operations teams can act on within hours, not quarters.

You are running A/B tests on processes or touchpoints and need fast signal.

You have many small touchpoints and need granularity rather than one strategic number.

Your customers find NPS too abstract for the moment (often the case after a single, small interaction).

TYPICAL INDUSTRY RANGES (TOP-2-BOX CSAT)

E-commerce delivery	85–92 %
Customer support / B2C	75–90 %
SaaS support	80–95 %
Banking branches	75–88 %
Telco call centres	65–80 %
Healthcare appointments	80–92 %
Retail in-store checkout	85–93 %

Indicative ranges from common industry studies; verify against your own historical baseline.

CHAPTER 4

CES Customer Effort Score.

4

The metric that finally predicts churn — if you ask it correctly.

The youngest of the big three — and the best predictor of retention.

CES was introduced in a 2010 Harvard Business Review article that argued something heretical at the time: customers do not want to be delighted, they want their problem solved with minimum friction. The research showed effort predicted loyalty more strongly than satisfaction or even NPS in service contexts. CES asks one question: **how hard did the customer have to work?**

THE QUESTION (MODERN, CES 2.0)

"[Company] made it easy for me to handle my issue."

— Rated 1 (strongly disagree) to 7 (strongly agree).

THE SCALE

Note that CES 2.0 is phrased as agreement with a positive statement — not as effort directly. The original CES 1.0 asked "How much effort did you personally have to put forth?" and was abandoned because customers found it ambiguous (was high effort good or bad?). Always use CES 2.0.

THE FORMULA (TWO COMMON VARIANTS)

$$CES = \text{average score (1–7)}$$

— or —

$$CES\ \% = \%\ \text{of responses that agree (5, 6 or 7)}$$

The percentage variant is easier to communicate to executives ("80 % of customers say it's easy"); the average is more sensitive to change.

A WORKED EXAMPLE

300 responses after a support call closed.

7 — Strongly agree	120 (40 %)
6 — Agree	90 (30 %)
5 — Somewhat agree	36 (12 %)
4 — Neutral	24 (8 %)
3 — Somewhat disagree	18 (6 %)
2 — Disagree	9 (3 %)
1 — Strongly disagree	3 (1 %)

Avg CES = 5.78 · CES % = **82 %**

WHY THIS MATTERS MORE THAN IT LOOKS

The HBR study found that 96 % of customers who reported *high effort* became less loyal — while delighting customers (going above and beyond) had almost no effect on loyalty. The lesson: removing friction matters more than adding magic. CES is the metric that operationalises that finding.

The best predictor of churn in service-heavy contexts.

STRENGTHS

Best churn predictor in service contexts. Multiple studies show CES outperforms both NPS and CSAT at predicting next-renewal or next-purchase behaviour in support-driven categories.

Diagnostically rich. A low CES on a specific touchpoint points directly at process: too many steps, unclear policy, repeated escalations. The metric tells you where to look.

Operationally cheap to act on. Removing effort is engineering work, not culture work — faster and more measurable than "improving the experience" in the abstract.

Works well in B2B. Where loyalty is driven by reliability and time-to-resolution more than enthusiasm, CES outperforms NPS as a leading indicator.

Cuts through politeness bias. "It was easy" is a fact-y statement; "I would recommend you" feels like a favour. CES tends to be less inflated.

WEAKNESSES

Narrow construct. CES only measures effort. It says nothing about emotional connection, brand affinity or whether the customer enjoyed the experience — just whether it was easy.

Wrong tool for delight-driven categories. In hospitality, luxury, premium retail or anything where customers *want* to invest effort (cooking with you, designing with you), CES misses the point entirely.

Less executive recognition. Board members know NPS. They sometimes know CSAT. Many do not know what CES is. Internal advocacy is more work.

No universal benchmark yet. The metric is young enough that industry benchmarks are still thin and inconsistent.

Requires touchpoint precision. "Was it easy?" only makes sense if "it" is a specific, defined task. Asking CES about a whole relationship produces noise.

USE CES WHEN

You operate in a service-heavy or support-heavy category (telco, SaaS, banking, insurance, healthcare admin).

You want a leading indicator of churn — CES drops months before contracts cancel.

You are running a process-redesign or self-service initiative and need to measure friction directly.

You already have CSAT or NPS and need a metric that actually points at *what to fix*.

Your business model depends on repeat usage where every additional click costs you.

