# EVALUATION

HOW DO WE DEMONSTRATE THAT THE WORK WE DO TOGETHER DELIVERS VALUE FOR YOU?

Our approach to evaluation, based on the Kirkpatrick Method[1], involves engagement with client sponsors, client participants, and client stakeholder to achieve:

- Agreement around evaluation criteria, data sources, and timelines.
- Appreciation of the strengths and weakness of the Kirkpatrick Method.
- Acceptance of our respective responsibilities.

Applying the Kirkpatrick Method to evaluate team development programmes

The Kirkpatrick Method provides a structured, multi-level framework for evaluating the effectiveness of a team development programme. Its strength lies in linking immediate learning experiences to longer-term behavioural change and organisational impact. This framework is captured as Level 0 – Level 4 below.
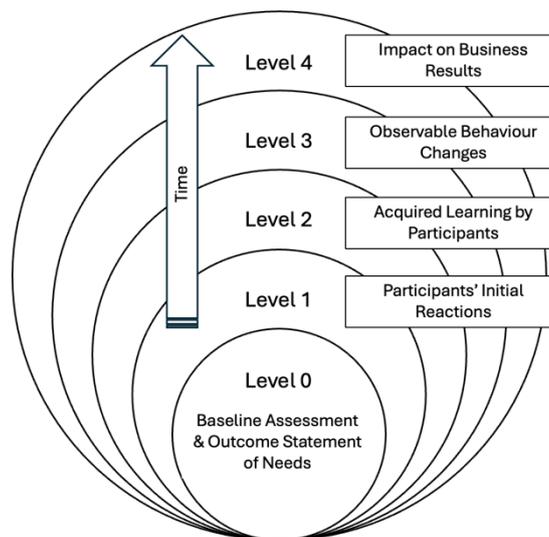
Level 0: Baseline Assessment & Outcome Statement

Establishing a credible shared baseline assessment and statement of desired outcomes is critical if actual outcomes are to be evaluated in a way that stakeholders regard as fair and useful.

---

[1] Kirkpatrick, D. L. (1994). *Evaluating Training Programs: The Four Levels*. Berrett-Koehler Publishers, San Francisco, CA, USA.

In contexts where performance is partly subjective and multi-causal, the goal is not perfect precision but a shared reference point against which change can be plausibly assessed.

Level 0 steps include:

- Clarify purpose, scope, and stakeholders i.e. who wants to see what changes where?
- Define observable indicators in behavioural and/or performance terms.
- Determine capture method i.e. questionnaire and/or interview.
- Use multiple data sources e.g. team members, key stakeholders, any objective data.
- Apply consistent rating scales e.g. a 1- 4 scale such as rarely/ sometimes/ usually/ consistently applied to the key behaviour and/or performance attributes in scope.
- Capture narratives so as to note any constraints or external factors impacting performance.
- Validate with team members and key stakeholders.
- Treat as a point of comparison rather than a scientifically controlled benchmark.

Level 1: Reaction

This gathers participants' experience via end-of-session "hot debriefs" and post-session surveys after any significant programme element such as a workshop. While Level 1 does not measure outcomes, it gives participants a voice and demonstrates our readiness to fine-tune the programme as it unfolds.

Level 2: Learning

This Level evaluates the extent to which team members acquired the intended knowledge, skills, or attitudes. This can be included in the "hot debrief" and post-session surveys mentioned above, and also the follow-up holding-to-account sessions, 4 – 6 weeks after a major workshop.

Level 3: Behaviour

This Level examines whether learning translates into observable changes in how the team operates in the workplace. This level is especially important as it demonstrates practical application rather than theoretical gain.

We will collect evidence against the baseline assessment as agreed in Level 0 through colleague and stakeholder feedback. This will be done via questionnaires and/or interviews at an agreed point after completion of the programme, typically 3 – 6 months.

Level 4: Results

This Level assesses the broader impact on business outcomes as per the success metrics established in the Level 0 Outcome Statement at the outset. This level enables stakeholders to make an assessment of return on investment. This will be done via questionnaires and/or interviews at an agreed point after completion of the programme, typically after 6 - 12 months.

CAVEATS

While it cannot eliminate subjectivity or prove causation, the Kirkpatrick Method provides a useful foundation against which changes in performance can be assessed and discussed with stakeholders in a logical and transparent way. That said, it is worth emphasising the following issues with this approach.

Difficulty of causal attribution

A key limitation is the challenge of establishing a clear 1-to-1 cause–effect relationship between the team development intervention and observed results. Team performance and behaviour are influenced by many variables - leadership changes, workload, organisational culture, external pressures etc. - making it difficult to isolate the impact of the development process alone. This is particularly problematic at Levels 3 (Behaviours) and 4 (Business Results).

Subjectivity of behavioural and results data

Many improvements in team development are perceived rather than objectively measurable. For example, assessments of improved trust often rely on self-reporting or managerial judgement. Both are subject to bias, differing perspectives, and contextual interpretation. As a result, findings may be contested by stakeholders seeking more definitive evidence.

Risk of over-simplification

The linear structure of the Kirkpatrick Method can imply that learning naturally leads to behaviour change and business results. This is not often the case in complex team environments. Without acknowledging intervening factors such as leadership support or organisational constraints, the model may oversimplify how change actually occurs.

Stakeholder expectations of "proof".

When stakeholders expect definitive, metric-based proof of impact, the Kirkpatrick Method may fall short. The model supports contribution rather than proof of causation, which can create tension when investment decisions depend on clear accountability.