The Journal of Positive Psychology: Dedicated to furthering research and promoting good practice
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rpos20

Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context

Tim Theeboom[a], Bianca Beersma[a] & Annelies E.M. van Vianen[a]
[a] Department of Work and Organizational Psychology, University of Amsterdam, Weesperplein 4, 1018 XA, Amsterdam, The Netherlands
Published online: 13 Sep 2013.

The Journal of Positive Psychology, 2013
http://dx.doi.org/10.1080/17439760.2013.837499

Routledge
Taylor & Francis Group

Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context

Tim Theeboom*, Bianca Beersma and Annelies E.M. van Vianen

Department of Work and Organizational Psychology, University of Amsterdam, Weesperplein 4, 1018 XA, Amsterdam, The Netherlands

(*Received 12 February 2013; accepted 15 August 2013*)

Whereas coaching is very popular as a management tool, research on coaching effectiveness is lagging behind. Moreover, the studies on coaching that are currently available have focused on a large variety of processes and outcome measures and generally lack a firm theoretical foundation. With the meta-analysis presented in this article, we aim to shed light on the effectiveness of coaching within an organizational context. We address the question whether coaching has an effect on five both theoretically and practically relevant individual-level outcome categories: performance/skills, well-being, coping, work attitudes, and goal-directed self-regulation. The results show that coaching has significant positive effects on all outcomes with effect sizes ranging from $g = 0.43$ (coping) to $g = 0.74$ (goal-directed self-regulation). These findings indicate that coaching is, overall, an effective intervention in organizations.

Keywords: coaching; coaching effectiveness; coaching interventions; coaching outcomes; meta-analysis

Introduction

The use of coaching methodologies as a means of enhancing performance and development in organizations has increased substantially over the last two decades. Since its foundation in 1995, the International Coach Federation (ICF) has seen its member count grow to over 20,000 members in over 100 countries in 2012 (ICF, 2012) and the total annual revenue from coaching is estimated at roughly $2 billion globally (ICF, 2012). Coaching can be defined as a 'result-oriented, systematic process in which the coach facilitates the enhancement of life experience and goal-attainment in the personal and/or professional lives of normal, non-clinical clients' (Grant, 2003, p. 254).

While coaching is often considered as a useful tool for individual and organizational development (Grant, Passmore, Cavanagh, & Parker, 2010), the lack of a systematic empirical review of research on the outcomes of coaching makes it prone to skepticism regarding its effectiveness (Bono, Purvanova, Towler, & Peterson, 2009; Bozer & Sarros, 2012). This skepticism seems valid in the light of the high costs of coaching. In a survey completed by 428 coaches, Bono et al. (2009) found that the average hourly fee for coaches was $237. Harvard Business Review even stated that 'coaches aren't monks bound to a vow of poverty, and they can earn up to $3500 an hour' (Couto & Kauffman, 2009, p. 1). However, the current literature on coaching is inconclusive on whether these high financial costs outweigh the benefits that coaching potentially has for organizations (Leonard-Cross, 2010). Thus, whereas coaching is very popular as a management tool and organizations are apparently willing to pay large amounts of money for it, an empirical review of coaching effectiveness is lagging (Bozer & Sarros, 2012).

In addition, most studies on coaching are conducted by practitioners. While these studies can provide valuable insights, most practitioners are not trained in research methods. As a result, validated outcome measures are seldom used and the studies generally lack a firm theoretical foundation (Grant, 2013). Consequently, extant research on coaching has focused on a large variety of processes and outcome measures (Latham, 2007) and this disjointed state of the literature on coaching hinders the establishment of a theoretical framework for future research (Spence & Oaedes, 2011; Sue-Chan & Latham, 2004).

All in all, there is a strong need for a quantitative summary and integration of existing coaching research. To date, several good literature reviews, rather than empirical reviews, have been published (Brock, 2008; Feldman & Lankau, 2005; Grant, 2001; Grant & Cavanagh, 2004; Grant, Passmore, et al., 2010; Kampa-Kokesch & Anderson, 2001; Passmore & Fillery-Travis, 2011). To our knowledge only De Meuse, Dai, and Lee (2009) conducted a quantitative review. Their review assessed the success of coaching in terms of its return on investment (ROI). Although ROI can be an indicator of

*Corresponding author. Email: t.theeboom@uva.nl

the effectiveness of coaching as a change methodology, it also has some serious limitations which we will discuss below.

