Academic Paper

Going full circle: Advancing coaching efficacy using 360° survey metrics – A case study

Hilton Rudnick ✉ (University of Johannesburg, South Africa)
Neville Goldin

Abstract

The ubiquitous proliferation of coaching requires practical mechanisms to measure efficacy at an individual intervention level. While several meta-analyses have generally demonstrated positive outcomes, these speak to the overall profession rather than individual interventions. This paper discusses 360° reviews as a viable method to measure coaching efficacy in the organisation. The research includes a single-subject case study applying pre and post 360° reviews and comparing the results using a paired samples test rating. The results indicate visible behavioural shifts in the post and pre 360° rater categories. Although 360° reviews might add to logistical complexity, the benefits outweigh the overhead.

Introduction

The need to facilitate workplace learning, leadership development and employee performance enhancement has led to the continuous proliferation of workplace coaching (De Haan & Nilsson, 2023; Jarosz, 2023). Yet, due to its nature, coaching stays a private and confidential intervention. Though there are varying degrees of disclosure to the funding organisation, the process remains predominantly insulated from its stakeholders. That is one of the reasons why a routine mechanism to measure the return on investment (ROI) has become increasingly necessary. However, as De Haan and Nilsson suggest, "Despite a growing base of empirical research, the effectiveness of workplace coaching has not been established beyond doubt" (2023, p. 1). Jarosz (2023) suggests 360° reviews as one of the ways to measure aspects of coaching effectiveness, referring to a study on integrated leadership development programmes by Coates (2013). While not discounted, utilising 360° reviews to track coaching progress still seems rare. It would serve the industry well if more objective tools to measure the efficacy of individual coaching interventions were accessible

and applicable. This paper presents 360° reviews as a quantitative approach to tracking coaching efficacy.

The coaching domain is lean on performance metrics to the extent that Cannon-Bowers, Bowers, Carlson, Doherty, Evans & Hall (2023, p. 1) argue that "a true science of coaching is yet to develop." It becomes even more pronounced at an intervention level (typically 6 to 10 sessions) where coaching practitioners have limited tools to inform the efficacy of their efforts. To this end, this paper uses a single-case study to demonstrate how 360° reviews and coaching can be combined to yield important measurement results for enlightening coaches, coachees and the sponsoring organisations and assist in evolving the coach's practice. Single-case studies have their detractors (Bonaiuto, De Gregorio, Sarrecchia & Gentile, 2008), mainly because of the lack of generalisability of results. However, single-case studies have been indicated to have a positive place in research and practice (Nock, Michel & Photos, 2007). For this paper, the single-case study was applied as a method of measurement principle and not to generalise statistical findings.

Literature review

The International Coaching Federation (ICF), only one of the organisations providing accreditation, indicates a membership of around 61,000 accredited coaches worldwide (ICF, 2023). Faire (2023) confirms the substantial proliferation of coaching, while De Haan, Duckworth, Birch and Jones (2013) allude to the "plethora of models and approaches" (p. 2), implying that professionals practise in vastly different ways according to their particular epistemologies. Though it seems that there are many coaches and many ways in which to coach, interestingly, the evidence indicates that the particular approach to coaching is far less important than the elements of coaching prevalent in all approaches (De Haan et al., 2013; Molyn, De Haan, van der Veen & Gray, 2022). Their studies confirm that the coachee-coach relationship (working alliance) and the coachee's self-efficacy expectations are the most pivotal in predicting successful outcomes.

Coaching and coaching efficacy

Coaching in this paper refers to coaching performed within the workplace context, confirming Grant's (2017) definition of coaching as "the use of coaching methodologies in the workplace for the purpose of helping employees, managers and leaders attain work-related goals in terms of skills, performance or developmental outcomes."(p. 3). It can be done at most role levels and involves one-on-one sessions with a trained professional (Jones, Woods & Guillaume, 2016). Coaching is widely used in leadership development (Anthony, 2017) in a range of industries such as healthcare (Henochowicz & Hetherington, 2016), education (Gallagher, Grierson & Susin, 2023) and the military (Greene, 2019).

The behaviours and competencies of leadership have an immense impact on performance within organisations particularly in maintaining positive team engagement while driving results (Herbison, 2023). Companies are, therefore, often intentional about helping groom staff members for leadership roles through development programmes or coaching. At the same time, the requirements to measure the effectiveness of such programmes more objectively have been well documented (Claudy, Lamera & Underhill, 2013; Phillips & Phillips, 2004; Stawiski, & McCauley, 2023; Tooth, Nielsen & Armstrong, 2013). Yet, determining effectiveness is challenging, while empirical research about measuring the efficacy of coaching has lagged behind the practice (De Haan et al., 2013; Pavur, 2013). De Meuse, Dai and Lee (2009) performed a meta-analysis on six of the twelve identified studies that met their inclusion criteria. They concluded that more rigorous empirical research is required to quantify coaching effectiveness and build better instruments.

