The effectiveness of workplace coaching: A meta-analysis of learning and performance outcomes from coaching

Rebecca J. Jones[1]*, Stephen A. Woods[2] and Yves R. F. Guillaume[3]

[1]University of Worcester, UK
[2]University of Surrey, UK
[3]Aston University, Birmingham, UK

This study presents a meta-analysis synthesizing the existing research on the effectiveness of workplace coaching. We exclusively explore workplace coaching provided by internal or external coaches and therefore exclude cases of manager–subordinate and peer coaching. We propose a framework of potential outcomes from coaching in organizations, which we examine meta-analytically ($k = 17$). Our analyses indicated that coaching had positive effects on organizational outcomes overall ($\delta = 0.36$), and on specific forms of outcome criteria (skill-based $\delta = 0.28$; affective $\delta = 0.51$; individual-level results $\delta = 1.24$). We also examined moderation by a number of coaching practice factors (use of multisource feedback; type of coach; coaching format; longevity of coaching). Our analyses of practice moderators indicated a significant moderation of effect size for type of coach (with effects being stronger for internal coaches compared to external coaches) and use of multisource feedback (with the use of multisource feedback resulting in smaller positive effects). We found no moderation of effect size by coaching format (comparing face-to-face, with blended face-to-face and e-coaching) or duration of coaching (number of sessions or longevity of intervention). The effect sizes give support to the potential utility of coaching in organizations. Implications for coaching research and practice are discussed.

Practitioner points

- Our meta-analysis supports the positive effects of workplace coaching as an approach to employee learning and development in organizations, with a variety of criteria.
- Our findings indicate that coaching was more effective when conducted by internal coaches and when multisource feedback was excluded.
- Workplace coaching was effective whether conducted face-to-face or using blended techniques (i.e., blending face-to-face with e-coaching).

Is workplace coaching effective in terms of delivering individual learning and development, and improvements in performance and results for organizations? Despite the huge growth in the use of coaching as a strategy for employee learning and development (ICF, 2012), there remains a paucity of scientific evidence examining its benefits for organizations, coupled with a generally poor specification of the types of outcomes that

*Correspondence should be addressed to Rebecca J. Jones, University of Worcester, Castle Street, Worcester WR1 3AS, UK (email: rebecca.jones@worc.ac.uk).

can be expected from coaching. In this study, we report a meta-analysis of the effects of coaching on performance, and other learning and development outcome criteria. We also examine moderators of these effects, focusing on the techniques and features of specific coaching interventions. Our systematic review of studies evaluating the benefits of coaching in organizations advances the literature in three main ways. First, we address problems of criterion specification in the coaching literature by proposing and examining a framework of coaching outcomes aligned to the learning, training, and development literature. Second, unlike previous studies (e.g., Theeboom, Beersma, & van Vianen, 2014), we clarify the effectiveness of workplace coaching specifically by focusing our analyses exclusively on organizational (and not general or educational) samples, and on coaching provided by practitioner coaches (and not managers or peers), closing an important empirical gap in the field. Finally, we examine a number of practice moderators of the effectiveness of coaching, analyses of which have important implications for practitioner coaches concerning the effectiveness of specific coaching tools and techniques.

Workplace coaching

Workplace or executive coaching (hereafter referred to as 'coaching') is a one-to-one learning and development intervention that uses a collaborative, reflective, goal-focused relationship to achieve professional outcomes that are valued by the coachee (Smither, 2011). Coaching has typically been perceived as an 'executive' development activity, but it need not necessarily be limited to managers and executives in organizations. Although the term coaching may be used to refer to a variety of one-to-one development activities, there is some emerging consensus about what constitute the core features or elements of coaching (e.g., see Bono, Purvanova, Towler, & Peterson, 2009; Smither, 2011). These are (1) formation and maintenance of a helping relationship between the coach and coachee; (2) a formally defined coaching agreement or contract, setting personal development objectives; (3) the fulfilment of this agreement (i.e., achievement of the objectives) through a development process focusing on interpersonal and intrapersonal issues; (4) striving for growth of the coachee by providing the tools, skills, and opportunities they need to develop themselves and become more effective (Bono *et al.*, 2009; Kilburg, 1996; McCauley & Hezlett, 2002; Peterson & Hicks, 1996; Smither, 2011; Witherspoon & White, 1996).

The coaching relationship is one that the coachee enters into for the specific purpose of fulfilling development objectives. It is important to differentiate coaching from other forms of developmental relationships in the workplace. Conceptually, it may first be distinguished from mentoring relationships (see Brockbank & McGill, 2012 for a review). A mentoring relationship is conventionally long term between a highly experienced mentor and an inexperienced mentee. The mentor is assumed to be highly experienced in the discipline or field in which the mentee is working, and in the workplace, the mentor typically provides guidance on career development and networking (Eby *et al.*, 2013). In a coaching relationship, there is no such expectation that the coach has expertise or experience of the coachee's area of work, and the term of the relationship is rather guided by specific objectives.

Similarly, there are relative status prerequisites in 'peer coaching' (Parker, Hall, & Kram, 2008; Parker, Kram, & Hall, 2013), in which development is a two-way reciprocal process between people of equal status in an organization. Peer coaching aims to provide mutually supportive personal and professional development of both people in the

peer-coaching dyad. The developmental focus in coaching, by contrast, is solely the coachee, and the relationship is free from the influence or boundaries of organizational status structures.

The coaching relationship is also typically considered to be distinct from formalized organizational performance management relationships (e.g., supervisor–subordinate; Feldman & Lankau, 2005). For example, there exists a power relationship between line managers and their subordinates, which is absent in the helping relationship a coachee would have with an independent coach. Managers and supervisors might propose coaching as a developmental intervention for their staff (and in this sense, coaching might be considered a part of performance management processes), but to fit with the relational definition of coaching, the developmental relationship that facilitates learning and development would be with a coach, rather than the manager or supervisor.

In this study, we conceptualize coaching as distinct from these other forms of developmental relationship. Correspondingly, in our analyses, we specifically focus on studies of coaching and exclude studies of other relationships (e.g., mentoring, supervisor coaching, and peer coaching).

In terms of practical utility, keeping coaching distinct from other organizational performance management and development relationships offers potential advantages. Sherman and Freas (2004) report that the relational nature of coaching provides an individual, customized feel to coaching, with coaches providing candour, and honest feedback to the coachee in relation to their performance and behaviour. This is frequently supplemented with feedback from the coachee's organization (e.g., through multisource feedback). However, the privacy, non-judgmental perspective, and confidentiality of the coaching session provide a safe environment for the coachee to reflect on that feedback and work on improving areas of weakness. The coach may discuss suggested tools and techniques to help the coachee develop and improve, the content of which is dependent on the background and approach of the coach. However, coaches generally avoid providing instructional or prescriptive solutions to coachees, because as highlighted above, they are often not technical experts in the coachee's occupational area of specialty (McAdam, 2005).

In the context of human resource management (HRM), coaching represents an approach to employee learning, training, and development, and although there remains lack of consensus about the core processes of coaching as a development intervention (a point that we return to later), coaching sessions are frequently characterized either directly or indirectly with a number of performance-promoting features.

First, coaches apply goal-setting: Well established as a technique of performance improvement (e.g., Locke & Latham, 1990, 2002; Morisano, Hirsh, Peterson, Pihl, & Shore, 2010; Wegge, Bipp, & Kleinbeck, 2007). Second, goals generally feature activities undertaken whilst at work, promoting experiential forms of practice and learning (e.g., Kolb, 1984). Related to this and thirdly, by encouraging learning through practice at work, coaching rather directly promotes translation of learning to work performance behaviour, addressing the issue of transfer, often cited as a barrier to performance benefits of training (e.g., Baldwin & Ford, 1988). In this respect, the personalized nature of coaching may provide a high-fidelity form of workplace learning (Kozlowski & DeShon, 2004; van der Locht, van Dam, & Chiaburu, 2013).

Coaching effectiveness

In the literature to date, the case has been building, based primarily on anecdotal evidence and uncontrolled studies, that coaching is effective at improving work-based outcomes including goal accomplishment (Fischer & Beimers, 2009), professional growth (McGuffin & Obonyo, 2010), improved professional relationships (Kombarakaran, Yang, Baker, & Fernandes, 2008), greater managerial flexibility (Jones, Rafferty, & Griffin, 2006), increased productivity (Olivero, Bane, & Kopelman, 1997), and improved resilience and workplace well-being (Grant, Curtayne, & Burton, 2009). Coaching is also aligned with recent emergent interest in active rather than passive learning, by which employees take responsibility for shaping their own learning processes (Bell & Kozlowski, 2008). Coaching is led by the coachee, giving them control over their learning and development, and the increasing popularity of coaching in organizations may therefore reflect a more general trend away from 'one size fits all' approaches to training (Salas & Kozlowski, 2010).

Despite the apparent potential advantages of coaching, research has not kept pace with its growth in practice, and the lack of conclusive evidence regarding the effectiveness of coaching is one of the most frequently cited problems in the field (Grant, Passmore, Cavanagh, & Parker, 2010). Shortcomings in the research evidence base for coaching include problems of empirical research design and criterion measurement in evaluation studies (Grant, Passmore, *et al.*, 2010).

Recognizing the need for systematic evidence in this field, Theeboom *et al.* (2014) reported a meta-analysis of the effects of coaching on several outcome criteria. Using Hedges' g, which corrects for potential bias due to overestimate of population effect size when small samples are included in the analysis (Hedges, 1981), they reported positive overall effects of coaching with aggregated outcomes ($g = 0.66$), and with specific kinds of criteria: Performance and skills ($g = 0.60$); well-being ($g = 0.46$); coping ($g = 0.43$); work attitudes ($g = 0.54$); and goal-directed self-regulation ($g = 0.74$). Positive effects were moderated by research design (within-subjects research design studies $g = 1.15$, compared to mixed design studies $g = 0.39$). Moderator testing also showed no moderation by number of coaching sessions leading Theeboom *et al.* (2014) to conclude that the number of coaching sessions is not related to the effectiveness of the intervention.