With the meta-analysis we present in this article, we aim to provide a comprehensive quantitative review and to answer the question: Does coaching[1] work when provided in an organizational context by professionally trained coaches? Furthermore, by means of a systematic review and integration of the types of coaching outcomes that were included in prior studies, we aim to give an initial impetus to further theoretical development of coaching research. That is, by organizing coaching outcomes that emerge from the literature into meaningful categories, we aim to enable future studies to build on or extend theory that explains the paths and processes that lead to these different categories of outcomes.

Coaching

Coaching has its roots in a multitude of disciplines, including philosophy, sociology, anthropology, sports, communication science and even natural sciences (Brock, 2008). However, in terms of the number of articles published in peer-reviewed journals, sub-disciplines of psychology have shown to be the most fruitful areas of research on coaching (Grant, Passmore, et al., 2010). Initially, most of the research on coaching was conducted within areas such as sports psychology (e.g. Gallwey, 1974; Whitmore, 1992) and clinical psychology (e.g. Berg & Szabo, 2005; De Shazer, 1988). More recently, research on coaching has particularly flourished in two strongly related areas of psychology that emerged within the last two decades: positive psychology (Seligman & Csikszentmihalyi, 2000) and coaching psychology (Passmore, 2010).

Positive psychology focuses on studying three aspects that constitute the scientific notion of happiness: positive emotion, meaning, and engagement (Seligman, 2007). Coaching psychology focuses on studying behavior, cognition, and emotion within coaching practice to deepen understanding of coaching processes and to enhance coaching techniques (Passmore, 2010). While there are some debates about how these areas of research are related to each other (some authors see coaching psychology as a sub-discipline or an applied form of positive psychology; e.g. Grant & Cavanagh, 2007), both areas share their focus on performance enhancement, positive aspects of human nature, and the strengths of individuals (Linley & Harrington, 2005). Therefore, positive psychology seems to offer a robust framework for researching coaching and as such may constitute 'one of the solutions to the lack of a theoretical framework in the coaching field' (Freire, 2013, p. 428)

While different psychological sub-disciplines initially developed their own specific conceptual framework, more recent literatures have gradually moved towards a generally accepted definition of coaching. Kilburg (1996) originally defined coaching of executives in organizations as 'a helping relationship between a managerial-client and a consultant that follows a formally defined coaching agreement'. Grant (2003) transformed this definition into a more general one and defined coaching as a 'result-oriented, systematic process in which the coach facilitates the enhancement of life experience and goal-attainment in the personal and/or professional life of normal, non-clinical clients' (p. 254).

This latter definition encompasses several important features, namely: it can be applied to a multitude of coaching domains (e.g. personal coaching and organization coaching) and coaches (executive and non-executive), it emphasizes the self-directedness of the coaching process, and it recognizes coaching as a systematic process rather than just being empathic and 'having good conversations' (Leonard-Cross, 2010). Furthermore, although the differences and similarities between coaching and therapy are still a topic of debate (Bono et al., 2009; Brunning, 2006; Hart, Blattner, & Leipsic, 2001), Grant's (2003) definition of coaching distinguishes coaching from therapy by its focus on a non-clinical population.

Coaching effectiveness: beyond return of investment measures

The literature on coaching has grown exponentially in the last 15 years. Whereas only 93 articles were published in the years between 1937 and 1999, the total number of articles and dissertations on coaching reached 634 in 2011 (Grant, 2013) and the number has been steadily growing ever since. However, the bulk of articles still consists of descriptive papers and/or case studies as well as practitioner articles primarily aimed at emphasizing the benefits of certain coaching interventions (De Meuse et al., 2009).

Additionally, this predominantly practitioner-generated research has resulted in an overemphasis on ROI measures. ROI as a measure of coaching effectiveness is appealing because it provides some direct insight into the tangible benefits of coaching interventions (Fillery-Travis & Lane, 2006; Leonard-Cross, 2010). However, the ROI measure has some serious limitations. For instance, the factors included in the most frequently used calculation of the ROI metric (benefits – costs/costs × 100) are highly idiosyncratic and tend to ignore context variables such as team input. Therefore, it is often impossible to determine the degree to which the financial benefits can be directly attributed to the coaching intervention (Grant, Passmore, et al., 2010, p. 26). Additionally, performance measures (direct benefits) are seldom available and a narrow focus on ROI and other

performance-related measures neglects other – more indirect – ways in which organizations could potentially benefit from coaching, such as employee well-being and health. Therefore, the current meta-analysis investigates coaching effectiveness by looking at well-validated, more distal indicators of functioning in addition to performance measures: well-being, coping, work and career-related attitudes, and goal-directed self-regulation.