The construct of 'coaching effectiveness' refers to achieving coaching goals and, according to Kauffman and Hodgetts (2016, p. 163), "helps individuals grow to higher stages of cognitive and

emotional complexity." A meta-analysis by Wang, Lai, Xu and McDowell (2021) on methods to test coaching effectiveness found only 20 studies that met their criteria for inclusion (which comprised using psychologically informed approaches). In another analysis, Burt and Talati (2017) only considered effectiveness studies that included control groups, of which only 11 met the criteria. De Haan and Nilsson's (2023) meta-analysis included only studies containing randomised controls and located 37 of these meeting their strict criteria. While the usefulness of these studies is not for debate here, the logistics of evaluating coaching effectiveness through qualitative and quantitative measures, using both self-report and other-report methods, can be onerous. It is perhaps one of the reasons that Burt and Talati (2017) point out how scarce it is in both the USA and UK, to evaluate coaching outcomes at all.

360° reviews

Essentially, a 360° review is a method of a central person being rated on a set scale by multiple people in different categories (for example, manager, peers and direct reports) across a set of items that cover a sample of role competencies. Accordingly, Bracken, Rose and Church (2016, p.764) supply the following definition:

> *360° Feedback is a process for collecting, quantifying, and reporting co-worker observations about an individual (i.e., a ratee) that facilitates/enables three specific data-driven/based outcomes: (a) the collection of rater perceptions of the degree to which specific behaviors are exhibited; (b) the analysis of meaningful comparisons of rater perceptions across multiple ratees, between specific groups of raters for an individual ratee, and for ratee changes over time; and (c) the creation of sustainable individual, group, and/or organisational changes in behaviors valued by the organisation.*

The use of 360° as a feedback tool in organisations is becoming more popular as a means of helping leaders develop insight and awareness about their management and leadership behaviours for purposes of performance management, leadership development, and, increasingly, to measure the efficacy of coaching interventions (Luthans & Peterson, 2003). However, the 360° approach is not without its criticisms. For example, De Meuse et al. (2009) raise issues such as the continuity of raters that may change over time and benchmarks for the variables that may change over time. Lawrence and Bachkirova (2023) argue that 360° reviews can oversimplify the nature of the self, not taking account of the multiplicity of selves, where the same person has different forms of agency with different people and in different environments. Yet, at the same time, while there is a built-in linearity to these evaluations, they do offer individuals a rare opportunity to see themselves through the lens of others.

There is no shortage of advocates of 360° reviews; for example, London and Beatty (1993) suggest they provide a competitive advantage to those organisations using them. Performance management has been one of the principal areas in which 360° reviews have found purchase (Ghorpade, 2000; Kanaslan & Iyem, 2016). Chen (2022) considers that 360° reviews with built-in multi-rater components, anonymity of raters and opportunities to grow from the feedback give them a greater advantage compared to other appraisal systems. Another benefit is that they are designed to provide real-world metrics from which merit can be quantified. For some researchers, 360°'s reductionism is a great concern (Harrington, 2013; Lawrence & Bachkirova, 2023). While Ghorpade is generally positive towards 360° reviews, arguing that the gains from using them "are substantial" (p. 149), he also cautions against some pitfalls. He suggests that although reviewers may be honest due to their anonymity, this does not necessarily mean that their reviews (or scores) are more valid. Moreover, he suggests that raters may not have a full appreciation for the complexities of the role and therefore respond from a position of limited knowledge or subjective bias.

Like most tools, 360° reviews must be used responsibly, dealing with as many derailers as possible. Bracken and Church (2013) note the positive evolution of the 360° process over time and attest to the benefits when the components are handled competently. According to them, attention should be paid to aspects such as the selection of items within the survey, the nomination of appropriate raters, clear communications regarding the purpose of the reviews and sensitivity in providing feedback. These elements can work towards creating meaningful outputs that could be thoughtfully harnessed in a development process.

Coaching and 360° reviews

Meta-analyses on coaching effectiveness add crucial information to the literature and generally speak to the profession but disclose little about the individual providing a service to an organisation. While not an ideal instrument for creating behaviour change, the 360° survey is highly flexible and offers clear temporal metrics that, at a minimum, provide a useful read on broad directional change. The 360° review with feedback on its own appears to be somewhat limited for development purposes (Nowack & Mashishi, 2012), yet when coupled with a coaching intervention, it offers considerably more promise (Smither, London, Flautt, Vargas & Kucine, 2003). In addition, usage demands of a 360° review are relatively modest, even when doing pre and post evaluations. That makes them a low-touch, scalable method for getting multiple voices weighing in on the effectiveness of a coaching intervention. A further advantage is the multiple vantage points from which visible behaviour change is viewed in a 360° survey. With self, managers, peers, direct reports, and sometimes other stakeholders doing the reviews, the extent of change metrics across these different categories is easily delineated.

Findings from a meta-analysis by Jones et al. (2015) indicated that multisource feedback in the coaching process resulted in a lower effect size on a positive coaching outcome. These results were counter to expectations, and the authors surmise that it could have created a preoccupation with negative feedback, diverting attention from the actual coaching. It is unclear from the study where the multisource feedback occurred. Still, we propose that 360° feedback occurs as an adjunct to the coaching sessions and is established as a separate process component.

Methodology

The intention of this study is to indicate how the measurement method of using pre and post 360° reviews, can highlight coaching efficacy relatively easily. Pre and post 360° methods have been used before to indicate development in other areas such as management training (Rosti & Shipper, 1998) and teacher development (Mahar & Strobet, 2010). In this case, the outcomes were positive but irrelevant as we are not attempting to demonstrate the efficacy of coaching per se but the method's usefulness in tracking coaching efficacy. It is important as coaching interventions are amorphous, carry a certain invisible mystique, and their qualitative nature in some way eschews interrogation. Using 360° reviews and translating coaching outcomes into simple behavioural terms allow for a relatively unambiguous quantification of the results. It holds benefits for the coach, the coachee, and the organisation.

