INTERNAL COACHING DISCUSSION FORUM

Discussion Paper 05

How Should We Measure ROI in Organisational Coaching?

A CIPD study found only 36% of organisations evaluating their coaching. In the cases where evaluation has persuaded sponsors, the evidence has seldom been a return-on-investment calculation.

Tom Battye

EMCC Global Master Practitioner | Accredited Supervisor | EMCC Global Supervision Award Winner

October 2026

tombattye.co.uk

In June 2003, the head of training and recruitment at a major UK financial services provider ran a small experiment. Twenty-four financial planning managers were followed for six months, with each manager's sales in the three months before a selling skills course compared with the three months after it. Half attended the course, of whom six were then given a coach for three months (the case does not say whether the coaches were internal or external); the rest formed a control group. Sales rose by 4% in the control group, by 8% among those trained but not coached, and by 27% among those who had training followed by coaching (Carter et al., 2005, as cited in St John-Brooks, 2014).

The company did not calculate a return on investment, because the rise in sales so plainly exceeded the cost of the training and coaching that a calculation would have added nothing. Nor did it try to prove that the coaching caused the rise, which groups of six, six and twelve people could not have shown with any rigour. The comparison was proof enough for the people who had to decide. St John-Brooks (2014) treats the case as an illustration of the gap between practitioner and academic research.

Most of the questions about measuring coaching are present in that small study. Sponsors ask for a return on investment. In the cases St John-Brooks and Hawkins report, they were more often persuaded by a comparison, or by a measure chosen to match the reason the coaching was bought.

Why so few organisations evaluate

Research by the CIPD (2010, as cited in St John-Brooks & Duncan, 2026) found that only 36% of the organisations in its study were evaluating their coaching. The figure is sixteen years old, but St John-Brooks and Duncan (2026) still describe most organisations as making no assessment of their coaching service. Where an assessment happens, it generally consists of feedback from the client (McDowall & Lai, 2024, as cited in St John-Brooks & Duncan, 2026).

Lead coaches give practical reasons for not evaluating. One organisation in Snape's (2012, as cited in St John-Brooks & Duncan, 2026) research had decided not to get too involved in measuring effectiveness because it was so difficult; it preferred "to spend the money on making our coaches better". Those running a small pool often hear from clients and their line managers that the coaching is working, and feel no need for quantitative data to tell them what they "already know" (St John-Brooks & Duncan, 2026).

Trusting that feedback works until a champion moves on or a budget review demands cuts. The lead coach then has to show value to someone who has never been coached. Hawkins (2012) quotes the Foreign and Commonwealth Office's coaching strategy: "Coaching is an investment, not a budget item." The sentence holds only while someone in authority accepts it. A new finance director may not.

Outputs and outcomes

Hawkins (2012) distinguishes outputs, which show up in the person coached, from outcomes, which follow in the organisation. The first include progress on objectives, new behaviour and a satisfied client; the second include performance, retention and the experience of customers and staff. Most coaching evaluation stops at outputs.

Hawkins illustrates the distinction with car retail. For many years, dealerships put considerable effort into surveying customer satisfaction and competing on it. Later research found very little correlation between satisfaction soon after purchase and whether the customer bought their next car from the same dealer. Dealers began to build an evaluation system that could support the creation of value only once they started measuring customer retention, which is an outcome.

In coaching, a client who reports progress on their goals has produced an output; whether that person stays, performs better or leads a stronger team is an outcome. Hawkins also warns against measuring short-term shifts in behaviour alongside long-term financial performance and then claiming, without testing the link, that one caused the other.

What an ROI figure contains

St John-Brooks and Duncan (2026) use the Kirkpatrick-Phillips model, in which Phillips added *return on investment* as a fifth level to Kirkpatrick's four: *reaction*, *learning*, *behaviour* and *results*. The standard calculation subtracts the cost of the coaching from the estimated value of its benefits, then expresses the difference as a percentage of the outlay.

Hawkins (2012) sets out a worked example using the Sherpa methodology. The benefits of an assignment are costed first: a rise in team productivity valued at £24,000 and a reduction in staff turnover valued at £36,000. The coachee then judges how much of each improvement the coaching produced, in this case half of the productivity gain and a fifth of the reduced turnover, which brings the attributed benefit to £19,200. That figure is discounted by the degree of confidence in the judgement, here 80%, to give £15,360. Subtracting the £5,000 cost of the assignment leaves a net benefit of £10,360, a return of 207%.

Every number in that chain apart from the cost is an estimate. The result looks precise only because the method multiplies them together. St John-Brooks and Duncan (2026) describe a senior diplomat, coached on relationship-building and negotiation, who helps form an alliance with long-term political and economic benefits and credits thirty per cent of the result to the coaching. "But thirty per cent of what?" they ask.