The overall goal of coaching in a work context is to optimize a person's work-related functioning. First, individuals in organizations function better if they feel well, that is, if their basic needs are fulfilled (Deci & Ryan, 1985, 2000) and if they do not struggle with health-related problems as caused by their job (Burton et al., 1999). With regard to the latter, work-related stress affects over 20% of workers in the European Union (Brunn & Milczarek, 2007) which costs organizations about 20,000 million Euros per year. Indeed, research has evidenced negative relationships between 'soft' outcomes, including individual health and well-being, and 'hard' performance measures (Bond & Bunce, 2003; Wright & Cropanzano, 2000). Specifically, Duijts, Kant, van den Brandt, and Swaen (2008) found that a coaching intervention in medical health care and educational organizations resulted in a positive change in employees' well-being as indicated by a decrease in sickness-related absenteeism and burnout and an increase in life satisfaction. Furthermore, coaching could also have a preventive role via its effect on individual' abilities to cope with stressors. Several studies have found coaching to have positive effects on coping mechanisms such as resilience (Grant, Curtayne, & Burton, 2009), mindfulness (Spence, Cavanagh, & Grant, 2008), and the use of self-enhancing attributions (Moen & Skaalvik, 2009).

Another way in which coaching could benefit organizational effectiveness is via its potentially beneficial effects on employees' work- and career-related perceptions and attitudes (e.g. job satisfaction). Coaching may facilitate the cognitive 'reframing' of work experiences and attitudes which is a central aspect of multiple coaching programs (Grant, 2001). In order to function better, employees may initially try to change their work tasks and interactions, but if organizational boundaries put constraints on making actual changes, a coach could help them adjust their job perceptions (Wrzesniewski & Dutton, 2001). Indeed, a preliminary study showed that coaching positively influenced work- and career-related attitudes (Bozer & Sarros, 2012) which, in turn, may positively affect work performance (Judge, Thoresen, Bono, & Patton, 2001; Organ & Ryan, 1995).

Finally, coaching enhances organizational effectiveness through its potentially beneficial effect on employees' goal-directed self-regulation (Grant, 2003). Most coaching programs target either one or more stages of goal-directed self-regulatory processes. For example, several studies have shown that coaching can positively influence goal-attainment expectancy (Evers, Brouwers, & Tomic, 2006; Moen & Skaalvik, 2009) and goal-progression and commitment (Green, Oades, & Grant, 2006). Because the relationship between goal-setting on the one hand and motivation and performance on the other hand is well established (Locke & Latham, 2002), improving employees' self-regulation by enhancing their ability to set and strive for goals is yet another way in which organizations could benefit from coaching.

All in all, coaching could benefit organizations by enhancing employees' performance and skills, well-being, coping, work attitudes, and goal-directed self-regulation. Because extant empirical studies (see Method section below) lack a clear conceptual framework for classifying coaching outcomes, we categorized available outcomes into these outcome dimensions, which are both theoretically and practically relevant and also well established within the psychological literature.

Method

Data collection

In order to decide which studies to include in our analyses, we undertook an extensive literature search, which consisted of five phases based on the PRISMA statement for reporting meta-analysis (Moher, Liberati, Tetzlaff & Altman, 2009). A PRISMA flowchart is displayed in Figure 1.

First, we searched several (social) science databases (Google Scholar, JSTOR, Mendely, PsycINFO, ScienceDirect, and Springerlink) using the following keywords: workplace coaching, developmental coaching, executive

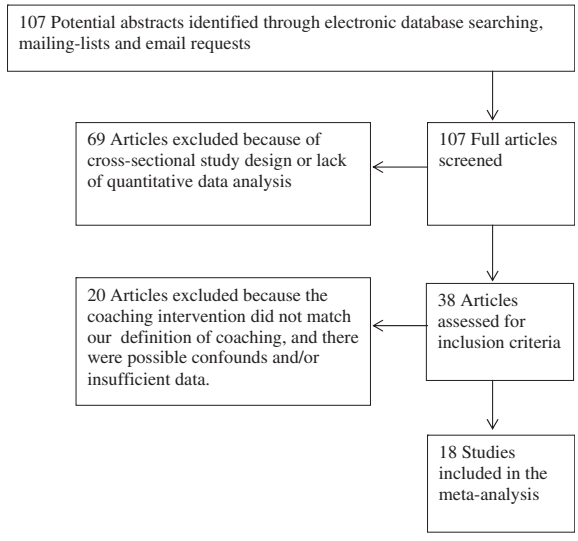

Figure 1. Search strategy used for the inclusion of studies considered in the present meta-analysis.