TYPICAL INDUSTRY RANGES (CES %, TOP-3-BOX)

SaaS / self-service	75–90 %
Banking digital	70–85 %
Telco support	55–75 %
Insurance claims	50–72 %
Healthcare admin	55–75 %
E-commerce returns	75–90 %
B2B onboarding	65–82 %

Ranges based on InsightSofa client data and published research; treat as orientation, not authority.

CHAPTER 5

Side-by-side comparison.

Eight dimensions, three metrics, one matrix.

The matrix — rated across eight practical dimensions.

No metric scores well on every dimension. Use the matrix to **eliminate**: strike out the row whose weakness would be fatal in your context, and pick from what remains. Most mature programmes use two of the three in combination (typically a relational NPS plus a transactional CSAT or CES), not one alone.

DIMENSION	NPS	CSAT	CES	WHAT IT TELLS YOU
What it measures	Loyalty / advocacy	Satisfaction with a moment	Effort to complete a task	NPS = relationship; CSAT = moment; CES = friction.
Best cadence	Relational (quarterly)	Transactional (per event)	Transactional (per task)	Mix relational + transactional for full coverage.
Predicts churn	Medium	Low	High (service)	CES wins for service-heavy categories.
Predicts referrals	High	Low	Low	NPS designed for this exact prediction.
Operational actionability	Low	High	High	Relational NPS tells you "where," not "what."
Executive recognition	High	High	Medium	CES still needs internal education in most companies.
Response rate (typical)	10–25 %	15–40 %	15–35 %	CSAT highest; NPS suffers from "abstract question" effect.
Risk of inflation / ceiling	Medium	High	Low	CSAT ceiling effect is its biggest practical flaw.

THE HONEST SUMMARY

If you can run only one metric, pick the one that matches your dominant business mechanic: NPS for loyalty-driven B2C and subscription, CSAT for high-volume operational touchpoints with rapid feedback loops, CES for service-heavy or support-heavy categories where friction is the primary cause of churn. If you can run two, pair a relational metric (NPS) with a transactional one (CSAT or CES).

Three scenarios — what each metric would tell you.

The matrix on the previous page is abstract. The scenarios below are how those tradeoffs feel in practice. In each, the metric you pick determines not just the number but the diagnosis — and therefore the fix.

SCENARIO 1 · SAAS SUPPORT TICKET

The moment: a customer has just had a billing issue resolved by support after two emails and a phone call.

NPS would say: "Would you recommend us?" feels jarring after a support call. Low response, noisy signal.

CSAT would say: 4/5 — "they fixed it, fine." Looks healthy.

CES would say: 3/7 — "two emails and a call for a billing question." Identifies the friction CSAT hid.

Best fit: CES.

SCENARIO 2 · HOTEL STAY

The moment: a guest has just checked out after a three-night stay.

NPS would say: "Would you recommend us?" feels natural — recommendation is exactly how hotels grow. Good signal.

CSAT would say: per-touchpoint CSAT (check-in, room, breakfast) gives operations granular feedback. Good complement.

CES would say: "Was it easy?" Wrong question — guests don't want frictionless, they want delightful.

Best fit: NPS + CSAT.

SCENARIO 3 · INSURANCE CLAIM

The moment: a customer has just finalised a claim after a car accident.

NPS would say: "Would you recommend us?" in a stressful moment — tone-deaf.

CSAT would say: "How satisfied were you?" — misleading; satisfaction is the wrong construct.

CES would say: "Was it easy to handle?" — precisely the question. Friction here is the entire customer experience.

Best fit: CES (with annual relational NPS for the wider relationship).

THE PATTERN

The right metric is the one whose underlying question would make sense to the customer in that exact moment. If the question feels weird, the data will be weird. This is the simplest filter and it eliminates two thirds of measurement design errors before they happen.

CHAPTER 6

Decision framework.

Four questions, one defensible answer.

6

How to pick the right metric — without religious wars.

Four questions, answered honestly, almost always narrow the choice to a single best metric or a clear pairing. In our experience across forty-plus client engagements, the first two questions dominate the answer in about 80 % of cases.

QUESTION 1 · WHAT ARE YOU TRYING TO PREDICT?

"What customer behaviour do you most need to forecast?"

Referrals / word-of-mouth growth → **NPS**

Churn / renewal in a service category → **CES**

Quality of a specific touchpoint → **CSAT**

QUESTION 2 · WHERE DOES YOUR RELATIONSHIP LIVE?

"What is the dominant moment of truth?"

Brand & emotion (luxury, hospitality, retail) → **NPS + CSAT**

Service delivery (telco, insurance, banking) → **CES + NPS**

Self-service (SaaS, fintech, digital) → **CES + CSAT**

One-off transactions (e-commerce, ticketing) → **CSAT**

QUESTION 3 · HOW OPERATIONAL IS YOUR PROGRAMME?

