Mystery Shopping & Other Outside Scoring

Tools For Operators · https://jeffreysummers.com/docs/mystery-shopping-other-outside-scoring/

A stranger comes into your building, eats a meal, fills out a form, and gets paid. A week later a report arrives with a score on it and a list of things that went wrong. You take that list into a pre-shift and correct your cast against it.

Nobody in that sequence read your operation. Not the shopper, who was following a form. Not the company, which sold you the form. Not you, because the report arrived instead.

What A Mystery Shop Actually Is

A mystery shop is one visit, by one person you did not choose, scored against criteria you did not write.

Take those three apart, because each one is fatal on its own.

One visit. A single sample of an operation that produces hundreds a week. Every shift has variance -- who was on, what the book did, what broke, who called in. A sample of one cannot separate a pattern from a Tuesday.

A person you did not choose. Not your Guest. Somebody performing an evaluation, which is a different activity than dining. They are watching for scoreable items instead of having an experience, which means the thing being measured is not the thing you produce.

Criteria you did not write. This is the whole game. The form came from the company that sold you the program, built to be usable across restaurants generally, which means it measures what is measurable everywhere. Greeting inside thirty seconds. Suggestive sell attempted. Restroom checked. Name used. Those are the items that survive being made portable, and they are not what your Guests are choosing you for.

The inference: **a trained observer scored the visit; therefore the score describes the operation.**

It Scores The Portable And Ignores The Load

Run the form's items against your own standard and look at what does not appear anywhere on it.

Whether the room felt like your room. Whether a Guest who wanted to be left alone got left alone. Whether the cast member read the table correctly and adjusted. Whether the thing your operation is actually chosen for happened. Whether a Guest who came in tense left better than she arrived.

None of that is on the form, because none of it can be scored the same way in a thousand buildings. What is on the form is the intersection of everywhere, which is the thinnest possible version of what your operation does.

So the report tells you how well your cast performed a generic script and tells you nothing about whether your standard held.

Six Ways It Damages The Operation

It manufactures repair work. The report is a list of defects. Your entire management attention for a week goes to returning to zero on items a stranger flagged. Nothing on that list moves the operation forward, and the time it consumes is the time your own read would have used.

It teaches your cast to perform for the wrong person. Cast members learn a shop could be at any table, so they run the scoreable behaviors on everybody. Name used, suggestive sell attempted, thirty-second greeting. Guests can feel it. What you have installed is a performance aimed at an evaluator, running in front of people who came in to eat.

It puts a stranger above your leads. Your lead on the stage watched forty tables that night. The shopper watched one. The report outranks the lead, and every person in your building learns what that means about whose read counts.

It corrects individuals on a sample of one. A cast member gets coached against a single anonymous visit, sometimes weeks after it happened, with no ability to say what was actually going on that night. That is not development, and your cast knows it.

It converts your standard into somebody else's checklist. Whatever the form measures becomes the operating standard over time, because that is what gets scored and scored is what gets managed. Your ruling quietly gets replaced by a vendor's item list.

It is the same form your competitor gets. He can buy the identical program this afternoon. Both operations get corrected toward the same generic script, and now the two of you are

converging on the same operation from opposite directions.

The Legitimate Thing Underneath It

There is a real need buried in the sale, and it is worth naming honestly.

An operator cannot be in the building for every shift, and cannot see his own operation the way a first-time Guest sees it. Both of those are true, and both are the actual problem the shop is sold against.

But the answer to cannot be everywhere is a cast that holds the standard without you, and the answer to cannot see it fresh is a read built for that, not a stranger with a form.

What Replaces It

Your own read, on a cadence, written down. Walk your own operation at the hours that matter, at the volume your book actually produces, watching against your own ruled standard. Nobody can sell you this and nobody can do it for you.

The cast's read, collected on purpose. Your lead and your kitchen manager know what went badly last night and why. That read is free, it is more accurate than any shop, and almost no operation has a habit for capturing it. Build the habit.

Named Guests, asked directly. Pick actual Guests -- regulars, and people on a first visit -- and ask them a specific question about a specific visit. Then write down what they said. Ten of those beat a thousand scored items because each one is attached to a person, a shift, and a real occasion.

A trusted outsider, briefed on your standard. If you want fresh eyes, use someone who has read your ruling and is looking at whether your standard held. Not a form. Your standard, in your words, and a conversation afterward.

Your own leading indicators. Return interval, request-by-name rate, recovery acceptance, cast retention. Those measure whether the operation is producing what you ruled, and no vendor owns them.

Video and your own eyes on the flow, not on the people. Where work backs up, where the room reads tense, where a station cannot keep pace. Structural findings, not individual defects.

What Changes Tomorrow

Take the most recent report you have and go through it line by line with one question against each item: does this measure something my operation is chosen for.

Most items will not. Cross them off. What survives is small, and it is the only part of that program that was ever worth what you paid.

Then write your own standard for one part of the operation -- the greet, the table read, the pace of a two-top on a Friday. Your words, your building, and specific enough that your lead could state it without looking.

Give it to your lead and ask him to read the stage against it for one week and tell you where it did not hold. That is a real read, produced by somebody with a name, who was in the building, who knows the cast, and who has to live with the answer.

Compare what that produces to what a hundred and forty dollars and a stranger's form produced. You will not buy another shop.

Digging Deeper

Every term in my framework lives in the Knowledge Base: kb.jeffreysummers.com

Terms used here: [Repair Work], [Repairman Syndrome], [Measurement Lock-In], [Cover Blindness], [Visibility Trap], [Operator's Visibility Problem], [Vendor Capture], [Default Architecture], [Designed Architecture], [Authority To Execute], [Safest Mediocre Execution], [People Constraint], [Product Constraint], [Constraint Architecture], [Transactional Cost-Plus]

More of my work on this, all of it in the Knowledge Base:

- 038 The Operator Who Manages From The Report -- the failure mode this instrument installs
- 4.ME.12 Transactional Instrumentation -- the category this belongs to
- 4.SH.11 Real-Time Coaching -- the replacement, done in the moment
- 2755 Call Your Critics -- going to the Guest directly
- 138 Something Changed In Your Room -- reading the building instead of the score

Discover more from Jeffrey Summers

Subscribe to get the latest posts sent to your email.