Questionnaire

From an existing leadership competency item bank, a 360° questionnaire was developed for the pre and post reviews. The questions were designed in conjunction with the organisation based on their bespoke leadership competencies, values, style and culture. Questions were adjusted based on the emerging coaching goals, specifically around a more democratic and consultative leadership style, self-management and resilience. The 67 items in the questionnaire covered the 24 sub-categories listed in Table 2. Each sub-category was made up of multiple items. Four of the sub-categories had four (4) items, 11 had three (3) items, and nine (9) consisted of two (2) items.

Each was connected to a 6-point Likert scale encompassing (1) Never, (2) Hardly Ever, (3) Sometimes, (4) Often, (5) Mostly, and (6) Always. There was also an *I Don't Know* option.

Raters

The raters consisted of a self-rating by Heather (the case study subject), her manager, several colleagues, and subordinates, as in Table 1. While ten (10) raters were used in total, only seven (7) were constant across pre and post reviews. All measurements pertain to the seven (7) constant raters.

Table 1: Number of raters in pre and post-reviews

	Pre	Post	Comments
Self	1	1	Heather, in both cases
Manager	1	1	Same Manager
Colleagues	3	3	Same Colleagues
Subordinates	3	3	Same Subordinates
All-Raters*	7	7	All 7 reviewers were constant across both surveys

*Note: The *All-Raters* category aggregates all the raters, excluding the self.

Coaching case study

Heather (pseudonym) was employed in a regional business manager role for a multinational corporation. She had ten direct reports working as sales representatives for the company, calling on professional clients in a specific geographic area. Operating in a highly competitive industry, there was a strong focus on performance in the organisation. Heather was highly regarded in the organisation as an outspoken manager who met and often exceeded sales targets, for which she won a highly prestigious and much-admired internal award in the year preceding her enrolment for coaching. By her own admission, she was a *competitive driver* and pushed her team relentlessly for results.

Heather had come into coaching about a month after a national lockdown had been declared due to the COVID-19 pandemic. She and her team were in the throes of adjusting to working from home and making remote calls to their vast base of clients. Another significant event in the lead-up to Heather coming into coaching was that in the months before, she had been the subject of a formal grievance process by one of her subordinates, having been accused of victimisation and having a poor leadership style.

The coaching process commenced with a three-way meeting between Heather, her manager and the coach. The purpose of such a leadership contracting session is contracting, onboarding, and achieving alignment on tentative coaching goals. A significant fact is that she and her manager were peers before his promotion (as her manager) a few years back. His hopes for her coaching journey were to develop a better relationship with her, for her to temper her coping mechanisms and to achieve a better work-life integration.

The next step was to initiate the psychometric assessment and preliminary 360° feedback exercise as coaching inputs. Based on the discussions and input, Heather formulated three coaching goals (paraphrased):

  i. Becoming a more inclusive leader and softening her interpersonal style.
 ii. Asking for others' opinions instead of being the one with all the answers.
iii. Improving her work-life integration, giving mental and physical time to herself.

Heather completed eight coaching sessions, by which time she believed she had made significant progress.

Results

Approximately six months after the vendor's coaching process, Heather participated in a post coaching 360° review. The items used in the post were identical to those used at the initial (pre) baseline. There were 67 items in total in the 360° survey. Cronbach alphas were calculated using the pre coaching reviews of all eight (8) respondents. The reliability of each sub-category is presented in Table 2 below, indicating a solidly constructed instrument.

Table 2: Pre-survey reliabilities of the coaching constructs

Sub-category	Pre Cronbach alpha	Number of items
Attention to detail	0.774	2
Problem solving	0.879	3
Able to delegate	0.877	3
Building relationships	0.957	3
Communication	0.917	4
Consulting	0.863	3
Diplomatic	0.929	3
Influencing	0.778	2
Recognising and acknowledging	0.836	3
Teamwork	0.888	4
Delivery	0.981	2
Flexible	0.920	3
Integrity	0.926	2
Interpersonal style	0.975	4
Performance	0.965	2
Principled	0.716	2
Resilience	0.888	3
Self-management	0.950	3
Decision to lead	0.885	2
Innovative	0.941	4
Practical leadership	0.904	3
Shepherd leadership	0.893	2
Strategic leadership	0.865	3
Talent management	0.901	2

Aligning the 67 items to the coaching goals resulted in identifying eight (8) items of particular interest assembled into three (3) sets for rating Heather's general behaviour.

The first set related to Heather being a more inclusive leader and softening her interpersonal style:

1. Constantly looks for ways to make team feel empowered.
2. Fosters teamwork through sharing knowledge and being inclusive.
3. Uses subtle language to deal with difficult interpersonal situations.

The next set of items related to asking for others' opinions instead of being the one with all the answers:

4. Considers the input from others when making decisions.
5. Listens well to people's different viewpoints to find the best solution to a problem.
6. Shows respect for people with views and ideas different to her own.

