RED BEAR: ENTERPRISE CULTURE EVOLUTION · DATA SHEET · G · MEASURING THE PROGRAMME

Did anything move?

Nobody can say whether the programme moved anything, so it is cut. Six measures, each with a *Never* clause, read on the pilot against its comparison. Usage is not one of them.

KEY POINTS

- Six measures: dial movement, toxic pattern decline, lesson rate, lesson application, signal trust and speak-up balance, each printed with the clause that says what it may never become
- The programme effect is a difference between pilot and comparison units with a band; a difference above the prior of one to two points carries a *measurement review* flag until the sponsor records the review
- Sessions, satisfaction and return rate appear only on the pilot report, labelled diagnostic, not an outcome; nothing is computed per manager
- A dial used as a target for anything outside the programme is flagged, its band widened, and a Campbell's-law review opened for the sponsor and the owner

What it is

Programme effect shows the sponsor the six measures of the Red Bear Culture Framework for the programme window, each with its *Never* clause beside it, the number of dials, units or lessons in the set, and the confidence of the readings underneath. It shows each measure for the pilot units against the comparison units, as a difference with a band, from the pilot the sponsor designed before the programme started, and the point at which the sponsor may decide to extend. No mean position across dials is shown; a level is never shown without its trend; and a measure whose set is below its floor shows *suppressed*.

The Calibration pulse is the instrument the other signal families are validated against: the seven Edmondson psychological-safety items and the DORA Westrum items, issued quarterly to the units in scope, aggregated above the floor. The item bank is locked, and a request to add a licensed item is refused with the licensing rule stated.

The pain point

Vendor-reported results in culture technology are usage and satisfaction. Valence and Delta reported sessions; IBM Watson Talent reported claims; and the one senior-leader trial with a control group found the AI arm beat control on nothing, with engagement and effect uncorrelated. A programme judged on usage is judged on nothing, and a programme without a comparison is cut at the first sponsor change.

The other failure is the opposite one. The more a social indicator is used for decisions, the more it is corrupted: Wells Fargo's cross-sell scorecard is the terminal case, and a psychological-safety score published per team changes the next survey's answers. A Culture Dial is designed to be read, not hit. So the product measures outcomes and never usage, requires a comparison before it reports an effect, treats a claimed effect above the honest prior as a measurement problem first, and flags any dial that becomes a bonus criterion, an objective or a league position for a review that asks whether the lever, not the reading, should carry the incentive.

How it works

The six measures. Dial movement counts dials that moved towards target, as a share with the number of dials. Toxic pattern decline reads coded prevalence by family at window start and end, corroborated by the calibration family. Lesson rate and lesson application are the coaching measures, in aggregate; lesson application is never shown with fewer than ten lessons. Signal trust is participation in the pulse and in story submission, never a target for managers to drive. Speak-up balance pairs case volume with the speak-up items, because volume alone is ambiguous (FR-MEA-01).

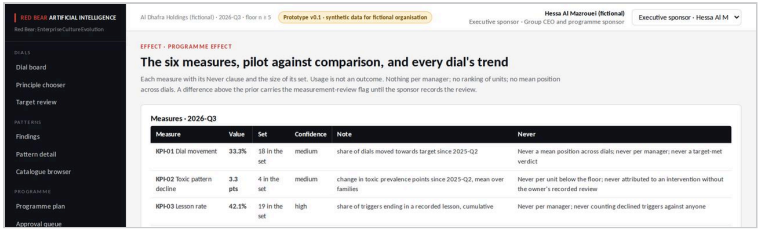
Effect over time. The pilot-against-comparison difference is shown with its band over the programme window; a difference above the prior is flagged for measurement review until the sponsor records it; the point at which extension may be decided is marked (FR-MEA-02). For every dial, the trend is shown as direction and rate with its band, coarsened so that every period clears the floor, with the owner's annotations and adopted elements on the timeline (FR-MEA-03).

Honesty rules. No screen presents session count, duration, stories per manager, driven participation, satisfaction or return rate as a measure; where any appears it is labelled a diagnostic (FR-MEA-06). Every measure is computed across groups above the floor, none per manager, and a breakdown by unit is an unordered set with each unit's confidence and group size (FR-MEA-07). When a dial is recorded as a target for anything outside the programme, the product flags it *used as a target* on every view, widens its band and opens the four-question review: which families are cheap to move, whether calibration still corroborates, whether participation has been driven, whether the lever should carry the incentive instead; the marker stays until the outcome is recorded as a decision (FR-MEA-04).

What you get

Capability	What it does	Requirements
The six measures	Each with its *Never* clause, set size and confidence; suppressed below the floor; no mean position	FR-MEA-01
Programme effect over time	Pilot against comparison as a difference with a band; measurement review above the prior; diagnostics labelled	FR-MEA-02
Dial trend	Direction and rate with a band, coarsened above the floor, annotated; never a level alone	FR-MEA-03
Campbell's-law review	*Used as a target* marker, widened band, the four questions, an outcome recorded as a decision	FR-MEA-04
Calibration pulse	Open Edmondson and DORA Westrum items, quarterly, above the floor; locked bank; licensed items refused	FR-MEA-05
Usage is not an outcome	No usage, satisfaction or return figure as a measure; participation never a manager's target	FR-MEA-06
Never per manager	Computed across groups above the floor; no breakdown by manager, no ranking of units	FR-MEA-07

Screens



S-22 Programme effect: Programme effect: the six measures with their *Never* clauses, the pilot against its comparison as a difference with a band, and the point at which the sponsor may decide to extend. Prototype render on synthetic data for fictional entities.

RETURN ON INVESTMENT

This pillar is the return on investment: it is where the programme's claim is made and where it is limited. All six measures (ECE-CF-01 KPI-01 to KPI-06) are baselined before the pilot, read against a matched or stepped-wedge comparison, and carry no target in this document; the honest prior on any effect is one to two points on a dial's scale, because the published effect of nudges is several times the effect that ships. What measurement makes possible is a pilot report a sponsor can defend to a successor and to a board; a decision to extend that is taken on recorded evidence; a portfolio that sheds what did not move a dial; and a dial that stays honest because the product notices when someone starts paying for it.

Platform engines

Measurement consumes the Red Bear Platform's Workplace Culture Evolution engine for programme measures, no ranking, the pilot comparison and the Campbell's-law flag (RBP FR-WE-08, FR-WE-10, FR-WE-14, FR-WE-15), the Workplace Culture Analysis engine for the calibration pulse (RBP FR-WA-14), the Organisation Information Gathering engine for issuing it (RBP FR-OG-05), the Decision Feedback Coaching engine for aggregate coaching measures (RBP FR-DC-05) and the Privacy and aggregation service (RBP FR-PA-06). Platform behaviour is specified in the platform functional specification, not here.