Blind Human Call Grading

An add-on to Peer: the blind human read on the calls Peer grades — a baseline your leaders can defend, and the check on the AI that scores them.

Trained assessors grade whole sales calls blind — never seeing an AI's score, a report, or the person's name — against the standard you choose. Every grade carries the quote and the timestamp that earned it. Because one conversation in five is marked twice by assessors working apart, you also get a measured reliability figure, published every cycle.

What it is — and what it isn't

IT IS	IT ISN'T
A baseline by team and by competency. Hundreds of whole calls a quarter, graded to one standard, show where the gaps concentrate — which team, which stage of the call, which skill.	**A verdict on an individual from one call.** A single call surfaces roughly a third of the skills in the standard. Nobody should be ranked or coached off one human read, and the portal says so wherever a number could be misread.
The human check on AI grading. Peer grades every call; this is how you know whether to rely on it — criterion by criterion, with an agreement figure.	**A coaching platform, or real-time.** It runs after the call, on a sample. Coaching depth per rep comes from Peer, which grades every call.
Evidence that survives being questioned. A grade a revenue officer, HR or a client can challenge is answered with the moment, the rubric line and the assessor reference.	**A substitute for grading everything.** Human hours are finite; the sample is designed to be reliable at team level, not exhaustive.

How it runs

1. **Conversations come in.** The recorded calls Peer has graded, with Peer's scores alongside. Every person becomes a code; names, companies and pricing are stripped before an assessor sees a word. Transcripts are never retained.
2. **The panel marks the sample, blind.** Whole call, every competency in the standard, evidence quoted and timestamped. One in five is marked twice.
3. **The read is reconciled.** Where humans and AI agree, where they don't, and why — per criterion, per team.
4. **You get the portal, the statement and the recommendations.** Grade profiles per rep, evidence chips, clip playback where your recordings allow it, the human-versus-AI view, and a signed reliability statement for the cycle.

The standard you grade against

Your own scorecard or methodology; the **Peer Standard** — a universal sales-skills standard that does not change between companies; or **both, side by side**. Your rubric tells you whether scoring was applied consistently. The Peer Standard tells you where your people stand against everyone's. Only this service ever uses a customer's own scorecard; Peer itself always grades to the universal standard.

Options — frequency and volume

CYCLE	WHOLE CALLS GRADED PER CYCLE	DOUBLE-MARKED	BEST FOR	PRICE PER CYCLE
Standard	150	30	Teams up to ~100 reps — every rep reached about twice a year on a quarterly cadence	**£6,500**
Enterprise	300	60	250+ reps — about one call per rep per quarter; adds per-team reliability reporting	**£11,500**

CYCLE	WHOLE CALLS GRADED PER CYCLE	DOUBLE-MARKED	BEST FOR	PRICE PER CYCLE
Baseline (one-off)	3 whole calls per rep for a chosen cohort — minimum 150 calls	1 in 5	The first read, before quarterly cycles begin; the only design that supports a per-rep starting point	**£40 per call** (60 reps ≈ £7,200 · 250 reps ≈ £30,000)

Frequency. Quarterly is the default and what the reliability statement is built around. Twice-yearly and annual cadences are available for a lighter programme; a one-off Baseline can precede any of them.

Setup. Calibration setup — the rubric mapping, the anonymisation pipeline and marker training — is **included** for Peer customers. Blind Human Call Grading is available to Peer customers only.

Sizing guide

REPS	WHAT A QUARTERLY STANDARD CYCLE GIVES YOU	WHAT WE RECOMMEND
Up to 100	~1.5 calls per rep per quarter — a genuine team baseline	Standard, quarterly
100–250	0.6–1.5 per rep per quarter — team-level only	Standard + a one-off Baseline, or Enterprise
250+	Under 0.6 per rep — too thin to call a baseline	Enterprise (≈1 per rep per quarter), with a Baseline for any cohort you need read in depth

With Peer

Peer grades every call, every moment, against the same standard — around twenty calls per rep a month at the cadence we observe — and serves that record to the assistant your leaders already use. Blind Human Call Grading is the blind human read of it: the machine covers everything, the humans check it, and the baseline holds up when someone asks "based on what?". **Peer + four Enterprise cycles a year** is the package for a 250-rep organisation.

Data and privacy

Assessors see codes and anonymised transcripts, never names. Transcripts are processed and discarded, not stored. Recordings stay wherever you keep them; clip playback in the portal links to the moment where your host supports it, and falls back to quote and timestamp where it doesn't.

PeerLab · www.peerlab.ai/assurance · greg@peerlab.ai