Whilst the meta-analysis of Theeboom *et al.* (2014) represents an important advance in the evidence base for coaching generally, the implications for coaching in organizations specifically are less clear. This is because in their analyses, Theeboom *et al.* (2014) compute effect sizes which are derived from studies of coaching in a variety of contexts. For example, results from studies of workplace coaching are combined with results from studies of coaching conducted for different purposes (e.g., of the 18 studies included in the analyses, six report the results of general life coaching, and one reports results of health coaching). Moreover, Theeboom *et al.* (2014) mix studies using organizational samples with studies based on educational and general non-organizational samples (e.g., seven of the 18 primary studies included were conducted in non-organizational samples such as student or general population convenience samples). As a consequence, the implications of the meta-analysis for organizations applying coaching as part of HRM, for performance improvement at work, are confounded and therefore inconclusive.

The problem of applying the findings of Theeboom *et al.* (2014) to the literature on learning, training, and development in organizations is further compounded by issues of criterion specification. Although five categories of criteria were analysed (performance/skills; well-being; coping; work attitudes; goal-directed self-regulation), they appear to have been defined bottom-up (i.e., based on those criteria measured in the included

studies) instead of top-down (i.e., by applying systematic criterion framework to classify outcomes). Uncertainty about the specific nature of the work outcomes from coaching that might be expected (e.g., Bennett, 2006; Brotman, Liberi, & Wasylyshyn, 1998; Lowman, 2005) therefore remains unaddressed in the literature.

To address these limitations, in our study, we seek to more clearly integrate the literature on coaching with wider literature on employee learning, training, and development at work. We therefore focus exclusively on studies using organizational samples that examine work-relevant outcomes (i.e., excluding variables such as general life satisfaction). This approach is consistent with recent calls in the organizational behaviour literature to consider more carefully the context and setting of research studies (Brutus, Gill, & Duniewicz, 2010; Stone-Romero, 2008). Moreover, we propose and examine a criterion framework of coaching outcomes at work, derived from existing criterion models in the learning, training, and development literature. This provides a more robust classification of criterion effects and enables our findings to be better integrated with evidence from industrial, work and organizational psychology, and HRM research on training effectiveness. Finally, addressing the need for research on the features and techniques of coaching that make it more or less effective (Grant, Passmore, *et al.*, 2010; Swart & Harcup, 2013), and adding to the practical utility of our study for coaching practice at work, we also test several coaching practice moderators that may influence the effectiveness of coaching.

Outcome criteria of coaching and their measurement

There is little consensus in the literature regarding the most appropriate outcome criteria for evaluating coaching (Grant, Passmore, *et al.*, 2010; MacKie, 2007; Smither, 2011). To advance in this respect, it is possible to draw on established criterion models from the literatures on learning, training, and development to propose a criterion framework for evaluating coaching.

The ultimate aim of learning, training, and development is to maximize the effectiveness of an organization's human capital (Ford, Kraiger, & Merritt, 2010) by improving performance at the individual level, on the assumption that this will subsequently result in organizational-level improvements (Swart & Harcup, 2013). Learning, training, and development interventions are therefore positioned as a central function of HRM, integrated with performance management processes more widely (e.g., Murphy & DeNisi, 2008). Coaching is compatible with this conceptualization because the aim of coaching is to aid the achievement of individual goals for improvement, in order to positively contribute to organizational-level goals and objectives. Given this compatibility, outcome criteria for training effectiveness may represent a sensible foundation for modelling the potential outcomes of coaching.

In the training literature, Kirkpatrick (1967) model of evaluation criteria proposes that the evaluation of training should be performed at four levels: Reactions, learning, behaviour, and results. Kirkpatrick's model is widely applied in research and practice (e.g., Alliger, Tannenbaum, Bennett, Traver, & Shotland, 1997; Arthur, Bennett, Edens, & Bell, 2003; Powell & Yalcin, 2010; Tharenou, Saks, & Moore, 2007) and represents a logical organization and progression of outcomes from basic individual reactions to training through to training transfer and organizational results. Kraiger, Ford, and Salas (1993) argued that when evaluating training, it was necessary to examine learning-based outcomes in a more sophisticated way, rather than combining learning and transfer outcomes as in the Kirkpatrick levels. Kraiger *et al.* (1993) proposed a model of three

classes of learning outcomes that occur following training: Cognitive; skill-based, and affective outcomes. This three-component model of learning outcomes has been applied in numerous studies (e.g., Kalinoski *et al.*, 2013; Mesmer-Magus & Viswesvaran, 2010; Taylor, Russ-Eft, & Chan, 2005).

Based on these existing criterion frameworks, we propose that the outcomes of coaching be modelled in the following ways. With respect to the Kraiger *et al.* (1993) three-component classification, we propose that the potential outcomes of coaching are similarly separated into cognitive, skill-based, and affective outcome criteria. Examples of cognitive outcomes from coaching include new declarative and procedural knowledge which could be acquired by self-directed learning and problem-solving (guided by goal-setting). The work-based application of improvement and development activity that is encouraged in coaching is likely to promote skill acquisition and enhancement, effectively resulting in skill-based outcomes. In addition, many of the intended benefits of coaching represent affective outcomes, such as the development of self-efficacy and confidence, reduction of stress, and increased satisfaction and motivation.

We also note the aims of coaching to contribute to achievement of organizational-level goals and objectives (e.g., Sonnentag & Frese, 2002). By aligning individual goal-setting to these organizational-level goals and objectives, coaching may impact performance, making it sensible to include some measure of results, as per Kirkpatrick's model (Kirkpatrick, 1996). Results might be measured in terms of impact on individual-, team-, and organizational-level performance. Our proposed set of evaluation criteria for coaching is summarized in Table 1. Drawing on Kraiger *et al.* (1993) and Kirkpatrick (1996), we also specify acceptable measurement methodology for assessing these outcome criteria, which were applied in our coding of studies for the meta-analysis.

Given its role as a learning and development intervention, we expected coaching to influence all of our proposed evaluation criteria. However, in our meta-analysis, we were only able to test effects of coaching on affective, skill-based, and individual-level results outcomes. We return to this point in our discussion of findings.

> *H1:* Coaching will demonstrate positive effects for affective, skill-based, and individual-level results outcome criteria.

Table 1. Summary of proposed coaching evaluation criteria

Outcome criteria	Description	Measurement methodology
Affective outcomes	Attitudes and motivational outcomes (e.g., self-efficacy, well-being, satisfaction)	Self-report questionnaires
Cognitive outcomes	Declarative knowledge; procedural knowledge; cognitive strategies (e.g., problem-solving)	Recognition and recall tests
Skill-based outcomes	Compilation and automaticity of new skills (e.g., leadership skills, technical skills, competencies)	Behavioural observation in the workplace (e.g., multisource feedback questionnaire)
Results	Individual, team, and organizational performance	Financial results, objective or goal achievement, and productivity

Research design

Alongside the nature of coaching outcome criteria, a further methodological consideration for studies examining the effects of coaching is the research design employed to measure those criteria. Within-subjects research design has frequently been utilized by other researchers in the field of training evaluation (e.g., Dierdorff, Surface, & Brown, 2010; Franke & Felfe, 2012; Patrick, Smy, Tombs, & Shelton, 2012). Effect sizes in within-subjects designs represent the differences between measurements of criteria taken before and after the coaching has taken place (with varying duration of time between measurements depending on the number and schedule of coaching sessions). An alternative design is the between-subjects design (e.g., Ayres & Malouff, 2007; Holladay & Quiñones, 2003; Neal *et al.*, 2006; Orvis, Fisher, & Wasserman, 2009). In these studies, effect sizes represent the differences between control and experimental (i.e., coaching) groups measured after the coaching has taken place.

In our meta-analysis, we adopt a position that these two methodologies (within-subjects and between-subjects) both represent acceptable and robust designs for studies of coaching outcomes. We therefore only include studies that adopt one or other of these designs in our meta-analysis. For analytic completeness, we compare the effect size of coaching in studies that adopt each methodology. Theeboom *et al.* (2014) found stronger effects for within-subjects designs compared to between-subjects designs and speculated that this may be due to the increased level of control over potential bias and confounds in between-subjects designs, reducing the magnitude of effects compared with within-subjects designs. For example, inclusion of control groups allows for control of the natural maturation of participants over time and selection effects in sampling. Following this reasoning, we hypothesize:

> *H2:* Coaching will demonstrate stronger positive effects on overall outcomes in studies using within-subjects research designs compared to between-subjects designs.

Practice moderators of coaching effectiveness

In addition to the meta-analyses of the effectiveness of coaching in relation to specific criteria, we further extend the scope of our study by examining potential practice moderators of coaching effectiveness. There are a variety of possible tools and techniques that specific coaches may use, which might be considered method factors in the practice and implementation of coaching. As with studies of training interventions (e.g., Bell & Kozlowski, 2010), it is reasonable to assume that these method factors in coaching have some impact on its effectiveness. Such factors therefore represent moderators of the benefits of coaching on the outcome criteria we have identified. In our meta-analysis, we were able to test how four coaching method factors (use of multisource feedback; coaching delivery format; internal vs. external coach; longevity of coaching) moderate the benefits of coaching. We selected these moderators based on their coverage in the practical literature on coaching and because our literature search yielded studies that permitted their analyses. However, we do acknowledge that these are not the only potential practice moderators in the practice of coaching, a point that we develop further in our discussion of findings.

Multisource feedback

Coaching is frequently accompanied by multisource feedback in the research literature. Coaching is often used to facilitate the interpretation of multisource feedback (Luthans & Peterson, 2003; Smither, London, Flautt, Vargas, & Kucine, 2003), and in some studies (including some in our meta-analysis), changes in multisource evaluations over time serve as an outcome measure for coaching (Kochanowski, Seifert, & Yukl, 2010). Previous research has shown that multisource feedback can be an effective method of improving work performance in its own right (Kluger & DeNisi, 1996; Smither, London, & Reilly, 2005). The utility of multisource feedback may also be considered in relation to evidence that feedback more generally has performance benefits. For example, performance feedback has been shown to influence goal regulation. Ilies and Judge (2005) and Johnson (2013) found that both evaluative and objective feedback had significantly positive impacts on task performance. Seifert, Yukl, and McDonald (2003) found that feedback had a positive impact on outcomes when given in conjunction with a facilitator during a workshop rather than as written feedback alone.

Given that the purpose of coaching is to facilitate self-insight and reflection, it seems logical that combining coaching with multisource feedback is likely to lead to a greater impact on outcomes than coaching alone. Moreover, the discursive exploration that comes from coaching is likely to represent a more effective presentation of the multisource feedback than simple written results, further enhancing the benefits to the coaching intervention. We hypothesize that:

> *H3:* Coaching utilizing multisource feedback as part of the coaching process (rather than solely as an outcome measure) will demonstrate a larger effect size than coaching without multisource feedback.

