

Alva Masterclass: The Noise Audit

Your practical guide

Five strategies. Three sections each: what good looks like, your smallest first step, and the traps that quietly pull teams back.

How to use this

This guide pairs with the scorecard you also got in this email. Score yourself first, then come here for what to do about it.

Don't try to fix all five strategies at once. Pick one. Specifically: pick the one you scored lowest and the one that bites your team most often. They're not always the same.

For that one strategy, plan to take the first step forward and run it during the fall of 2026. The traps section is what to watch out for as you go.

If you only have ten minutes today, read the smallest first step for your weakest strategy and put a 30-minute slot in your calendar this week to start.

Guidelines

Define what good looks like

Weighted, role-specific criteria, written down before sourcing starts, with a concrete definition of "strong" for each. The job everyone's hiring for.

What good looks like

- An intake meeting before sourcing starts. Hiring manager and TA in the room together
- Around four to six attributes specific to this role, each with a weight
- Every attribute has a concrete "what strong looks like" description
- The same document reaches every assessor before they meet the candidate
- Weights are signed off in writing, not implied in conversation

Your smallest first step

Pick one upcoming role. Block 30 minutes for an intake meeting. Walk out with distinct and weighted attributes and a concrete definition of "strong" for each. Send the doc to every assessor before the process starts.

Common traps to avoid

Intake happened, the document never reached the assessors

Test it by asking each assessor what they're looking for. If you get four different answers, the doc didn't land.

Attributes like "culture fit"

These mean different things to different people. Push for behaviour-level definitions: "explains technical decisions to non-technical stakeholders without losing the point."

Weights agreed verbally

Memory shifts within a week. Write the weights down at intake and reference them at every debrief.

The same template across every role

Junior IC, senior IC, and team lead most likely need different attributes and different weights. Resist the urge to copy-paste.

Structuring

Break the decision down

Each attribute scored on its own scale before any overall view forms. The hire/no-hire call gets built from the parts.

What good looks like

- Each attribute scored independently on its own scale
- Scorecards submitted before the debrief, not in it
- A combination rule agreed in advance (weighted average, minimum threshold per attribute, etc.)
- The overall hire/no-hire view forms last, after the parts are in

Your smallest first step

Build an attribute scorecard for one role. Require every assessor to score each attribute separately. Look at the spread before the debrief. Start the meeting with where the scores disagree.

Common traps to avoid

Decorative scorecards

Filled in after the decision is made. The form gets completed because process says so, not because it shaped the call. Watch for scorecards that all read the same.

Halo/devil effect

One attribute drives the others. Look for assessors who score a candidate 5-5-5-5 or 1-1-1-1 across genuinely distinct attributes. Flag it, calibrate.

An "overall" field that overrides the attribute scores

Remove it. Or weigh it last, after the attribute scores are submitted.

Scores submitted in the debrief itself

Anchoring on the first voice happens within seconds. Submit separately, before the meeting opens.

Shared scale

Same number, same meaning

A scale where every point means the same thing to every assessor. Calibration is what gets you there.

What good looks like

- Ideally a 1-5 (or 1-7) scale with written behavioural anchors at each point.
- Every new assessor runs through one calibration session before they sit on a real panel.
- Periodic calibration on shared cases: same candidate, three assessors, compared scores.
- Drift between teams and roles gets surfaced and corrected, not ignored.

Your smallest first step

Take one role. Send three of your assessors the same recorded interview, role play or written case. Have them score it independently on your existing scale. Look at the spread. If it's wide, you've got your starting data.

Common traps to avoid

Anchors that just say "1 = weak, 5 = strong"

That's a label, and it gives assessors nothing to align on. Write what 1 looks like for this role. Write what 3 looks like. Write what 5 looks like.

New assessors onboarded by reading a document

Reading anchors doesn't calibrate the brain. The first calibration has to happen against a real case, with other assessors in the room.

Calibration meetings get cancelled when calendars are tight

These are the most important hours you'll run all quarter. Protect them. Treat them as non-negotiable, not as a nice-to-have.

Different teams running different scales without noticing

Tech, sales, and ops can quietly diverge. Audit once a quarter. Pull a sample of scorecards from each team, compare the distributions.

Sequencing

Protect each read from the next

Each stage stays clean of the next. Once an assessor reads someone else's score, they can't unread it.

What good looks like

- Each stage runs blind to upstream scores.
- A short, neutral probe list passes between stages: areas to explore, not opinions formed.
- Individual scores submitted before any debrief opens.
- The ATS or scorecard tool locks earlier scores from view until the next assessor has submitted theirs.

Your smallest first step

Pick one upcoming role. Tell every assessor: don't read the previous scorecard before your session. Check whether your tooling enforces it. If it doesn't, use a manual block (close the tab, no peeking) and hold each other to it.

Common traps to avoid

Slack messages between assessors before the panel

"She's strong, you'll love her" kills sequencing before the panel even starts. Have an explicit team rule: no candidate impressions in side channels until the debrief.

"I just want to peek at the last scorecard before I go in"

No. The whole point is that you don't peek. If the urge is strong, it's a sign sequencing matters here.

The hiring manager reads everyone's scorecards live in the debrief.

Lock them first. Share them only after each assessor has submitted their own. Otherwise the loudest scorecard sets the anchor.

Probe lists that contain opinions.

"Probe her communication, she was vague last time" is contamination dressed as a probe. Probes describe what to explore, not what to expect.

Aggregating

Combine independent reads

Independent scores first, group discussion second. The hard part isn't the math, it's protecting independence before the room opens its mouth.

What good looks like

- Scores submitted independently, 24 hours before the debrief where possible.
- The debrief opens with the spread, not with the consensus.
- A combination rule agreed in advance (mean, weighted average, threshold).
- When you only have one assessor, they explicitly argue against their own first impression before locking in.

Your smallest first step

For one role, require all scorecards 24 hours before the debrief. Print them or share them in the meeting. Open the conversation with: "Here's where we disagree." Run the meeting on the spread, not on the average.

Common traps to avoid

"Let's hear what everyone thinks before we score"

Flip the order. Score first, share second. Discussion before scoring contaminates the scores.

The most senior voice speaks first in the debrief

Inverts the seniority anchor. Have the most junior assessor share their score and reasoning first. The senior voice can react, not lead.

Smoothing disagreement to keep the room comfortable

Disagreement between competent assessors is the most useful signal you'll get. Stay with it, don't paper over it.

"I changed my score after the discussion."

Fine, but log the original. Patterns show up in the change deltas. If the same person always shifts toward the senior voice, you've found a structural problem worth fixing.

One last thing

Pick one move

You won't fix everything in a quarter. You don't need to. Pick one strategy, make one change, and let it bed in during the fall of 2026 before adding the next.

The five strategies build on each other. Guidelines feed into structuring. Structuring needs a shared scale. Shared scales need calibration, which needs sequencing, which needs aggregating. Wherever you start, the rest gets easier as you go.

Want a follow-up?

If you want to walk through your audit one-on-one, or talk through what to fix first, [click here to schedule a meeting](#), or send an email to anna@alvalabs.io.