The last set of items related to Heather improving her work-life integration, allowing mental and physical time to herself:

7. Is able to delegate work to others.

8. Gives space for members of the team to do their jobs, avoiding micro-management.

As observed from the pre columns in Table 3 below, there were notably low scores from several raters for some items. Although people typically use the scale idiosyncratically, it is relatively uncommon to see scores of 2 and 3 in a 360° review. In Heather's case, 14 of the 32 rater pre-scores (or 44%) are lower than 4, indicating the extent of the coaching leadership issues. Only one (1) of the 32 scores in the post columns is below 4. That suggests a solid improvement on the pre-measure.

Table 3: The average scores of the ratings on the pre and post reviews

	Self-score		Manager		Colleagues		Subordinates		All-Raters	
Question	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Listens well to people's different viewpoints to find the best solution to a problem	3.0	4.0	4.0	5.0	3.7	4.7	3.3	5.5	3.7	5.1
Uses subtle language to deal with difficult interpersonal situations	4.0	3.0	3.0	4.0	3.3	3.3	3.0	5.0	3.1	4.1
Considers the input from others when making decisions	2.0	4.0	2.0	5.0	4.3	4.7	4.0	6.0	3.4	5.2
Fosters teamwork through sharing knowledge and being inclusive	3.0	5.0	6.0	5.0	4.3	5.3	4.3	5.5	4.9	5.3
Shows respect for people with views and ideas different to her own	5.0	4.0	2.0	5.0	3.3	4.0	3.7	5.0	3.0	4.7
Is able to delegate work to others	4.0	5.0	3.0	5.0	4.7	5.3	4.3	5.5	4.0	5.3
Gives space for team members to do their jobs, avoiding micro-management	5.0	5.0	4.0	5.0	4.3	4.0	4.7	6.0	4.3	5.0
Constantly looks for ways to make the team feel empowered	5.0	5.0	6.0	5.0	5.0	4.3	4.0	6.0	5.0	5.1

Table 3 indicates that Heather's subordinates experienced a new version of her as all eight (8) items increased. Two (2) of the items, *Considers the input from others when making decisions* and *Constantly looks for ways to make the team feel empowered*, went up two (2) points and were scored at the maximum. Her colleagues' scores were more mixed, though the focus of the coaching would have been less visible to them. Nevertheless, the ratings of seven (7) items increased positively, and one, *Constantly looks for ways to make team feel empowered,* regressed slightly.

Coaching did more than impact the behaviours indicated by the eight (8) key coaching items from the 360° review. As Figure 1 below indicates, Heather also benefitted across many other dimensions. The two lines in the graph represent baseline (pre) versus post dimensions using the all-raters ratings (an average of manager, colleague and subordinate scores for each item making up the sub-categories).

Figure 1: All pre and post sub-category scores
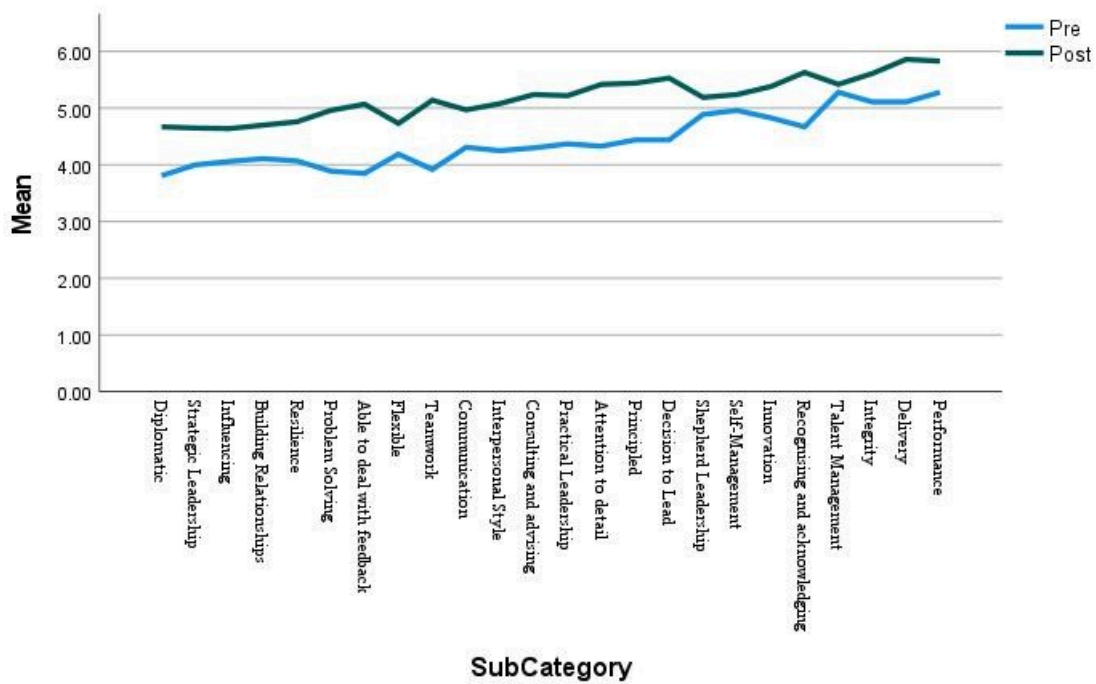
While focused on a few set goals, the coaching intervention had a ripple effect across all the dimensions, with the most profound changes occurring where the gaps between the lines are most significant (e.g. Teamwork, Recognising and acknowledging, and Principled). Paired-sample t-tests can be used to compare the means of two matched groups of people (Ross & Willson, 2017). In our case, it was used to compare the means of all rater types at two different points in time (pre- and post).