Format of coaching

Coaching is most commonly conducted face-to-face; however, the use of alternative formats such as videophone, telephone, and internet coaching is also prevalent (collectively referred to as e-coaching; the use of technology to conduct coaching). Research in the field of health psychology has found that telephone discussion is an effective way of delivering coaching for lifestyle change in older men (Aoun, Osseiran-Moisson, Shahid, Howat, & O'Connor, 2011) and for reducing people's health risks (Terry, Seaverson, Staufacker, & Gingerich, 2010). However, a study by Yan, Wilber, and Simmons (2011) found that when compared to telephone-delivered health coaching, face-to-face coaching was more effective at improving exercise performance in high-risk older adults.

A study by Wentz, Nyden, and Krevers (2012) demonstrated that internet-based coaching led to improvements in self-efficacy and subjective quality of life in a sample of young people with autism spectrum disorder and/or attention-deficit/hyperactivity disorder. Also focusing on the use of internet-mediated development, research has demonstrated the utility of e-mentoring for a variety of outcomes (Ensher, Heun, & Blanchard, 2003; Ensher & Murphy, 2007). For example, in a student sample, de Janasz and Godshalk (2013) found that e-mentoring was positively related to a variety of learning outcomes and Kyrgidou and Petridou (2013) found that e-mentoring of a sample of women entrepreneurs had a positive impact on mentees' knowledge, skills, and attitudes.

Although these findings taken together suggest the developmental support can be effective when provided in a variety of formats, there are no studies that directly compare

delivery formats for workplace coaching. Indeed, studies often combine approaches such as telephone and face-to-face, such that the format of the coaching may be considered mixed method or 'blended' format (e.g., Bozer & Sarros, 2012). We have earlier emphasized the dependence of coaching on the formation of a helping relationship, which must be established without *a priori* foundations (i.e., there is no existing relationship or power or status dynamic on which to build). We therefore propose that the relational nature of coaching is likely to be more effectively fostered in a face-to-face personal environment than in e-coaching formats. Studies in our meta-analyses enabled us to compare face-to-face with blended formats (which all comprised a mixture of face-to-face and telephone coaching). We therefore hypothesize:

> *H4:* Face-to-face coaching will demonstrate a larger effect size than blended (i.e., blending face-to-face with telephone coaching).

Type of coach

Coaches may either be external or internal to the coachee's employing organization. Internal coaches are different from managers and supervisors (who may also give performance feedback and suggest ways to improve as part of performance management) because they do not have supervisory responsibilities for the coachee. External coaches are often engaged as consultants to the organization to work with specific employees. Whilst again there are no published organizational studies of this method factor, a study by Sue-Chan and Latham (2004) in an educational context found that students receiving coaching from an external coach achieved significantly higher course grades than students who had been coached by a peer. Students also rated the external coaches as more credible than the peer coaches.

In the context of work, Higgins and Kram (2001) propose that the changing nature of careers is an underlying reason why individuals are likely to need to seek developmental support from outside the organization. As careers become more protean, acquisition of skills and perspectives from outside the employing organization becomes more critical. This is underscored by Arthur, Khapova, and Wilderom (2005) who identify the importance of extra-organizational support (through coaching or mentoring) in career success from the boundaryless career perspective. External coaches have the advantage in this respect of offering the coachee support that is not restricted to the boundaries of the employing organization.

A counterview from the mentoring literature rather supports the superiority of internal mentors. Haggard, Dougherty, Turban, and Wilbanks (2011) suggest that mentors working externally to the mentees organization cannot provide the full range of career assistance functions that an internal mentor could. In particular, the external mentor is unlikely to be able to provide sponsorship, protection or challenging assignments for their mentees. This observation was supported by Murrell, Blake-Beard, Porter, and Perkins-Williamson (2008) who suggest that despite the benefits of external mentors, an internal mentor provides a number of functions such as opportunities for exposure, visibility to key management and coaching around political landmines in a way that an external mentor could not. However, it is questionable whether these benefits necessarily transfer to coaching. Unlike mentoring, coaching does not rely on the organizational and career experience and expertise of the coach with respect to the work of the coachee, and so effective achievement of development objectives is not necessarily predicated on

knowledge of the internal workings of the organization. The advantages of working with an external coach may therefore outweigh the benefits afforded by an internal coach.

We propose that compared to those working with internal coaches, coachees who receive coaching from an external coach may feel more confident in the wider credibility and perspective of the coach, and also that the coaching is more likely to be confidential. Releasing development from the boundaries and constraints of organizational politics and barriers may also be beneficial in working through development objectives. In combination, these factors are likely to have a positive impact on the outcomes of the coaching. We hypothesize that:

> H5: *Coaching provided by an external coach will demonstrate greater effect sizes than coaching provided by an internal coach.*

Coaching schedule

Research examining the effects of the number of coaching sessions or the overall longevity of coaching interventions is limited. This is despite calls to address these questions (e.g., Smither, 2011). Theeboom *et al.* (2014) tested for the moderation of the number of coaching sessions in their meta-analysis on coaching and found that a greater number of coaching sessions did not significantly impact on outcomes. Theeboom *et al.* propose that this counterintuitive effect may be due to individuals with less serious or less complex issues needing fewer coaching sessions and experiencing more positive effects of coaching than individuals with more serious or complex issues. These findings are consistent with the mentoring literature. Ensher, Thomas, and Murphy (2001) found that the longevity of the mentoring relationship had a small but significant association with social support, but found no relationships with vocational support, role-modelling support, reciprocity (i.e., level of give and take between mentor and mentee), satisfaction with mentor, job satisfaction, and perceived career success. A possible explanation for this is that the mentor themselves (i.e., their success and expertise) is more important than the longevity of the relationship (Tonidandel, Avery, & Phillips, 2007).

Notably, the above studies tested only for linear effect of coaching schedule variation (number of sessions and longevity of the relationship). This would prevent detection of nonlinear relationships, which may occur if, for example, the impact of number of sessions or longevity plateaus after a certain point. The weight of the research evidence suggests on balance that coaching schedule does not have an impact on coaching outcomes, but nevertheless we tested for moderation (including for curvilinear effects) in our analyses. Given that we did not expect to observe effects, we did not set a hypothesis regarding these moderation analyses.

Method

Literature search

A range of search strategies were utilized to identify relevant published and unpublished studies (including doctoral dissertations). Firstly, various electronic databases were searched including ProQuest, EBSCO, Emerald Full Text, JSTOR Business, SAGE Journals Online, ScienceDirect, SwetsWise, Taylor and Francis, Wiley Online Library, and Oxford Journals. The following search terms were used: (*coaching*) and (*effectiveness* or *outcome* or *impact* or *influence* or *evaluation*). In addition to this database search,

frequent contributors to coaching research literature were contacted directly by e-mail to ensure that any unpublished data or work in progress were included in the meta-analysis. Finally, a manual review of the reference lists of all of the articles identified in the database search and also the reference list of all relevant reviews was completed. This initial search identified a total of 54 studies.

Criteria for inclusion

To be included in the meta-analysis, studies had to meet six criteria. First, the study had to examine workplace coaching effectiveness (i.e., studies in which coaching was provided with the objective of generating affective, cognitive, skill-based or results outcomes at work). Studies were included if they adequately described coaching activity (i.e., one-to-one development intervention based on a coach–coachee relationship) in a work context, and evaluated that activity in terms of its effectiveness. Studies that measured the impact of coaching on non-work outcomes (such as exercise or healthy eating) were excluded, as were studies where coaching was provided by a line manager. Secondly, we included only studies that evaluated coaching using the within- and between-subjects designs we reviewed earlier. Thirdly, studies had to have been conducted within an organizational setting, to ensure all participants were employed working adults. Fourthly, studies needed to report sample sizes. Fifthly, a d statistic or other statistic (e.g., means and standard deviations) that could be converted into a d statistic must have been reported between coaching and the outcome variable. Finally, the dependent variable or coaching outcome had to be measured at the individual level of analysis.

Data set

Of the 54 studies identified in the literature search, 17 met our inclusion criteria ($n = 2,267$ individuals). The average sample size of these studies was 133 with a range from 14 to 1,361. Seven studies were conducted in the United States; two in the United Kingdom; three in Australia; two in Norway; one in Egypt; one in Israel; and one in Denmark. The studies were conducted in different organization types and industries including service, manufacturing, construction, and public/government sectors. Occupations of participants were varied; however, the majority held senior management ($k = 5$, $n = 1,527$) or management roles ($k = 5$, $n = 326$). Other occupations were nurses ($k = 1$, $n = 120$), high school principals ($k = 1$, $n = 8$), teachers ($k = 1$, $n = 44$), construction foremen ($k = 1$, $n = 51$), and various occupations within single studies ($k = 2$, $n = 69$). The earliest study was reported in 1997, and the other studies were reported after 2003 (2003 = two studies; 2005 = two studies; 2006 = one study; 2007 = one study; 2009 = three studies; 2010 = four studies; 2011 = one study and 2012 = one studies). All studies included in the meta-analysis are summarized in Table 2 and listed in the references marked with asterisks (*).