"Strategic indicator or daily-action signal?"

Strategic / board-level → **Relational NPS**

Operational / daily ops → **Transactional CSAT or CES**

Both → pair one relational + one transactional.

QUESTION 4 · HOW MATURE IS YOUR INFRASTRUCTURE?

"Can you actually act on what comes back?"

Just starting → one metric, transactional. Usually CSAT.

Scaling → add NPS for strategic view.

Mature → NPS + CES/CSAT + verbatim AI.

Typical landing patterns

COMPANY TYPE	RECOMMENDED PRIMARY	RECOMMENDED PAIRING
B2C e-commerce, mid-sized	CSAT (post-delivery, post-support)	+ relational NPS twice a year
SaaS, growth-stage	CES (per support ticket and onboarding step)	+ NPS quarterly, segmented by plan
B2B enterprise services	NPS (relational, twice a year)	+ CES on critical workflows
Telco / utility	CES (every customer-facing event)	+ relational NPS quarterly
Insurance / claims-driven	CES (claims journey)	+ annual NPS for the relationship
Hospitality / luxury	NPS (post-stay)	+ per-touchpoint CSAT

CHAPTER 7

Six pitfalls that wreck CX numbers.

7

The same six errors, in nearly every programme we audit.

The structural mistakes that kill CX measurement.

These are not edge cases. Across roughly forty client diagnostics, the same six errors show up repeatedly — and they sabotage even well-chosen metrics. If you only fix one thing after reading this paper, fix whichever of these six is currently breaking your programme.

01

Measuring the wrong moment.

The metric and the timing don't fit the construct. Asking NPS right after a support call, asking CES about a brand relationship, asking CSAT once a year. Match the metric to the question; match the question to the moment.

02

One metric for everything.

Companies adopt one metric (usually NPS) and then try to make it answer every question — loyalty, satisfaction, operational quality, churn risk. The result is a number that moves around for ten reasons and nobody can act on it.

03

No verbatim, no diagnosis.

The score moves; nobody can explain why; the next quarter the score moves again; nobody can explain that either. A CX programme without structured verbatim analysis is a dashboard, not a feedback loop. The "why" question matters more than the score.

04

Score-gaming and incentives.

Tie bonuses to NPS or CSAT and watch what happens. Surveys get rounded up, awkward customers get excluded, support agents nudge for "5 stars only" before closing the ticket. Incentivise the work that produces good scores, never the scores themselves.

05

Sampling bias hidden in plain sight.

Voluntary email surveys. 5 % response rate. Almost all responses are from passionate fans or angry detractors. The "average" of those is meaningless. Build sampling logic deliberately: stratified, event-triggered, multichannel — not just whoever volunteers to fill in an email.

06

Closing the loop — with no actual loop.

A detractor responds. Nothing happens. They get a generic thank-you email. A working loop has a named owner, an SLA, a recovery action, and visibility back to the executive who saw the original score. Without that, surveys are just polling.

THE HONEST TEST

For each of these six, write down "yes / no / partially" against your current programme. If three or more are "no" or "partially," you do not have a measurement problem — you have an operating-model problem. No new metric will fix it.

CHAPTER 8

Beyond the big three.

OES, ISI, CLV, churn — and why composite metrics now matter.

When NPS, CSAT and CES are not enough.

The big three were designed for one-question simplicity. That is also their limit. As AI and modern feedback platforms make richer analysis cheap, a new generation of **composite** metrics has emerged — combining multiple signals to answer the questions a single number cannot.

OES · OVERALL EXPERIENCE SCORE

An AI-derived score (0–10) that reads the *entire* survey response — numeric answers, multiple-choice picks, and verbatim — and produces a single experience indicator. Unlike NPS or CSAT, OES isn't tied to one question; it summarises the full feedback record.

Strength: picks up signal that single questions miss — for example, a 9/10 NPS paired with verbatim describing five problems is downgraded automatically.

Limit: only as good as the AI model behind it; needs a trusted platform.

OES is an InsightSofa-developed methodology.

ISI · INSIGHTSOFA SATISFACTION INDEX

A composite that combines NPS, CSAT and CES into a single 0–10 number, weighted by relevance for the moment. Designed so executive teams have one comparable indicator even when underlying programmes use different metrics for different touchpoints.

Strength: resolves the "we measure five different things, what's the headline?" problem at the board level.

Limit: like any composite, it can hide divergent signals — always read the components alongside it.