Table 4 shows the means for each rater group across all items. In every case except for the Self review, the post item means are higher than the pre item means. The table also indicates that all the changes from pre to post were significant, excluding Heather's, which was not significant and regressed slightly. A possible explanation is that Heather learned more about effective leadership during her coaching and became more self-critical, leading to her more conservative post-rating. De Meuse et al. (2009) describe these as beta changes (the rater's yardstick for the variable shifts or stretches between the two assessments) and gamma changes (which involve a reconceptualization or redefinition of the key variables)

Table 4: Paired samples test

		Paired Differences					Significance
		Mean	N	Std. Deviation	t	df	Two-sided p
Pair1	PostSelf- PreSelf	4.60 4.69	67 67	0.87 1.05	-0.973	66	0.334
Pair2	PostManager-PreManager	5.18 4.48	66 66	0.74 1.23	5.043	65	<,001
Pair3	PostColleagues-PreColleagues	4.75 4.61	67 67	0.68 0.75	2.276	66	0.026
Pair4	PostSubordinates-PreSubordinates	5.52 4.07	67 67	0.32 0.44	26.750	66	<,001
Pair5	PostAllRaters-PreAllRaters	5.00 4.38	67 67	0.50 0.66	11.471	66	<,001

The Bonferroni correction (Emerson, 2020) was applied to account for multiple comparisons. The adjusted p-value after applying the Bonferroni correction was (0.05/5; p=0.01). Based on the adjusted p-value, manager, subordinate and all-raters differences remain statistically significant. Finally, establishing the effect sizes, the Point Estimate column of the differences provides a useful metric on the extent of change as evidenced by the 360° item shifts in Table 5.

Table 5: Paired samples effect sizes

			Standardizer	Point Estimate	95% Confidence Interval	
					Lower	Upper
Pair1	PostSelf- PreSelf	Cohen's d	.75	-.12	-.36	.12
Pair2	PostManager-PreManager	Cohen's d	1.12	.62	.36	.88
Pair3	PostColleagues-PreColleagues	Cohen's d	.50	.28	.03	.52
Pair4	PostSubordinates-PreSubordinates	Cohen's d	.44	3.27	2.66	3.87
Pair5	PostAllRaters-PreAllRaters	Cohen's d	.45	1.40	1.06	1.74

In the general rules of thumb provided by Cohen (1988) and Sawilowsky (2009), anything below 0.50 is a weak difference and unlikely to be significant. However, Parker and Vannest (2009) argue that for single-case studies, group-level thresholds are too lenient and calibrate their guidelines as follows: small < 0.87, medium 0.87 to 2.67, and large > 2.67. Using these criteria, self, manager, and colleagues would all be considered small effects. Only the 3.27 for subordinates would meet the 'large' threshold, indicating that Heather's changes as a leader were most amplified for those reporting directly to her, which aligns with the main focus of the coaching intervention.

Discussion

The case study indicated how the pre and post 360 ° review process can produce clear metrics on changes during a coaching intervention. The coaching intervention can harvest items and themes from the pre 360° review and revisit how they have changed in a post-review. In Heather's case, she was able to learn and experiment with adapted behaviours to function more effectively. Notably, the post 360° using the same items and scale allows for direct statistical analysis of measured change between the onset and termination of coaching.

The method's most apparent barriers and limitations are that while services and platforms offering 360° reviews are widely available (Slashdot, 2023), they add complexity. Logistics, cost and compliance (e.g. survey fatigue) contribute to the reluctance to include 360° reviews in the coaching process (McCarthy and Garavan, 2001). Issues are compounded by the requirement that reviews be performed at least twice. The results of the pre and post surveys then need to be collated, interpreted and fed back to the coach, coachee and sponsor. Having said that, once these processes are in place, they may be streamlined, routine, and used to measure sustained behaviour change. Coaching itself is not a readily scalable intervention, requiring one-on-one sessions tailored to the development needs of the individual. It follows that 360° reviews for tracking coaching efficacy would generally only be required for reasonably sized groups of individuals. Perhaps more importantly, few other and better options are available if scientific methods are needed to measure coaching outcomes.

An additional limitation may be the applicability of 360° content across multiple coaching contexts. Wide variability would mean the onerous creation of bespoke items for each coachee. However, within organisations, there is generally a common culture and set of values (Locander, Hamilton, Ladik & Stuart, 2002) to create a convergence of leadership and staff behaviours. That indicates that organisations often have much common ground, and 360° content could share a wide overlap. In such circumstances, only a few specific items may need to be tailored for individual coaching. However, coaching practitioners often work across companies, but even in these circumstances, widely held values and beliefs transcend organisations and cultures. Several authors have referred to the universality of leadership competencies (Bass, 1997; Dorfman, Javidan, Hanges, Dastmalchian & House, 2012). Dickson, Castaño, Magomaeva and Den Hartog (2012, p. 491) confirm, "Viewed from a level of abstraction, there are aspects of leadership that appear universal across cultures." One way or another, coaches can build up a 360° item bank for reuse and adaptation across varying circumstances.