Description of variables

All useable studies were coded on a number of specific variables including source of study, sample size, year of publication, occupation of coachee, and organizational context. In addition, we developed a coding scheme to code for the moderators highlighted in our hypotheses. Firstly, research design was coded for (i.e., between-subjects design, 14 studies; within-subjects design, three studies). In between-subjects studies, outcome data

Table 2. Summary of coaching effectiveness studies included in the meta-analysis

Authors	Sample	Country	Organizational context	Occupation of participants	Outcomes measured	Research design	Coaching technique	Longevity of intervention (weeks)[a]	Duration of sessions (min)	Number of sessions[b]	Multisource feedback	Format of coaching	Type of coach
Bozer and Sarros (2012)	96	Israel	Various	Management	Affective Skill-based	Between-subjects	Cognitive-behavioural	15.2	30–120	6-8	Not specified	Face-to-face, telephone and e-mail	External
Bright and Crockett (2012)	115	United States	Various	Various	Affective	Between-subjects	Not specified	4	30	1	No	Telephone	External
Cerni et al. (2010)	14	United States	Education	Principals	Skill-based	Between-Subjects	Epstein's constructive thinking programme (1998)	10	60	10	No	Face-to-face	External
Evers et al. (2006)	60	United States	Government Agency	Management	Affective	Between-subjects	GROW model (Whitmore, 1992)	17.3	Not specified	Not specified	No	Not specified	External
Finn (2007)	17	Australia	Not specified	Senior management	Affective Skill-based	Between-subjects	Not specified	12	60	6	Yes	Face-to-face and telephone	External
Grant et al. (2009)	41	Australia	Health	Executives and senior managers	Affective Results	Between-subjects	Cognitive-behavioural solution focussed (Grant, 2003) GROW model (Whitmore, 1992)	10	Not specified	4	Yes	Face-to-face and telephone	External
Grant, Green, et al. (2010)	44	Australia	Education	Teachers	Affective Skill-based Results	Between-subjects	Cognitive-behavioural solution focussed (Grant, 2003) GROW model (Whitmore, 1992)	20	Not specified	10	Yes	Face-to-face	External
Gyllensten and Palmer (2005)	31	United Kingdom	Finance	Not specified	Affective	Between-subjects	Not specified	34.7	Not specified	Not specified	No	Face-to-face	Internal
Kines et al. (2010)	51	Denmark	Construction	Construction foremen	Skill-based	Between-subjects	Not specified	42	Not specified	8	No	Face-to-face	External
Kochanowski et al. (2010)	30	United States	Retail	Store managers	Skill-based	Between-subjects	Not specified	6	30-60	6	Yes	Face-to-face and telephone	External
	20	United States	Manufacturing	Various		Within-subjects	Not specified	13	Not specified	1	Yes	Face-to-face	External

Continued

Table 2. (*Continued*)

Authors	Sample	Country	Organizational context	Occupation of participants	Outcomes measured	Research design	Coaching technique	Longevity of intervention (weeks)[a]	Duration of sessions (min)	Number of sessions[b]	Multisource feedback	Format of coaching	Type of coach
Luthans and Peterson (2003)					Affective Skill-based								
Moen and Allgood (2009)	127	Norway	Not specified	Executives and managers	Affective	Between-subjects	Not specified	52	Not specified	Not specified	No	Not specified	Not specified
Moen and Skaalvik (2009)	19	Norway	Not specified	CEOs	Affective	Between-subjects	Not specified	13	60–90	7	No	Face-to-face and telephone	External
Olivero et al. (1997)	31	United States	Health	Management	Results	Within-subjects	Not specified	8.7	60	8	No	Face-to-face	Internal
Smither et al. (2003)	1,361	United States	No specified	Senior managers	Skill-based	Between-subjects	Not specified	Not specified	Not specified	2–3	Yes	Face-to-face	External
Taie (2011)	120	Egypt	Health	Nurses	Skill-based	Between-subjects	Not specified	Not specified	Not specified	Not specified	No	Face-to-face	Internal
Toegel and Nicholson (2005)	89	United Kingdom	Finance	Senior managers	Skill-based	Within-subjects	Not specified	39	90	2	Yes	Face-to-face and telephone	External

Note. [a]Where the longevity was specified in months, we converted to weeks based on an assumption of average 4.33 weeks per month. For our analyses, if a range of longevity was reported, we coded as the mid-point of the range.

[b]Where a range of number of sessions was reported (e.g., 6–8), we coded the data as the mid-point of the range (e.g., 7).

were collected for all participants at time 1. The coaching group then received the coaching intervention, whilst the control group received no coaching. Outcome data were then collected for all participants at time 2. Between-subjects studies compared the change in outcomes from time 1 to time 2 for both groups (e.g., Evers, Brouwers, & Tomic, 2006; Kochanowski *et al.*, 2010; Taie, 2011). In within-subjects studies, all participants received coaching and outcomes were recorded before and after the coaching. Within-subjects studies examined the change in outcomes from time 1 to time 2 (e.g., Luthans & Peterson, 2003; Olivero *et al.*, 1997; Toegel & Nicholson, 2005).

We also coded for classification of outcome in accordance with the model outlined earlier, and summarized in Table 1: Outcome measures were classified as affective (e.g., job satisfaction; Luthans & Peterson, 2003) (10 studies), skill-based (e.g., competency skills; Taie, 2011) (10 studies), or results (e.g., productivity; Olivero *et al.*, 1997) (three studies). All three studies within the results category measured individual-level rather than team-level or organizational-level results. Note that some studies utilized multiple outcome measures falling within different outcome categories. Further, we coded for whether coaching was accompanied by multisource feedback. Studies were classified as either utilizing multisource feedback as part of the coaching process (e.g., Smither *et al.*, 2003) or coaching where multisource feedback was not used (e.g., Bright & Crockett, 2012). Studies that utilized multisource feedback solely as an outcome measure and not part of the coaching process were included in the latter category as the details of the multisource feedback were not disclosed to the coachee (e.g., Cerni, Curtis, & Colmar, 2010). We coded the coach as either external to the organization (e.g., a consultant as a coach; Grant, Green, & Rynsaardt, 2010) or internal to the organization (e.g., Gyllensten & Palmer, 2005). In the instances where an internal coach was used, the coach did not have line management responsibility for the coachee. Our coding for coaching format classified coaching as either face-to-face (e.g., Kines *et al.*, 2010) coaching or 'blended' coaching format. Blended coaching format was the category created for any format not solely face-to-face. Studies were grouped in this way due to the small number of studies present that used a format other than traditional face-to-face. Five studies utilized a combination of telephone with face-to-face coaching (Finn, 2007; Grant *et al.*, 2009; Kochanowski *et al.*, 2010; Moen & Skaalvik, 2009; Toegel & Nicholson, 2005), and one study used a combination of telephone, face-to-face, and email coaching (Bozer & Sarros, 2012). Finally, we coded for the number of coaching sessions (mean = 5.56; median = 6.00; SD = 3.18; range = 9) and the longevity, in weeks, of the coaching intervention (mean = 18.80; median = 12.00; SD = 13.85; range = 44). We also coded for the duration of each coaching session; however, we did not explore this variable as Table 2 shows the duration of coaching sessions was relatively invariant.

Coding accuracy and inter-rater agreement

All studies were coded independently by two coders. The first was the first author, and the second, a doctoral student with experience of meta-analytic techniques (who was briefed on the coding procedure, and provided with the coding scheme and instructions). No discrepancies between coders were encountered.

Meta-analytic procedure

The meta-analysis relied on the widely used Hunter and Schmidt (1990, 2004) approach: A random effects model that accounts for sampling bias and measurement error. We

calculated sample-weighted effect sizes and corrected for reliability of dependent variables. Given that we were interested in the treatment effects of coaching, we converted all effect sizes obtained from the primary studies to r, conducted a meta-analysis on r, and then converted the final results to d (Hunter & Schmidt, 2004). Missing artefact values (i.e., reliability of dependent variables) were estimated by inserting mean values across the studies where information was not given or could not be obtained from the authors after contacting them, as recommended by Hunter and Schmidt (2004). Objective performance data were not corrected for unreliability because it has been frequently argued that measures based on objective performance data are unbiased (Riketta, 2005), and also as no procedure is currently available to correct for unreliability of such measures.

Additionally, we report the 90% confidence intervals (90% CI) of the sample-weighted effect size d, and the 80% credibility intervals (80% CV) of the corrected population d statistic δ. Confidence intervals estimate variability in the sample-weighted effect size that is due to sampling error, whereas credibility intervals estimate variability in the individual correlations across studies that are due to the moderator variables (Whitener, 1990). If the 90% confidence interval of the sample-weighted effect size does not include zero, we can be 95% confident that the sample-weighted effect size is different from zero. Confidence intervals can also be used to test whether two estimates differ from each other; two estimates are considered different when their confidence intervals are non-overlapping.

If the 80% credibility interval of the corrected population effect size is large and includes zero, it indicates that there is considerable variation across studies and moderators are likely operating. To further corroborate that moderators are present, we assessed whether sampling error and error of measurement accounted for more than 75% of the variance between studies in the primary estimates (Hunter & Schmidt, 1990). To do this, we report the percentage of variance accounted for in the corrected population effect size by sampling and measurement error (% VE). Moderators are assumed to be operating when sampling and measurement error account for <75% of the variance. The majority of the moderators in our study are categorical moderators (only the number of coaching sessions and longevity of the intervention were continuous moderators). The categorical moderators were computed using Hunter and Schmidt (1990, 2004) subgroup analyses techniques by conducting separate meta-analyses at each of the specified moderator level. To examine whether there is significant difference between the mean corrected effect sizes of sublevels of the hypothesized moderator variable, we compared their confidence intervals as discussed above. To test the continuous moderators, we conducted weighted least squares (WLS) regression as suggested by Steel and Kammeyer-Mueller (2002). Unlike subgroup analysis of moderators as suggested by Hunter and Schmidt (2004), this approach avoids artificial categorization of continuous moderator variables. For the WLS regression, we have regressed the uncorrected correlations of overall coaching effectiveness in SPSS on each moderator variable using the inverse square root of the sampling error for each correlation as the weighting factor (as specified by Steel & Kammeyer-Mueller, 2002) to see whether the moderator explained variance in the uncorrected correlations of overall coaching effectiveness.

Results

Table 2 presents an overview of the coaching effectiveness studies included in our analysis. We report our results in two sections. Firstly, we report overall coaching

effectiveness and with respect to different outcome criteria. This is followed by exploration of the impact of moderator variables.[1]

Criterion effects of coaching

Coaching had a positive effect on all outcomes (supporting hypothesis 1). As can be seen in Table 3, none of the 90% CI included zero and coaching had a positive effect on overall outcomes ($\delta = 0.36$, 90% CI [0.16, 0.50]).

We distinguished between affective, skill-based, and individual-level results outcomes in our analyses. Table 3 suggests that coaching had positive effects on affective outcomes ($\delta = 0.51$, 90% CI [0.17, 0.78]), skill-based outcomes ($\delta = 0.28$, 90% CI [0.07, 0.44]), and individual-level results outcomes ($\delta = 1.24$, 90% CI [0.88, 1.47]). The non-overlapping confidence intervals of these three outcomes suggest that coaching has a stronger effect on individual-level results outcomes than on affective and skill-based outcomes. There are no differences with regard to affective and skill-based outcomes.

The low amount of variance accounted for by artefacts, and the large credibility intervals around the effects of coaching on overall outcomes (17.68%, 80% CV [−0.16, 0.97]) suggest that moderating variables are operating.