Hawkins (2012) summarises the studies behind the headline figures for coaching as claiming returns of between 600 and 700 per cent. The two most widely quoted rest on participants' own estimates of value. McGovern et al. (2001) surveyed 100 coached executives, of whom the 43 willing to put a value on the benefits produced an average return of 5.7 times the investment. Anderson (2001) reported a return of 529 per cent at a single Fortune 500 company, from 30 returned surveys; Parker-Wilkins (2006) reported 689 per cent at Booz Allen Hamilton. All three concern executive coaching and say little about the value of an internal coaching pool. Few HR professionals attempt the calculation. In a 2010 survey by Sherpa Executive Coaching that Hawkins cites, only 18% of more than 200 HR professionals calculated a return on their coaching spend.

The calculation carries practical hazards as well. A negative return is common in the first year of a programme, because the outlay falls in that year and the benefits arrive later (Soulsby, 2012, as cited in St John-Brooks & Duncan, 2026). Trained coaches without clients add cost with no matching benefit. On one coach management platform in 2025, 35% of the 1,763 internal coaches registered, from around 83 organisations, had no current client (St John-Brooks & Duncan, 2026). One lead coach advises tracking every expense from the first day, whichever budget it falls in, so that the figures exist when a future sponsor asks for a calculation (St John-Brooks & Duncan, 2026).

Returns particular to internal coaching

Internal coaching services have a simpler financial argument available, one that avoids estimating benefits altogether. In 2012 the BBC calculated that its internal coaching cost £50 an hour, after counting two full-time staff to manage the programme, in-house training, CPD and accreditation, though not the time the coaches gave (Macann, 2012, as cited in St John-Brooks & Duncan, 2026). A snapshot survey in 2021 put average external executive coaching rates in the UK at around £265 an hour (Passmore, 2021, as cited in St John-Brooks & Duncan, 2026). The comparison shows what the same coaching would have cost if bought in, which answers a sponsor who asks whether an internal service is cheaper. A sponsor who asks whether the sessions did any good learns nothing from it.

A second return falls to the coaches themselves. A government agency of 1,400 staff launched an internal coaching service in 2020; by 2023 it had settled at around 20 active coaches, trained to ILM Level 5. In 2024 the lead coach looked back over three years and found that over half the coaches had been promoted; they said that being a coach had made them more confident at interview. Nearly a quarter of the clients had also been promoted (St John-Brooks & Duncan, 2026). No comparison was made with the agency's usual promotion rate, which leaves the coaches' own account as the only evidence that coaching played a part. Even so, if internal coaches gain as much as their clients, an evaluation that counts only the clients may miss much of what the service produces.

Comparison with data the organisation already holds

The financial services study convinced the company because it compared like with like. An organisation can make the same kind of comparison at scale, and at little extra cost, using data it already collects.

St John-Brooks and Duncan (2026) describe a large accountancy firm that had trained 150 internal coaches over ten years. In 2022 it compared HR records for 400 employees who had completed coaching over a three-year period with those for 4,600 who had not. Coached employees were 40% more likely to achieve a higher performance rating and just over 50% more likely to be promoted. Retention improved marginally. Because the data already sat in the firm's HR systems, nobody had to design a survey or ask a coachee to estimate anything.

The firm did not control for who was offered coaching or who chose to take it up, which leaves the comparison open to an obvious objection. The people selected may well have been those already most likely to be promoted, in which case the gap could reflect the selection more than the coaching.

A multinational telecoms company described in the same chapter took more care over selection, matching 61 coached senior leaders with 56 managers at the same level, in the same country, with at least two years' service. On the engagement survey's measure of "manager action-taking", the coached leaders' teams reported a 9% rise, while the comparison group's reported a 9% fall. The gap was statistically significant (St John-Brooks & Duncan, 2026).

Measuring against purpose

Coaching bought to solve a specific problem comes with its own measure. In 2008, a medium-sized law firm in the City of London was steadily losing good young lawyers to larger firms (St John-Brooks, 2014). Exit interviews showed that money was not the main driver. Two to three years after qualifying, many began to feel seriously undervalued: long hours, partnership still far off, little feedback, no sense of how they were doing. The managing partner had three people in the L&D department trained as internal coaches to hold career development interviews with every lawyer between one and eight years qualified. Monthly monitoring of retention soon showed that the defections had dried up.

At the same time, the firm introduced an "associate partner" designation to give its junior lawyers something nearer to aim for. Nobody can say how much of the improvement belonged to the coaching and how much to the new title. The firm had no need to know, because it had bought a combined response to a retention problem and judged it by whether people stayed.

Southern Railway tied its measures to business objectives across a whole organisation. A leadership course for 250 managers led to coaching programmes and, in 2008, an internal coaching service (Hawkins, 2012). The company's steering group tracked a chain of linked measures: staff engagement in the staff survey; turnover, absence, grievances and disciplinary hearings; service performance and reliability; and customer satisfaction. Over the period, staff turnover fell by 30 per cent, while absence, grievances and disciplinary hearings declined and the network recorded its best ever scores for service, reliability and customer satisfaction. The managing director claimed only that "an element" of these results was "definitely attributable" to the change in leadership and management style, with the coaching programme "a big contributory factor". Hawkins records that establishing the chain was essential to keeping the support of senior executives. Like the law firm, Southern Railway could not isolate the effect of coaching; both organisations tied it to outcomes they were already tracking.