This discussion also adds another connection to our findings. A seminal work by Ericsson, Krampe and Tesch-Römer (1993) popularised the idea of extended hours of deliberate practice being necessary for superior performance. However, critiqued since then (Hambrick, Oswald, Altmann, Meinz, Gobet & Campitelli, 2014), it has also evolved and been applied to multiple fields of expertise, including music, sports, business and medicine (Chow, Miller, Seidel, Kane, Thornton & Andrews, 2015). While deliberate practice does not explain nearly all the variance in predicting superior performance (Hambrick et al.), it is necessary for expertise. Perhaps most germane to the discussion on coaching is that psychotherapists who routinely seek out deliberate practice, of which receiving feedback is a key component, are rated as being significantly more effective than those who do not engage in deliberate practice (Chow et al., 2015), and thus, are deprived of the opportunity to adapt their methods from regular feedback. From a coaching perspective the implications are that those coaches receiving regular feedback on coaching outcomes can lift their performance. The pre and post 360° reviews create such feedback opportunities.

Conclusion

By embarking on a granular analysis and using statistical tools to interrogate the scores from pre to post, this study was able to establish that the extent of change comfortably met the criteria for significance. This goes beyond the general collective anecdotal messages more commonly used to appraise coaching and thereby reduces the mystique surrounding coaching interventions. The method used in this study can easily be replicated, providing an efficient set of metrics to determine coaching efficacy and ROI. In a coaching industry that is still loosely regulated (Garvey & Stokes, 2023), this will offer sponsoring organisations a set of metrics to help reinforce support for and confidence in the change process.

The research has shown that effect sizes provide a well-established metric indicating the extent of change. What adds to the method's limitation is that organisations can be fluid in staffing, and the raters at the pre-stage may not be in the same positions at the post-stage. Hence, there could be changes in the raters, which may affect the outcome. In addition, there may be demand characteristics in which raters' scores trend upwards at the post-stage, which could skew the results. While the 360° review is not a perfectly calibrated measuring apparatus, it can help determine ROI within a coaching process. One of the common complaints regarding coaching is that it can be an expensive and intangible process. Performed behind closed doors and subject to regulated confidentiality, the sponsoring organisation can feel aggrieved at the minimal feedback it often receives. The pre and post 360° reviews somewhat respond to those criticisms. While reductionistic to some extent, its metrics provide a valuable lens to help determine whether all the efforts put into the coaching process have been worthwhile. Creating a bespoke 360° instrument assists with the alignment between the organisation's values, competency framework and the specific behaviour gaps of the coachee. The 360° feedback can assist the coach in a more holistic understanding of the coachee, providing an opportunity for building and deepening the relationship. In addition, the process creates identifiable, measurable behaviours which can be prioritised during the coaching process. Using pre and post 360° reviews provides feedback to at least four critical stakeholders about the success or otherwise of the coaching – the coachee, the coachee's sponsoring line manager, the coach and, where applicable, the coach's associate organisation. The post 360° feedback may validate the progress made but also highlights remaining areas of needed development for the coach, coachee and line manager regarding ongoing development plans for the coachee. Finally, a post 360° provides a way for coaches to regularly obtain feedback on their coaching interventions. Without this kind of feedback, a coach's subjective intuition on the success of the coaching is likely to be overestimated, as indicated in a study by Hubble, Duncan and Millar (1999) with seasoned psychotherapists. The practitioner coach can thus objectively learn and grow from every coaching encounter.

In summary, using well-designed pre and post 360° instruments could foreground strengths and areas requiring development and enhance the credibility of the coaching profession in highlighting the need for more objective, transparent, quantifiable measures for the efficacy of coaching interventions. For example, consolidated pre and post 360° feedback for a leadership team could track coaching efficacy and highlight strengths and development areas for team coaching interventions. While we know the coaching relationship supersedes coaching modalities and techniques (De Haan et al., 2013), a post 360° can be useful in identifying which coaching interventions have been more or less effective in addressing developmental needs in specific cases.

References

Anthony, E.L. (2017). The impact of leadership coaching on leadership behaviors. *Journal of Management Development*, 36(7), 930-939. DOI: 10.1108/JMD-06-2016-0092.

Bass, B. M. (1997). Does the transactional–transformational leadership paradigm transcend organisational and national boundaries? *American Psychologist*, 52(2), 130–39. DOI: 10.1037/0003-066X.52.2.130.

Bonaiuto, M., De Gregorio, E., Sarrecchia, A. & Gentile, D. (2008). The evaluation of executive coaching effectiveness: Theory, research and critical issues. *Bollettino di Psicologia Applicata*, 254, 3-17.

Bracken, D. W. & Church, A. H. (2013). The "new" performance management paradigm: capitalising on the unrealised potential of 360 degree feedback. *People and Strategy*, 36(2), 34.

Bracken, D. W., Rose, D. S. & Church, A. H. (2016). The evolution and devolution of 360 feedback. *Industrial and Organisational Psychology*, 9(4), 761-794. DOI: 10.1017/iop.2016.93.

Burt, D. & Talati, Z. (2017). The unsolved value of executive coaching: A meta-analysis of outcomes using randomised control trial studies. *International Journal of Evidence Based Coaching and Mentoring*, 15(2), 17-24. DOI: 10.24384/000248.

Cannon-Bowers, J. A., Bowers, C. A., Carlson, C. E., et al., (2023). Workplace coaching: a meta-analysis and recommendations for advancing the science of coaching. *Frontiers in Psychology*, 14, 1204166. DOI: 10.3389/fpsyg.2023.1204166.