Moderators of coaching effectiveness

Research design

Most of the studies within our sample were between-subjects design ($k = 14$, $n = 2,109$). A small group of studies used a within-subjects design ($k = 3$, $n = 158$). Coaching had a positive effect on outcomes for both types of research design. Table 3 suggests that coaching had a greater effect on outcomes for within-subjects ($\delta = 0.57$) compared to between-subjects design ($\delta = 0.35$). However, the 90% CI were overlapping ([0.43, 0.62]; [0.12, 0.51]), which suggests that the two effects are not different; therefore, research design does not appear to moderate the effects of coaching on outcomes (therefore, hypothesis 2 was not supported).

Multisource feedback

We examined the effects of whether the coaching included multisource feedback as part of the coaching process. Table 3 suggests that coaching had a positive effect on outcomes for both groups; however, this effect was greater on outcomes when coaching was provided without multisource feedback ($\delta = 0.88$ vs. $\delta = 0.21$). The non-overlapping confidence intervals suggest that coaching has a stronger effect when it is provided without multisource feedback ([0.40, 1.29]; [0.12, 0.26]); therefore, hypothesis 3 was not supported.

---

[1] *An anonymous reviewer of an earlier draft of our paper rightly noted that one study (Smither et al., 2003) had a markedly high sample size compared to others in our meta-analysis. Based on the idea that there is a trade-off between weighting individual effect sizes by their sample size and individual effect sizes obtained from very large samples overly influencing the weighted average effect size, it has been suggested to limit the N of any primary study to 500, and to substitute any sample size with a value larger than that with 500 (Riketta, 2005). When we applied this modification, the effect sizes for the analyses that included the Smither et al. study increased slightly. However, we observed no other changes, and our conclusions regarding our hypotheses were unaffected. Parsimoniously, we therefore report the meta-analytic results with the original sample size for the Smither et al. study.*

Table 3. Meta-analytic results

Variable	k	n	d	90% CI Lower	90% CI Upper	δ	SD_δ	% Var. acc. for	80% CV Lower	80% CV Upper
Overall effectiveness	17	2,267	0.33	0.16	0.50	0.36	0.42	17.68	−0.16	0.97
Affective outcomes	10	592	0.46	0.17	0.78	0.51	0.55	22.84	−0.15	1.39
Skill-based outcomes	10	1,784	0.26	0.07	0.44	0.28	0.35	19.02	−0.16	0.76
Individual-level results outcomes	3	116	1.15	0.88	1.47	1.24	0.00	100.00	1.25	1.25
Research design										
Overall outcomes – between-subjects	14	2,109	0.31	0.12	0.51	0.35	0.44	14.88	−0.20	0.99
Overall outcomes – within-subjects	3	158	0.53	0.43	0.62	0.57	0.00	100.00	0.56	0.56
Multisource feedback (MSF)										
Overall outcomes – MSF not used	9	569	0.80	0.40	1.29	0.88	0.82	12.53	0.00	2.73
Overall outcomes – MSF used	7	1,620	0.19	0.12	0.26	0.21	0.00	100.00	0.20	0.20
Format of coaching										
Overall outcomes – face-to-face coaching	8	1,691	0.27	0.06	0.48	0.29	0.36	15.49	−0.16	0.80
Overall outcomes – blended coaching	6	274	0.25	0.09	0.43	0.28	0.00	100.00	0.24	0.24
Type of coach										
Overall outcomes – external coach	13	1,958	0.18	0.11	0.26	0.20	0.00	100.00	0.20	0.20
Overall outcomes – internal coach	3	182	1.27	0.65	2.36	1.40	0.89	14.23	0.54	3.81

Note. k = number of studies; n = number of respondents; d = sample-weighted mean effect size; 90% CI = 90% confidence interval of the d; δ = corrected population d statistic; SD_δ = standard deviation of the corrected population d statistic; % var. acc. for = percentage of variance attributed to sampling error and artefact corrections; 80% CV = 80% credibility interval of the δ.

Format of coaching

We compared the effects of coaching provided in the face-to-face format with blended formats. Table 3 shows that both effect sizes were positive and neither coaching format demonstrated stronger criterion effects; the two effects on outcomes are similar and their 90% CI are overlapping ($\delta = 0.29$, [0.06, 0.48] vs. $\delta = 0.28$ [0.09, 0.43]; therefore, hypothesis 4 was not supported).

Type of coach

We examined effects for coaches that were employed internally by the organization, and compared with effects for external consultant coaches. Table 3 shows that the effects of coaching on outcomes were weaker when the coach is external to the organization ($\delta = 0.20$ vs. $\delta = 1.40$) and that the 90% CI were non-overlapping ([0.11, 0.26]; [0.65, 2.36]). This suggests that contrary to our prediction, the effects of coaching on outcomes were weaker for external coaches in comparison with internal coaches (therefore, hypothesis 5 was not supported).

Coaching schedule

We tested whether the longevity in weeks of the coaching intervention and number of coaching sessions moderated overall coaching effectiveness. As longevity of the coaching intervention and number of coaching sessions might have a decreasingly positive effect on coaching effectiveness, we tested for both linear and curvilinear effects. To test for linear effects, we entered the moderator variable as predictor variable, and to test for curvilinear effects, we entered the moderator variable and the squared term of the moderator variable as predictor variables. None of the effects were significant (linear effect of longevity in weeks of the coaching intervention: $B = .01$, $SE = 0.00$, ns, $R^2 = .14$; curvilinear effect of longevity in weeks of the coaching intervention: $B = .00$, $SE = 0.00$, ns, $R^2 = .18$; linear effect of number of coaching sessions: $B = .01$, $SE = 0.02$, ns, $R^2 = .01$; curvilinear effect of number of coaching sessions: $B = .01$, $SE = 0.01$, ns, $R^2 = .07$). Therefore, neither longevity in weeks of the coaching intervention nor number of coaching sessions moderated overall coaching effectiveness.

Discussion

In the present study, we reported results of a meta-analysis to synthesize the existing research on coaching effectiveness at work in order to understand the impact that coaching has on a variety of workplace outcomes. We examined outcome criteria classified on a framework modelled on the learning, training, and development literature and tested whether particular techniques or tools of coaching moderated its effectiveness. Our analyses demonstrated that for all outcomes, coaching had a positive impact. Effect sizes nevertheless varied for different types of outcome criteria, and based on some specific applied method features of coaching. Collectively, our findings have important implications for research and practice of coaching.

Criterion effects of coaching

To align our analyses with literatures on training and development, and HRM, we proposed a criterion framework for the outcomes of coaching. We hypothesized that coaching would have an overall positive impact on all outcome criteria, specifically in our meta-analyses for affective, skill-based, and individual-level results outcomes. We found that in the combined analyses of all 17 studies, coaching demonstrated a positive impact on outcome criteria ($\delta = 0.36$). Within this finding, there was variation for different types of outcome criteria ranging from 0.28 to 1.24. These effect sizes appear to be comparable to those obtained for other types of organizational interventions. For example, in their meta-analysis of the training effectiveness literature, Arthur *et al.* (2003) obtained effect sizes ranging from 0.60 to 0.63, whereas Powell and Yalcin (2010) found a smaller effect size of 0.24 for managerial training interventions. Looking at wider interventions, findings from Arthur *et al.* (2005), Smither *et al.*'s (2005) meta-analysis of multisource feedback on performance reported much smaller effect sizes ranging from 0.05 to 0.15. Therefore, compared to the impact of a popular intervention such as multisource feedback, coaching appears to have larger and more consistent positive effects on outcome criteria.

Interestingly, the largest effect size in our study (1.24) was observed for individual-level results outcomes, a key organizational focus representing translation of learning through to performance benefits. Evidence that coaching has a significantly positive impact on individual-level results in particular indicates that businesses can expect positive performance and impact improvements from investment in coaching. By contrast, Powell and Yalcin (2010) reported the smallest effects for results criteria in their meta-analysis of managerial training interventions explaining their findings with reference to potential issues in training transfer, leading to a smaller impact of training on results outcomes. We earlier described how coaching encourages development activities that are personalized to individual need, and carried out in day-to-day work. These may therefore serve to promote development that is directly relevant to the workplace, and is therefore more straightforward for the coachee to implement in their performance behaviour. These processes may promote transfer of learning to work activity resulting in improved performance results, a proposition that appears to be supported in our findings, representing a potential advantage of coaching over other forms of training.

We did not find moderation of the criterion effects of coaching by research design, contrary to our hypothesis 2, and the results of Theeboom *et al.* (2014). Although the effect size for within-subjects designs was higher than for between-subjects designs, the difference was non-significant. We must conclude therefore that whilst there is some evidence that research design may influence criterion effect sizes in studies of coaching, the specific nature or importance of those effects remains to be confirmed.

Practice moderators of coaching criterion effects

Coaches often employ different practice factors in the form of techniques or tools into their coaching approaches, and we tested the moderating effects of some of these in our analyses. Firstly, our analyses revealed that when controlling for moderation effects of these practice factors, the positive effects of coaching were maintained. Moreover, we did not find significant moderation by the use of face-to-face compared to blended coaching. This finding, although contrary to our original prediction, is consistent with findings from the field of e-mentoring (de Janasz & Godshalk, 2013; Kyrgidou and Petridou (2013) who demonstrated that e-mentoring has a positive impact on academic and workplace outcomes. This finding has positive implications for the practice of coaching. One of the

main advantages of providing coaching in alternative formats such as telephone or e-mail is the efficiencies in terms of cost and time commitment. Our finding that both face-to-face and blended coaching were effective provides some justification for applying blended formats in order to extend its accessibility to a wider audience. Notwithstanding this point, our data do not permit us to draw conclusions about the effectiveness of e-coaching (i.e., the use of technology such as internet or telephone to conduct coaching) alone. Only one study used solely e-coaching (in which coaching was conducted exclusively by telephone), so we were unable to compute a meta-analytic effect. Future research on the effectiveness of e-coaching is therefore clearly warranted.

Our analysis showed that there was no significant moderation of outcomes by the number of coaching sessions or the longevity of the coaching intervention, consistent with the previous findings of Theeboom *et al.* (2014). Our tests for curvilinear effects also indicated that there is not a plateauing of the impact of coaching (by which additional sessions or weeks of intervention would make an impact up until a certain point, after which more coaching would cease to result in improved impact). Taken together, our data indicate that even short-term coaching has a beneficial impact. These findings once again have positive implications for the practice of coaching as they suggest that shorter and potentially more cost effective coaching interventions are likely to be effective. However, like previous studies (e.g., Theeboom *et al.*, 2014), we caution that future research should qualify this finding by factoring in the severity of the presenting employee development issues at the outset of coaching. Some employee development issues may require a greater number of sessions, and a longer term coaching intervention than others. Data analysed in this study did not permit this possibility to be tested.