Adjacent metrics worth tracking

METRIC	WHAT IT MEASURES	WHEN IT MATTERS	WATCH OUT FOR
CLV	Customer lifetime value — expected revenue from a customer over the relationship.	Always. The financial counterweight to CX scores.	Estimation method differs wildly; agree on one.
Churn rate	% of customers who leave in a defined period.	Subscription, SaaS, B2B services.	Pair with CES — CES leads, churn lags.
Retention rate	% of customers retained period-over-period.	Mirror image of churn; often friendlier internally.	Beware definitional games — "active" hides a lot.
First Contact Resolution	% of issues resolved on the first interaction.	Support-heavy categories; correlates with CES.	Pressure to close tickets can fake it.
Sentiment score	AI-derived emotional tone of verbatim and transcripts.	Programmes with high verbatim volume.	Sentiment ≠ satisfaction; track both.

SELF-ASSESSMENT SCORECARD

Where is your measurement programme today?

Twenty statements across five dimensions of a working CX measurement programme. Score one point per statement you can honestly answer **true today** — not "aspirationally," not "we're working on it."

METRIC CHOICE (5 STATEMENTS)

- ☐ We have one primary metric per major touchpoint — not three.
- ☐ The choice of metric matches what we are trying to predict.
- ☐ We can name (in one sentence) the reason we picked each metric.
- ☐ We run at least one relational metric and at least one transactional.
- ☐ The metric definition (scale, formula, top-box) is written down and stable.

SURVEY DESIGN (5 STATEMENTS)

- ☐ Surveys fire on event triggers, not on a calendar schedule.
- ☐ The timing window between the moment and the ask is ≤ 24 hours.
- ☐ We use at least two channels (e.g. email + SMS or in-app).
- ☐ Each survey has a follow-up "why?" verbatim field.
- ☐ Response rates by channel and segment are tracked monthly.

ANALYSIS (4 STATEMENTS)

- ☐ Verbatim is categorised systematically — not read individually.
- ☐ Scores are reported by segment, not only as a headline number.
- ☐ Trends are reviewed against a stated benchmark or baseline.
- ☐ We can identify the top three drivers of detractors / dissatisfaction.

ACTION (3 STATEMENTS)

- ☐ Detractor responses trigger a defined recovery workflow with named owner.
- ☐ Each CX score has a named accountable executive.
- ☐ Operational changes are verified against the originating metric.

GOVERNANCE (3 STATEMENTS)

- ☐ Scores are reviewed in the executive cadence at least quarterly.
- ☐ CX metrics are referenced in cross-functional goals or OKRs.
- ☐ Score-gaming risks are explicitly designed out of incentive plans.

HOW TO READ YOUR SCORE

17–20: mature programme. Focus on composite metrics (OES, ISI) and verbatim AI. · **12–16:** solid foundation; focus on closing the loop and verbatim diagnostics. · **7–11:** the metric is fine; the operating model around it is the bottleneck. · **Below 7:** start with one metric, one touchpoint, one named owner. Do not redesign everything at once.

Pick the metric. Then earn it.

No metric is right by default. The right metric is the one that matches the question you are trying to answer, in the moment you are asking it, with the operational capacity to act on what comes back.

The companies that get CX measurement right are not the ones that picked the theoretically best metric — they are the ones that picked any reasonable metric, configured it carefully, built the operating cadence to act on it, and committed to one definition over time. Stability beats sophistication.

Three takeaways worth more than the rest of this paper:

01

Match the metric to the construct.

NPS for loyalty. CSAT for moment-quality. CES for friction. Anything else is decoration.

02

Pair relational with transactional.

One number for direction, one number for action. Programmes with only one of the two are always blind in one direction.

03

The score is not the work.

The work is the verbatim, the recovery, the operational change. The score is just the thermometer. Run a programme that reads the temperature *and* treats the patient.

How InsightSofa can help.

Beyond the platform itself, our CX strategy team runs a focused **measurement audit** with clients in the middle of rebuilding their metrics stack. The deliverables — metric choice, survey-design review, closing-the-loop map and a 90-day implementation plan — are the ones we use ourselves across roughly forty client engagements.

InsightSofa supports NPS, CSAT, CES, OES and ISI out of the box, with AI-powered verbatim analysis and a workflow engine that turns the score into action automatically — not just a chart on a dashboard.

The fastest start is a 30-minute working session with our CX strategy team.

+420 777 661 368 · insightsofa.com

This paper draws on the InsightSofa CX Strategy Team's experience across mid-sized and enterprise clients in B2B, B2C and the public sector. Industry benchmark ranges are indicative and should be validated against your own historical data.