Chen, Y. (2022). What are the main objectives of reward management systems, and how has best practice influenced the development of them? *Studies in Social Science & Humanities*, 1(4), 43-49. DOI: 10.56397/SSSH.2022.11.06.

Chow, D. L., Miller, S. D., Seidel, J. A., et al., (2015). The role of deliberate practice in the development of highly effective psychotherapists. *Psychotherapy*, 52(3), 337–345. DOI: 10.1037/pst0000015.

Claudy, K. P., Lamera, A. I. & Underhill, B. O. (2013). Evaluating leadership improvement: using executive coaching to drive change. In *Integrated talent management scorecards: Insights from world-class organizations on demonstrating value*, 221. ASTD Press.

Coates, D. (2013). Integrated leadership development programmes: Are they effective, and what role does coaching play? International Journal of Evidence Based Coaching and Mentoring, S7, 39-55. Available at: https://radar.brookes.ac.uk/radar/items/894af787-235a-414d-a2e5-4b5364a3e888/1/.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

De Haan, E., Duckworth, A., Birch, D. & Jones, C. (2013). Executive coaching outcome research: The contribution of common factors such as relationship, personality match, and self-efficacy. *Consulting Psychology Journal: Practice and Research*, 65(1), 40-57. DOI: 10.1037/a0031635.

De Haan, E., & Nilsson, V. O. (2023). What can we know about the effectiveness of coaching? A meta-analysis based only on randomized controlled trials. *Academy of Management Learning & Education*, 22(4), 1-21. DOI: 10.5465/amle.2022.0107.

De Meuse, K.P., Dai, G. & Lee, R.J (2009) Evaluating the effectiveness of executive coaching: beyond ROI?, *Coaching: An International Journal of Theory, Research and Practice*, 2(2), 117-134. DOI: 10.1080/17521880902882413.

Dickson, M. W., Castaño, N., Magomaeva, A. & Den Hartog, D. N. (2012). Conceptualising leadership across cultures. *Journal of World Business*, 47(4), 483-492. DOI: 10.1016/j.jwb.2012.01.002.

Dorfman, P., Javidan, M., Hanges, P., et al., (2012). GLOBE: A twenty year journey into the intriguing world of culture and leadership. *Journal of World Business*, 47, 504–518. DOI: 10.1016/j.jwb.2012.01.004.

Emerson, R. W. (2020). Bonferroni correction and type I error. *Journal of Visual Impairment & Blindness*, 114(1), 77-78. DOI: 10.1177/0145482X20901378.

Ericsson, K. A., Krampe, R. T. & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363-406. DOI: 10.1037/0033-295X.100.3.363.

Faire, M. (2023). The three Cs of coaching and professional practice. In *The Coach's Casebook: Skills, Tools and Techniques for Effective Coaching*. London: Kogan Page.

Gallagher, T. L., Grierson, A. & Susin, C. A. (2023). Facilitating professional learning for technology coaches through cross-district collaboration. *International Journal of Mentoring and Coaching in Education*, 12(2), 216-230. DOI: 10.1108/IJMCE-07-2022-0051.

Garvey, B. & Stokes, P. (2023). Ethics and professional coaching bodies. In *The Ethical Coaches' Handbook: A Guide to Developing Ethical Maturity in Practice*. Routledge.

Ghorpade, J. (2000). Managing five paradoxes of 360-degree feedback. *Academy of Management Perspectives*, 14(1), 140-150. DOI: 10.5465/ame.2000.2909846.

Grant, A. M. (2017). The third 'generation' of workplace coaching: Creating a culture of quality conversations. *Coaching: An International Journal of Theory, Research and Practice*, 10(1), 37-53. DOI: 10.1080/17521882.2016.1266005.

Greene, H. C. (2019). Coaching military special-operations forces. *Consulting Psychology Journal: Practice and Research*, 71(2), 107. DOI: 10.1037/cpb0000131.

Hambrick, D. Z., Oswald, F. L., Altmann, E. M., et al., (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence*, 45, 34-45. DOI: 10.1016/j.intell.2013.04.001.

Harrington, A. (2013). *Numbers, words and anonymity in 360-degree feedback: A qualitative study* (Doctoral dissertation, Loughborough University).

Henochowicz, S. & Hetherington, D. (2006). Leadership coaching in health care. *Leadership & Organization Development Journal*, 27(3), 183-189. DOI: 10.1108/01437730610657703.

Herbison, G. (2023). Developing leadership capabilities to drive engagement. *Journal of Financial Planning*, 36(1), 46-48.

Hubble, M. A., Duncan, B. L. & Miller, S. D. (1999). *The heart and soul of change: What works in therapy*. APA. DOI: 10.1037/11132-000.

ICF (2023). *Professional coaches membership and credentialing sheet*. Available at: https://coachingfederation.org/.

Jarosz, J. (2023). The cube of coaching effectiveness. *International Journal of Evidence Based Coaching & Mentoring*, 21(1). DOI: 10.24384/gkny-df71.

Jones, R. J., Woods, S. A. & Guillaume, Y. R. (2016). The effectiveness of workplace coaching: A meta-analysis of learning and performance outcomes from coaching. *Journal of Occupational and Organisational Psychology*, 89(2), 249-277. DOI: 10.1111/joop.12119.