The case against ROI

Grant (2012, as cited in St John-Brooks & Duncan, 2026) treats ROI as an unreliable measure of coaching outcomes and argues that coaching does itself a disservice by concentrating on financial results. He proposes wellbeing and engagement instead, on the grounds that "organisations function better with mentally healthy employees who are engaged in their work activities". Validated instruments exist for both, among them the short Warwick-Edinburgh Mental Wellbeing Scale and the Utrecht Work Engagement Scale; most are free to use, sometimes with permission, and can be given before and after coaching (St John-Brooks & Duncan, 2026).

Measuring wellbeing produced one of the more rigorous evaluations that St John-Brooks and Duncan (2026) report. The Institute for Employment Studies evaluated a virtual coaching service for primary care staff in England, although the source does not say whether its coaches were internal. The evaluation tracked 261 people who were coached and 195 who were not, measuring the same individuals' wellbeing and resilience before coaching, immediately after it and at a later follow-up. Both measures rose significantly in the coached group and fell in the other. The more a person's wellbeing increased, the less likely they were to intend to leave. The results formed part of a successful case for continued funding.

Hawkins (2012) also records the view that the value of a coaching culture is self-evident. Sherman and Freas (2004, as cited in Hawkins, 2012) argue that the result "may not be directly measurable in dollars", while every organisation benefits from more candour and less denial. Hawkins anticipates the reply that moral goodness is not enough of a foundation to sustain such an important enterprise.

Measuring wellbeing does not settle the sponsor's question; it asks the sponsor to accept a different unit of value. A board that already sees engagement as a driver of

performance will probably accept an engagement measure, while one that does not will ask again what the coaching returned.

The measure and the offer

Each of the evaluations that persuaded sponsors used a measure tied to something the organisation already counted or cared about. The law firm counted departures. The primary care service measured wellbeing, the thing its coaching was introduced to support, while the accountancy firm drew on ratings and promotions already held in its HR records.

Matching the measure to the reason for the coaching works when there is one reason. Many internal coaching services offer something broader: coaching for anyone who asks, on whatever the client brings. The outcomes of an open offer vary as widely as its clients, so no single measure is likely to capture them. A sponsor is entitled to ask what such a service returns. The answers available are an aggregate of estimates, a cost comparison that says nothing about benefit, or a measure that covers one part of what the coaching does.

Questions for discussion

1. If your coaching is open to anyone, for any purpose, is there a measure that could capture what it does? Or does coaching that can be evaluated need a narrower purpose?

2. What does your organisation measure about its coaching now? In Hawkins' terms, are those measures outputs or outcomes?

3. The financial services company accepted a comparison between groups of six, six and twelve people as proof. What standard of evidence would your sponsors accept, and who decides?

4. The law firm could not separate the effect of coaching from the effect of a new job title, and did not need to. When does it matter whether an improvement can be attributed to the coaching?

5. Grant argues for wellbeing and engagement in place of financial return. Would that argument carry weight with your senior leadership team?

6. Is the saving against external coaching rates a return, or a way of avoiding the question?

References

Anderson, M. C. (2001). *Executive briefing: Case study on the return on investment of executive coaching*. MetrixGlobal.

Hawkins, P. (2012). *Creating a coaching culture*. Open University Press.

McGovern, J., Lindemann, M., Vergara, M., Murphy, S., Barker, L., & Warrenfeltz, R. (2001). Maximizing the impact of executive coaching: Behavioral change, organizational outcomes, and return on investment. *The Manchester Review, 6*(1), 1–9.

Parker-Wilkins, V. (2006). Business impact of executive coaching: Demonstrating monetary value. *Industrial and Commercial Training, 38*(3), 122–127. https://doi.org/10.1108/00197850610659373

St John-Brooks, K. (2014). *Internal coaching: The inside story*. Karnac Books.

St John-Brooks, K., & Duncan, J. (2026). *Internal coaching: The inside story* (2nd ed.). Routledge. https://doi.org/10.4324/9781003519911

About the author

Tom Battye is an EMCC Global Master Practitioner, Accredited Supervisor, and EMCC Global Supervision Award winner. He has spent 25 years training internal coaches across the public and private sectors. He is the founder of Tom Battye Coaching and the creator of CoachInsights, an AI-powered coaching assessment platform. His programmes include the Certificate in Executive Coaching (EMCC Practitioner, AC accredited) and the Leader as Coach programme (EMCC Foundation, AC accredited).

tombattye.co.uk | tom@tombattye.com

About this series

This paper is part of the Internal Coaching Discussion Forum, a six-session discussion series for L&D, OD, and talent leaders running from May to November 2026. Each session explores a question about what it takes to create a coaching culture, drawing on published research, named case studies, and 25 years of practical experience. Format: 20 minutes of content, 40 minutes of discussion, via Zoom. Details and registration at tombattye.co.uk/forum.