Two significant moderator variables were identified in our meta-analyses: Use of multisource feedback, and type of coach. In both cases, the results ran counter to our hypotheses.

We hypothesized that coaching combined with multisource feedback would have a greater effect on outcomes. However, our findings showed that coaching without multisource feedback actually had a significantly stronger positive impact on outcomes. This finding may possibly be explained by previous research findings on the impact of feedback (including multisource feedback) on performance. For example, in their meta-analysis of the effects of performance feedback, Kluger and DeNisi (1996) reported that whilst on average feedback interventions resulted in higher performance, over a third of the studies included in their analyses actually reported that feedback resulted in lower performance. They concluded that their findings were meaningful, and could not be explained by sampling error, feedback sign or existing theory, but speculated that attentional processes may help to explain the negative impact of feedback on performance. In the context of coaching, feedback (especially negative feedback) received by the coachee in the coaching session may divert attention so that the coachee becomes pre-occupied with the content of the feedback. This could leave insufficient attentional resources to engage in the subsequent coaching process, limiting the positive benefits of the session. An additional factor may be the content of the multisource feedback. Such instruments are often focused on leadership behaviour, which may be rather distal from the development objectives of the coachee. Collectively, these potentially negative mechanisms associated with feedback and its content, may explain why in our meta-analyses, we observed weaker performance benefits when multisource feedback was combined with coaching.

We also hypothesized that external coaches would have a greater impact on outcomes as they would provide an external perspective, unrestricted by the constraints of the

employing organization, and more consistent with contemporary boundaryless models of work and careers. This was not confirmed in our findings, which rather showed that although coaching by both internal and external coaches was beneficial for learning and performance, the effects of coaching by internal coaches had a stronger effect compared to external coaches. This is an interesting finding given the growing scale of the coaching industry, which is strongly based around a model of contracted external (e.g., consulting) coaches. Some potential explanations as to why internal coaches appear to be more effective than external coaches are provided by Strumpf (2002), who questions the assumption that bringing in an external coach is necessarily the best option for organizations. Strumpf suggests that the choice to use either an internal or external coach is dependent on a number of factors. These factors include a cultural bias and readiness, as some organizations prefer to use 'homegrown' solutions versus those generated by external consultants and a strong, strategically placed head of human resources who may provide an excellent role model for internal coaches and consequently increase the credibility of an internal coach. A further potential explanation is that internal coaches inevitably have a better understanding of the organization's culture and climate and may therefore be better placed to enable the coachee to be more productive in their specific workplace (i.e., by setting goals in such a way that organization-specific barriers or facilitators to their achievement are realistically discussed, addressed and incorporated into development). Our earlier arguments that an internal perspective would be less critical for coaching as compared to mentoring, for example, are not borne out in the data we analysed. Our interpretations here should be viewed appropriately cautiously given the relatively low number of studies using internal coaches.

Implications for research and future directions

Our study represents an important advance in the evidence base for the effectiveness of coaching in organizations. To consider the implications of our findings for future research, we highlight three themes emerging from our analyses: (1) criterion effects measured in primary research on coaching; (2) practice methodology in studies of coaching; (3) theoretical explication of the processes of coaching.

Considering first the criterion effects of coaching, we propose that our framework of criterion effects could be used by researchers to organize more clearly the criteria that are measured in coaching evaluating studies. We presented some methods of measurement for each of the criteria we include in our framework, which could guide researchers in choosing measurement methods. Our framework also highlighted gaps in the literature for certain kinds of outcome criteria. Specifically, none of our included studies examined cognitive, or team or organizational-level results outcome criteria, areas where further studies are clearly warranted. Moreover, future research might examine how the relative impact of coaching on different kinds of criteria, and compare this with other forms of learning and development intervention (e.g., training).

Second, we acknowledged earlier that our selection of practice moderators was somewhat governed by those factors that were described and operationalized in the studies we examined. This observation speaks to a broader limitation of many studies of coaching effectiveness, namely a lack of detail in the descriptions of coaching interventions employed. Failing to completely describe coaching techniques and approaches employed in empirical studies prevents later classification in meta-analyses such as ours. We therefore encourage greater thoroughness on the part of researchers in this area to specify precisely the nature and format of coaching employed in empirical

studies. Such specification should include description of who provided coaching, to whom, in what format, for how long, over how many individual sessions, using which coaching approaches or models, and including which (if any) tools or techniques.

Greater level of detail in the specification of coaching undertaken in empirical studies is also essential for the development of theory and understanding of the processes of coaching. Like Theeboom *et al.* (2014), and echoing similar observations for all training and development intervention literatures (Ford *et al.*, 2010), we strongly feel that there is a need for the development of theory concerning why, how and in what ways coaching leads to the positive effects we reported in this study. We earlier argued that potential processes that are common features of coaching are goal-setting (Locke & Latham, 2002), encouragement of experiential learning (Kolb, 1984), and setting of development activities that are completed in day-to-day work activity. Coaching may therefore be considered to be a high-fidelity form of training and development intervention (e.g., Baldwin & Ford, 1988). However, it remains impossible to test these or other processes in meta-analyses because individual studies do not consistently describe, for example, if objectives or goals were set, what kind of development or learning activities were recommended, or how they were formulated. There are of course methodological challenges to adequately describing multiple sessions with multiple clients, but failure to address these will continue to hold back development of the coaching literature.

With regard to future research directions, the implications of these observations are twofold. First, there is a need for experimental studies that manipulate the features of coaching to examine directly the impact of particular coaching processes. Theoretical and conceptual development should clearly accompany or precede such research. Second, all empirical studies of coaching should adequately describe in detail any particular general processes that applied to the coaching sessions under study.

Applied implications for organizations and coaches

Our study has a number of implications for practitioners and coaches in organizations. Firstly, our meta-analysis demonstrated that coaching had a positive effect on all outcome criteria we examined, providing an evidence base from which practitioners can draw confidence.

Second, our proposed framework of outcome criteria from coaching provides researchers and organizations with a method of categorizing the types of outcomes that can be expected from coaching. These can be applied in evaluation studies by practitioners, or by organizations to examine the impact of coaching.

Finally, our findings are informative for decisions about using particular tools and coaching techniques. Although our results show that overall, coaching appears to be effective irrespective of the format of the coaching and the longevity of the coaching intervention (including number of sessions), practitioners and organizations should consider carefully the use of multisource feedback, and the engagement of external coaches, both of which resulted in smaller positive effects of coaching. If multisource feedback is used, practitioners should review and apply evidence in the literature about the determinants of effective use of feedback. When engaging external coaches, organizations could ensure that a thorough familiarization process is undertaken to enable coaches to have a full and complete understanding of the organizational context of employee learning and performance.

Limitations and strengths

There are a number of limitations of our study. First, with respect to the coding of data, due to incomplete reporting of sample characteristics and coaching variables, a number of the coded variables had to be listed as either unspecified or, in the case of reliability data, estimated. Although estimation from means is consistent with recommendations for replacing other forms of missing data (e.g., Hunter & Schmidt, 1990), such replacements are less satisfactory than reported data. In some cases, lack of reporting in research articles meant that we did not have enough studies in each group to fully analyse moderators. In addition, as we comment on earlier, incomplete reporting meant that the moderators we were able to explore were limited to those included in the primary studies.

Second, due to the relatively nascent nature of coaching research, our meta-analysis included a relatively small number of studies. Whilst the number of studies included in our analyses is similar to some other meta-analyses in work and organizational psychology (e.g., Riketta, 2008), some caution is warranted in interpreting and generalizing from our results, which invite replication as the number of research studies in this area grows.

Alongside these limitations, our study has a number of significant strengths. Ours is the first systematic examination of the effectiveness of coaching to exclusively focus on coaching in organizations, marking an important advance in the literature on coaching at work. The positive findings prompt future development of the coaching literature, and coaching practices. Moreover, we align our analyses of coaching outcomes with existing taxonomies of outcomes from learning, training, and development, making our findings easy to view alongside these literatures. Our examination of moderators is particularly informative for practising coaches to understand the impact of particular tools and techniques of coaching, enabling them to make informed choices about their work.

Conclusion and final comments

At the outset of this study, we asked whether workplace coaching was effective in terms of delivering individual learning and development, and improvements in performance and results for organizations. Our meta-analysis has made significant steps in addressing the lack of systematic scientific evidence about the benefits of coaching for organizations. Our findings suggest that the answer to our question is yes, and our analyses indicated that coaching resulted in a number of key positive effects for learning and performance outcomes in the studies we examined. Our study gives support to the further development of coaching research and practice, providing evidence of the potential utility of coaching at work.

References

*Studies marked with an asterisk contained data used in the meta-analysis.

Alliger, G. M., Tannenbaum, S. I., Bennett, W., Traver, H., & Shotland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology*, *50*, 341–358. doi:10.1111/j.1744-6570.1997.tb00911.x

Aoun, S., Osseiran-Moisson, R., Shahid, S., Howat, P., & O'Connor, M. (2011). Telephone lifestyle coaching: Is it feasible as a behavioural change intervention for men? *Journal of Health Psychology*, *17*(2), 227–236. doi:10.1177/1359105311413480

Arthur, W., Bennett, W., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design & evaluation features. *Journal of Applied Psychology*, *88*, 234–235. doi:10.1037/0021-9010.88.2.234

Arthur, M. B., Khapova, S. N., & Wilderom, C. P. M. (2005). Career success in a boundaryless career world. *Journal of Organizational Behavior*, 26, 177–202. doi:10.1002/job.290

Ayres, J., & Malouff, J. M. (2007). Problem-solving training to help workers increase positive affect, job satisfaction, and life satisfaction. *European Journal of Work and Organizational Psychology*, 16, 279–294. doi:10.1080/13594320701391804

Baldwin, T. T., & Ford, J. F. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, 4, 63–105. doi:10.1111/j.1744-6570.1988.tb00632.x

Bell, B. S., & Kozlowski, S. W. J. (2008). Active learning: Effects of core training design elements on self-regulatory processes, learning, and adaptability. *Journal of Applied Psychology*, 93, 296–316. doi:10.1037/0021-9010.93.2.296

Bell, B. S., & Kozlowski, S. W. J. (2010). Toward a theory of learner-centered training design: An integrative framework of active learning. In S. W. J. Kozlowski & E. Salas (Eds.), *Learning, training, and development in organizations* (pp. 3–64). New York, NY: Routledge.