Kanaslan, E. K. & Iyem, C. (2016). Is 360 degree feedback appraisal an effective way of performance evaluation? *International Journal of Academic Research in Business and Social Sciences*, 6(5), 172-182. DOI: 10.6007/IJARBSS/v6-i5/2124.

Kauffman, C. & Hodgetts, W. H. (2016). Model agility: Coaching effectiveness and four perspectives on a case study. *Consulting Psychology Journal: Practice and Research*, 68(2), 157–176. DOI: 10.1037/cpb0000062.

Lawrence, P. & Bachkirova, T. (2023). Challenging traditional approaches: 360° feedback and theories of the multiplicity of self. *International Journal of Evidence Based Coaching & Mentoring*, 21(1). DOI: 10.24384/d294-gy98.

Locander, W. B., Hamilton, F., Ladik, D. & Stuart, J. (2002). Developing a leadership-rich culture: The missing link to creating a market-focused organisation. *Journal of Market-Focused Management*, 5, 149-163. DOI: 10.1023/A:1014048111158.

London, M. & Beatty, R. W. (1993). 360-degree feedback as a competitive advantage. *Human Resource Management*, 32(2-3), 353-372. DOI: 10.1002/hrm.3930320211.

Luthans, F. & Peterson, S.J. (2003). 360-degree feedback with systematic coaching: Empirical analysis suggests a winning combination. *Human Resource Management*, 43, 243-256. DOI: 10.1002/hrm.10083.

Mahar, J. A. & Strobert, B. (2010). The use of 360-degree feedback compared to traditional evaluative feedback for the professional growth of teachers in K-12 education. *Planning and Changing*, 41, 147-160. Available at: https://files.eric.ed.gov/fulltext/EJ952378.pdf.

McCarthy, A. M. & Garavan, T. N. (2001). 360 feedback process: Performance, improvement and employee career development. *Journal of European Industrial Training*, 25(1), 5-32. DOI: 10.1108/03090590110380614.

Molyn, J., de Haan, E., van der Veen, R. & Gray, D. E. (2022). The impact of common factors on coaching outcomes. *Coaching: An International Journal of Theory, Research and Practice*, 15(2), 214-227. DOI: 10.1080/17521882.2021.1958889.

Nock, M. K., Michel, B. D. & Photos, V. (2007). Single-case research designs. In D. McKay (Ed.). *Handbook of research methods in abnormal and clinical psychology* (337-350). Thousand Oaks, CA: Sage Publications.

Nowack, K. M. & Mashihi, S. (2012). Evidence-based answers to 15 questions about leveraging 360-degree feedback. *Consulting Psychology Journal: Practice and Research*, 64(3), 157-182. DOI: 10.1037/a0030011.

Parker, R. I. & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, 40(4), 357-367. DOI: 10.1016/j.beth.2008.10.006.

Pavur, E. J. (2013). Why do organisations want their leaders to be coached? *Consulting Psychology Journal: Practice and Research*, 65(4), 289–293. DOI: 10.1037/a0035414.

Phillips, J. & Phillips, P. (2004). Measuring ROI in executive coaching. *Executive Coach Magazine*, 18-24.

Ross, A. & Willson, V. L. (2017). Paired samples T-test. In *Basic and advanced statistical tests*. (17-19). Brill.

Rosti, R. T. & Shipper, F. (1998), A study of the impact of training in a management development program based on 360 feedback. *Journal of Managerial Psychology*, 13(1/2), 77-89. DOI: 10.1108/02683949810369147.

Sawilowsky, S.S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8(2), 597–99. DOI: 10.22237/jmasm/1257035100.

Slashdot (2023). *Best 360 degree feedback software of 2023*. Available at: https://slashdot.org/software/360-degree-feedback/.

Smither, J. W., London, M., Flautt, R., et al., (2003). Can working with an executive coach improve multisource feedback ratings over time? A quasi-experimental field study. *Personnel Psychology*, 56(1), 23-44. DOI: 10.1111/j.1744-6570.2003.tb00142.x.

Stawiski, S. & McCauley, C. D. (2023). Maximising the effectiveness of assessment-rich leadership development programs. In *Talent assessment: Embracing innovation and mitigating risk in the digital age*, 312-C20P72. DOI: 10.1093/oso/9780197611050.003.0020.

Tooth, J. A., Nielsen, S. & Armstrong, H. (2013). Coaching effectiveness survey instruments: Taking stock of measuring the immeasurable. *Coaching*, 6(2), 137–151. DOI: 10.1080/17521882.2013.802365.

Wang, Q., Lai, Y. L., Xu, X. & McDowall, A. (2021). The effectiveness of workplace coaching: A meta-analysis of contemporary psychologically informed coaching approaches. *Journal of Work-Applied Management*, 14(1), 77-101. DOI: 10.1108/JWAM-04-2021-0030.

About the authors

Hilton Rudnick is a research associate at the University of Johannesburg and the Managing Director of Omnicor, a consultancy working in organisational development and psychometric services. Hilton is a counselling psychologist and a coaching practitioner. He has a particular interest in metrics related to work performance and data analytics.

Neville Goldin is a Work Psychologist with a long career in corporate Human Resources (Organisation and Leadership Development). Neville currently practices as a Coaching and Consulting Psychologist to individuals and teams in organisational settings. He is a certified Tavistock coach, using the systems psychodynamic method and enjoys coaching leaders through career transitions.