Bennett, J. L. (2006). An agenda for coaching-related research: A challenge for researchers. *Consulting Psychology Journal: Practice and Research*, 58, 240–249. doi:10.1037/1065-9293.58.4.240

Bono, J. E., Purvanova, R. K., Towler, A. J., & Peterson, D. B. (2009). A survey of executive coaching practices. *Personnel Psychology*, 62, 361–404. doi:10.1111/j.1744-6570.2009.01142.x

*Bozer, G., & Sarros, J. C. (2012). Examining the effectiveness of executive coaching on coachees' performance in the Israeli context. *International Journal of Evidence Based Coaching and Mentoring*, 10, 14–32.

*Bright, D., & Crockett, A. (2012). Training combined with coaching can make a significant difference in job performance and satisfaction. *Coaching: An International Journal of Theory, Research and Practice*, 5(1), 4–21. doi:10.1080/17521882.2011.648332

Brockbank, A., & McGill, I. (2012). *Facilitating reflective learning: Coaching, mentoring and supervision*. London, UK: Kogan Page.

Brotman, L. E., Liberi, W. P., & Wasylyshyn, K. M. (1998). Executive coaching: The need for standards of competence. *Consulting Psychology Journal: Practice & Research*, 50, 40–46. doi:10.1037/1061-4087.50.1.40

Brutus, S., Gill, H., & Duniewicz, K. (2010). State of the science in industrial and organizational psychology: A review of self-reported limitations. *Personnel Psychology*, 63, 907–936. doi:10.1111/j.1744-6570.2010.01192.x

*Cerni, T., Curtis, G. J., & Colmar, S. H. (2010). Executive coaching can enhance transformational leadership. *International Coaching Psychology Review*, 5, 81–85.

de Janasz, S. C., & Godshalk, V. M. (2013). The role of e-mentoring in Protégés' learning and satisfaction. *Group & Organization Management*, 38, 743–774. doi:10.1177/1059601113511296

Dierdorff, E. C., Surface, E. A., & Brown, K. G. (2010). Frame-of-reference training effectiveness: Effects of goal orientation and self-efficacy on affective, cognitive, skill-based, and transfer outcomes. *Journal of Applied Psychology*, 95, 1181–1191. doi:10.1037/a0020856

Eby, L. T. D. T., Allen, T. D., Hoffman, B. J., Baranik, L. E., Sauer, J. B., Baldwin, S., & Evans, S. C. (2013). An interdisciplinary meta-analysis of the potential antecedents, correlates, and consequences of protégé perceptions of mentoring. *Psychological bulletin*, 139, 441–476. doi:10.1037/a0029279

Ensher, E. A., Heun, C., & Blanchard, A. (2003). Online mentoring and computer-mediated communication: New directions in research. *Journal of Vocational Behavior*, 63, 264–288. doi:10.1016/S0001-8791(03)00044-7

Ensher, E. A., & Murphy, S. E. (2007). E-mentoring: Next generation research strategies and suggestions. In B. R. Ragins & K. E. Kram (Eds.), *Handbook of mentoring* (pp. 299–322). Thousand Oaks, CA: Sage.

Ensher, E. A., Thomas, C., & Murphy, S. E. (2001). Comparison of traditional, step-ahead, and peer mentoring on protégés support, satisfaction, and perceptions of career success: A social

exchange perspective. *Journal of Business and Psychology*, *15*, 419–438. doi:10.1023/A:1007870600459

Epstein, S. (1998). *Constructive thinking: The key to emotional intelligence*. London, UK: Praeger.

*Evers, W. J. G., Brouwers, A., & Tomic, W. (2006). A quasi-experimental study on management coaching effectiveness. *Consulting Psychology Journal: Practice and Research*, *58*, 174–182. doi:10.1037/1065-9293.58.3.174

Feldman, D. C., & Lankau, M. J. (2005). Executive coaching: A review and agenda for future research. *Journal of Management*, *31*, 829–848. doi:10.1177/0149206305279599

*Finn, F. A. (2007). *Leadership development through executive coaching: The effects on leaders' psychological states and transformational leadership behaviour* (Unpublished doctoral thesis). Queensland University of Technology, Brisbane, Australia.

Fischer, R. L., & Beimers, D. (2009). "Put me in, Coach" a pilot evaluation of executive coaching in the nonprofit sector. *Nonprofit Management and Leadership*, *19*, 507–522. doi:10.1002/nml.234

Ford, J. K., Kraiger, K., & Merritt, S. M. (2010). An updated review of the multidimensionality of training outcomes: New directions for training evaluation research. In S. W. J. Kozlowski & E. Salas (Eds.), *Learning, training, and development in organizations* (pp. 3–64). New York, NY: Routledge.

Franke, F., & Felfe, J. (2012). Transfer of leadership skills: The influence of motivation to transfer and organizational support in managerial training. *Journal of Personnel Psychology*, *11*, 138–147. doi:10.1027/1866-5888/a000066

Grant, A. M. (2003). The impact of life coaching on goal attainment, metacognition and mental health. *Social Behavior and Personality: An International Journal*, *31*, 253–264. doi:10.2224/sbp.2003.31.3.253

*Grant, A. M., Curtayne, L., & Burton, G. (2009). Executive coaching enhances goal attainment, resilience and workplace well-being: A randomised controlled study. *The Journal of Positive Psychology*, *4*, 396–407. doi:10.1080/17439760902992456

*Grant, A. M., Green, L. S., & Rynsaardt, J. (2010). Developmental coaching for high school teachers: Executive coaching goes to school. *Consulting Psychology Journal: Practice and Research*, *62*, 151–168. doi:10.1037/a0019212

Grant, A. M., Passmore, J., Cavanagh, M. J., & Parker, H. (2010). The state of play in coaching today: A comprehensive review of the field. *International Review of Industrial and Organizational Psychology*, *25*, 125–167. doi:10.1002/9780470661628

*Gyllensten, K., & Palmer, S. (2005). Can coaching reduce workplace stress? A quasi-experimental study. *International Journal of Evidence Based Coaching and Mentoring*, *3*, 75–85.

Haggard, D. L., Dougherty, T. W., Turban, D. B., & Wilbanks, J. E. (2011). Who is a mentor? A review of evolving definitions and implications for research. *Journal of Management*, *37*, 280–304. doi:10.1177/0149206310386227

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, *6*, 107–128. doi:10.2307/1164588

Higgins, M. C., & Kram, K. E. (2001). Reconceptualizing mentoring at work: A developmental network perspective. *Academy of Management Review*, *26*, 264–288. doi:10.5465/AMR.2001.4378023

Holladay, C. L., & Quiñones, M. A. (2003). Practice variability and transfer of training: The role of self-efficacy generality. *Journal of Applied Psychology*, *88*, 1094–1103. doi:10.1037/0021-9010.88.6.1094

Hunter, J., & Schmidt, F. (1990). *Methods of meta-analysis: Correcting for error and bias in research findings*. Newbury Park, CA: Sage.

Hunter, J., & Schmidt, F. (2004). *Methods of meta-analysis: Correcting for error and bias in research findings*. Thousand Oaks, CA: Sage.

Ilies, R., & Judge, T. A. (2005). Goal regulation across time: The effects of feedback and affect. *Journal of Applied Psychology*, *90*, 453–467. doi:10.1037/0021-9010.90.3.453

International Coach Federation (2012). *2012 ICF global coaching study: Executive summary*. Retrieved from http://www.coachfederation.org/coachingstudy2012

Johnson, D. A. (2013). A component analysis of the impact of evaluative and objective feedback on performance. *Journal of Organizational Behavior Management*, *33*, 89–103. doi:10.1080/01608061.2013.785879

Jones, R. A., Rafferty, A. E., & Griffin, M. A. (2006). The executive coaching trend: Towards more flexible executives. *Leadership & Organization Development Journal*, *27*, 584–596. doi:10.1108/01437730610692434

Kalinoski, Z. T., Steele-Johnson, D., Peyton, E. J., Leas, K. A., Steinke, J., & Bowling, N. (2013). A meta-analytic evaluation of diversity training outcomes. *Journal of Organizational Behavior*, *34*, 1076–1104. doi:10.1002/job.1839

Kilburg, R. R. (1996). Toward a conceptual understanding and definition of executive coaching. *Consulting Psychology Journal: Practice & Research*, *48*, 134–144. doi:10.1037/1061-4087.48.2.134

*Kines, P., Andersen, L. P. S., Spangenberg, S., Mikkelsen, K. L., Dyreborg, J., & Zohar, D. (2010). Improving construction safety through leader-based verbal safety communication. *Journal of Safety Research*, *41*, 399–406. doi:10.1016/j.jsr.2010.06.005

Kirkpatrick, D. L. (1967). Evaluation of training. In R. L. Craig & L. R. Bittel (Eds.), *Training and development handbook* (pp. 87–112). New York, NY: McGraw-Hill.

Kirkpatrick, D. L. (1996). Great ideas revisited. *Training and Development*, *50*, 54–59.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*, 254–284. doi:10.1037/0033-2909.119.2.254

*Kochanowski, S., Seifert, C. F., & Yukl, G. (2010). Using coaching to enhance the effects of behavioral feedback to managers. *Journal of Leadership & Organizational Studies*, *17*, 363–369. doi:10.1177/1548051809352663

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

Kombarakaran, F. A., Yang, J. A., Baker, M. N., & Fernandes, P. B. (2008). Executive coaching: It works!. *Consulting Psychology Journal: Practice and Research*, *60*, 78–90. doi:10.1037/1065-9293.60.1.78

Kozlowski, S. W. J., & DeShon, R. P. (2004). A psychological fidelity approach to simulation-based training: Theory, research and principles. In E. Salas, L. R. Elliott, S. G. Schflett, & M. D. Coovert (Eds.), *Scaled worlds: Development, validation, and applications* (pp. 75–99). Burlington, VT: Ashgate.

Kraiger, K., Ford, J. K., & Salas, E. D. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology*, *78*, 311–328. doi:10.1037/0021-9010.78.2.311

Kyrgidou, L. P., & Petridou, E. (2013). Developing women entrepreneurs' knowledge, skills and attitudes through e-mentoring support. *Journal of Small Business and Enterprise Development*, *20*, 548–566. doi:10.1108/JSBED-04-2013-0061

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35 year odyssey. *American Psychologist*, *57*, 705–717. doi:10.1037/0003-066X.57.9.705

Lowman, R. L. (2005). Executive coaching: The road to Dodoville needs paving with more than good assumptions. *Consulting Psychology Journal: Practice & Research*, *57*, 90–96. doi:10.1037/1065-9293.57.1.90

*Luthans, F., & Peterson, S. J. (2003). 360-degree feedback with systematic coaching: Empirical analysis suggests a winning combination. *Human Resource Management*, *42*, 243–256. doi:10.1002/hrm.10083

MacKie, D. (2007). Evaluating the effectiveness of executive coaching: Where are we now and where do we need to be? *Australian Psychologist*, *42*, 310–318. doi:10.1080/00050060701648217

McAdam, S. (2005). *Executive coaching*. London, UK: Thorogood.

McCauley, C. D., & Hezlett, S. A. (2002). Individual development in the workplace. In N. Anderson, D. Ones, H. K. Sinangil & C. Viswesvaran (Eds.), *Handbook of industrial, work and organizational psychology: Personnel psychology* (Vol. 1, pp. 313–335). Thousand Oaks, CA: Sage.

McGuffin, A. A., & Obonyo, E. (2010). Enhancing performance: A case study of the effects of employee coaching in construction practice. *Construction Management and Economics*, *28*, 141–149. doi:10.1080/01446190903460672

Mesmer-Magus, J., & Viswesvaran, C. (2010). The role of pre-training interventions in learning: A meta-analysis and integrative review. *Human Resource Management Review*, *20*, 261–282. doi:10.1016/j.hrmr.2010.05.001

*Moen, F., & Allgood, E. (2009). Coaching and the effect on self-efficacy. *Organization Development Journal*, *27*, 69–82.

*Moen, F., & Skaalvik, E. (2009). The effect from executive coaching on performance psychology. *International Journal of Evidence Based Coaching and Mentoring*, 7, 31–49.

Morisano, D., Hirsh, J. B., Peterson, J. B., Pihl, R. O., & Shore, B. M. (2010). Setting, elaborating and reflecting on personal goals improves academic performance. *Journal of Applied Psychology*, *92*, 255–264. doi:10.1037/a0018478

Murphy, K. R., & DeNisi, A. (2008). A model of the appraisal process. In A. Varma, P. S. Budhwar & A. DeNisi (Eds.), *Performance management systems: A global perspective* (pp. 81–94). Abingdon, Oxon, UK: Routledge.

Murrell, A. J., Blake-Beard, S., Porter, Jr, D. M., & Perkins-Williamson, A. (2008). Interorganizational formal mentoring: Breaking the concrete ceiling sometimes requires outside support. *Human Resource Management*, *47*, 275–294. doi:10.1002/hrm.20212

Neal, A., Godley, S. T., Kirkpatrick, T., Dewsnap, G., Joung, W., & Hesketh, B. (2006). An examination of learning processes during critical incident training: Implications for the development of adaptable trainees. *Journal of Applied Psychology*, *91*, 1276–1291. doi:10.1037/0021-9010.91.6.1276

*Olivero, G., Bane, K. D., & Kopelman, R. E. (1997). Executive coaching as a transfer of training tool: Effects on productivity in a public agency. *Public Personnel Management*, *26*, 461–469. doi:10.1177/009102609702600403

Orvis, K. A., Fisher, S. L., & Wasserman, M. E. (2009). Power to the people: Using learner control to improve trainee reactions and learning in web-based instructional environments. *Journal of Applied Psychology*, *94*, 960–971. doi:10.1037/a0014977

Parker, P., Hall, D. T., & Kram, K. E. (2008). Peer coaching: A relational process for accelerating career learning. *Academy of Management Learning & Education*, 7, 487–503. doi:10.5465/AMLE.2008.35882189

Parker, P., Kram, K. E., & Hall, D. T. (2013). Exploring risk factors in peer coaching: A multilevel approach. *The Journal of Applied Behavioral Science*, *49*, 361–387. doi:10.1177/0021886312468484

Patrick, J., Smy, V., Tombs, M., & Shelton, K. (2012). Being in one's chosen job determines pre-training attitudes and training outcomes. *Journal of Occupational and Organizational Psychology*, *85*, 245–257. doi:10.1111/j.2044-8325.2011.02027.x

Peterson, D. B., & Hicks, M. D. (1996). *Leader as coach: Strategies for coaching and developing others*. Minneapolis, MN: Personnel Decisions International.

Powell, K. S., & Yalcin, S. (2010). Managerial training effectiveness: A meta-analysis 1952–2002. *Personnel Review*, *39*, 227–241. doi:10.1108/00483481011017435

Riketta, M. (2005). Organization identification: A meta-analysis. *Journal of Vocational Behavior*, *66*, 358–384. doi:10.1016/j.jvb.2004.05.005

Riketta, M. (2008). The causal relation between job attitudes and performance: A meta-analysis of panel studies. *Journal of Applied Psychology*, *93*, 472–481. doi:10.1037/0021-9010.88.6.1057

Salas, E., & Kozlowski, S. W. J. (2010). Learning, training, and development in organizations: Much progress and a peek over the horizon. In S. W. J. Kozlowski & E. Salas (Eds.), *Learning, training, and development in organizations* (pp. 3–64). New York, NY: Routledge.

Seifert, C. F., Yukl, G., & McDonald, R. A. (2003). Effects of multisource feedback and a feedback facilitator on the influence behavior of managers toward subordinates. *Journal of Applied Psychology*, *88*, 561–569. doi:10.1037/0021-9010.88.3.561

Sherman, S., & Freas, A. (2004). The wild west of executive coaching. *Harvard Business Review*, *82*(11), 82–93.

Smither, J. W. (2011). Can psychotherapy research serve as a guide for research about executive coaching? An agenda for the next decade. *Journal of Business Psychology*, *26*, 135–145. doi:10.1007/s10869-011-9216-7

*Smither, J. W., London, M., Flautt, R., Vargas, Y., & Kucine, I. (2003). Can working with an executive coach improve multisource feedback ratings over time? A quasi-experimental field study. *Personnel Psychology*, *56*, 23–42. doi:10.1111/j.1744-6570.2003.tb00142.x

Smither, J. W., London, M., & Reilly, R. R. (2005). Does performance improve following multisource feedback? A theoretical model, meta-analysis, and a review of empirical findings. *Personnel Psychology*, *54*, 33–66. doi:10.1111/j.1744-6570.2005.514_1.x

Sonnentag, S., & Frese, M. (2002). Performance concepts and performance theory. In S. Soonentag (Ed.), *Psychological management of individual performance* (pp. 4–19). West Sussex, UK: John Wiley.

Steel, P. D., & Kammeyer-Mueller, J. D. (2002). Comparing meta-analytic moderator estimation techniques under realistic conditions. *Journal of Applied Psychology*, *87*, 96–111. doi:10.1037/0021-9010.87.1.96

Stone-Romero, E. F. (2008). Strategies for improving the validity and utility of research in human resource management and allied disciplines. *Human Resource Management Review*, *18*, 205–209. doi:10.1016/j.hrmr.2008.07.012

Strumpf, C. (2002). Coaching from the inside: When, why, and how. In C. Fitzgerald & J. Berger (Eds.), *Executive coaching: Practices & perspectives* (pp. 27–57). Palo Alto, CA: Davies-Black.

Sue-Chan, C., & Latham, G. P. (2004). The relative effectiveness of external, peer, and self-coaches. *Applied Psychology: An International Review*, *53*, 260–278. doi:10.1111/j.1464-0597.2004.00171.x

Swart, J., & Harcup, J. (2013). 'If I learn do we learn?': The link between executive coaching and organizational learning. *Management Learning*, *44*, 337–354. doi:10.1177/1350507612447916

*Taie, E. S. (2011). Coaching as an approach to enhance performance. *Journal for Quality and Participation*, *34*, 34–38.

Taylor, P. J., Russ-Eft, D. F., & Chan, D. W. L. (2005). A meta-analytic review of behavior modeling training. *Journal of Applied Psychology*, *90*, 692–709. doi:10.1037/0021-9010.90.4.692

Terry, P. E., Seaverson, E. L. D., Staufacker, M. J., & Gingerich, S. B. (2010). A comparison of the effectiveness of a telephone coaching program and a mail-based program. *Health Education & Behavior*, *37*, 895–912. doi:10.1177/1090198110367876

Tharenou, P., Saks, A. M., & Moore, C. (2007). A review and critique of research on training and organizational-level outcomes. *Human Resource Management Review*, *17*, 251–273. doi:10.1016/j.hrmr.2007.07.004

Theeboom, T., Beersma, B., & van Vianen, A. E. M. (2014). Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context. *The Journal of Positive Psychology*, *9*, 1–18. doi:10.1080/17439760.2013.837499

*Toegel, G., & Nicholson, N. (2005). *Multisource feedback, coaching, and leadership development: gender homophily in coaching dyads*. Academy of Management Best Conference Paper, MC, F1.

Tonidandel, S., Avery, D. R., & Phillips, M. G. (2007). Maximizing returns on mentoring: Factors affecting subsequent protégé performance. *Journal of Organizational Behavior*, *28*, 89–110. doi:10.1002/job.418

van der Locht, M., van Dam, K., & Chiaburu, D. S. (2013). Getting the most of management training: The role of identical elements for training transfer. *Personnel Review*, *42*, 422–439. doi:10.1108/PR-05-2011-0072

Wegge, J., Bipp, T., & Kleinbeck, U. (2007). Goal setting via video conferencing. *European Journal of Work and Organizational Psychology*, *16*, 169–194. doi:10.1080/13594320601125567

Wentz, E., Nyden, A., & Krevers, B. (2012). Development of an internet-based support and coaching model for adolescents and young adults with ADHD and autism spectrum disorders: A pilot study. *European Child Adolescent Psychiatry*, *21*, 611–622. doi:10.1007/s00787-012-0297-2

Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology*, *75*, 315–321. doi:10.1037/0021-9010.75.3.315

Whitmore, J. (1992). *Coaching for performance*. London, UK: Nicholas Brealey.

Witherspoon, R., & White, R. P. (1996). Executive coaching: A continuum of roles. *Consulting psychology Journal: Practice and Research*, *48*, 124–133. doi:10.1037/1061-4087.48.2.124

Yan, T., Wilber, K. H., & Simmons, W. J. (2011). Motivating high-risk older adults to exercise: Does coaching matter? *Home Health Care Services Quarterly*, *30*, 84–95. doi:10.1080/01621424.2011.569670