coaching, effects of coaching, and outcomes of coaching. Initially, we searched for articles (i) that included quantitative data on the effects of coaching; (ii) in which coaching was provided by professionally trained external coaches or trained peers; and (iii) in which the coaches belonged to a non-clinical population. Second, we performed both backward and forward searches on the studies we retrieved in phase 1. Third, we contacted several scholars known to be active in the field of coaching psychology research in order to retrieve (yet) unpublished results. At the same time, we sent out a request for published and unpublished studies via the mailing-list services of Academy of Management, the Society for Industrial and Organizational Psychology, and the European Association of Social Psychology.

In the fourth phase, we screened the 107 full articles retrieved in the first three phases and excluded all cross-sectional studies since these do not allow controlling for most threats to internal validity (Cook & Campbell, 1979). Furthermore, we excluded all studies in which the authors did not perform quantitative analyses (e.g. case studies). In the fifth and final phases, we used a priori conceptual and methodological criteria to decide which of these studies would be included in the final analysis. For example, we excluded all studies in which the described coaching process did not match our definition of coaching or for which a description of the coaching process could not be obtained.

In addition to Grant's (2003) definition of coaching, we took one more characteristic of coaching into account when selecting studies, namely that the coaching was provided by a (professionally trained) coach with no formal authority over the coachee. The main reason for this is that research on mentoring has shown that a mentor's formal authority over the protégée can affect the way in which the protégée behaves during coaching sessions (Mullen, 1994; Tepper, 1995; Waldron, 1991). Furthermore, the goals in managerial coaching (in which a manager coaches an employee) are often strongly related to organizational performance (Cox, Bachkirova, & Clutterbuck, 2009) which might influence the degree to which the coaching process is self-directed. Another reason for excluding studies on managerial coaching is that our aim for this research was to set a first step for answering the question whether the high costs of hiring *external* professionally trained coaches are justifiable for both individual clients and organizations. Thus, we excluded all studies in which the coach had a formal authority over the coachee.

We also excluded all studies in which the influence of other interventions (e.g. leadership development programs) could not be ruled out as a confounding factor. Because we aim to provide more insight into the usefulness of coaching in organizations, we only included studies conducted in a work or educational context. Studies conducted in an educational context (e.g. undergraduate students) were included because the clients' characteristics and needs are similar to those of clients in organizational settings, that is, they are similar with respect to their demographics and the challenges they face (working in teams, meeting deadlines etc.). Finally, we excluded all studies in which not enough statistical information was available or could not be made available after contacting the corresponding authors.

This selection process resulted in a total of 18 studies included in the final analysis. All studies that were included in the final analysis are indicated with an * in our list of references. An overview of basic characteristics of these studies is displayed in Table 1.

Outcome categorization

The categorization of outcomes was conducted in three steps. In the first step, the first and the second authors mutually assigned all study outcomes into one of the following categories: performance/skills, well-being, coping, work attitudes, or goal-directed self-regulation while the third author categorized the outcomes independently of the first two authors. In the second step, the interrater agreement was calculated. The interrater agreement (Cohens' κ) was 0.80, which is considered to be large (see Landis & Koch, 1977). In the third step, the authors discussed their discrepancies and agreed on a final categorization. Several studies in the meta-analysis included multiple measures within the same outcome category (e.g. measures of stress and burnout both fall within the well-being category). An average of these effect sizes was included in order to prevent violation of the independent sample assumption (Hunter & Schmidt, 2004).

The performance/skills category includes both subjective and objective outcome measures that either directly reflect performance (e.g. number of sales, supervisory rated job performance) or reflect the demonstration of behaviors needed for an organization to be effective (e.g. transformational leadership behaviors). The well-being category includes both subjective and objective outcome measures that are a direct representation of peoples' well-being, health, need fulfillment, and affective responses. Examples of these are measures of psychopathology (e.g. Depression Anxiety and Stress Scale; Lovibond & Lovibond, 1995) and burnout (e.g. Maslach Burnout Inventory; Maslach & Jackson, 1986).

The coping category includes outcome measures related to the ability to deal with present and future job demands and stressors. Examples of these are measures of self-efficacy (e.g. General Self-Efficacy Scale; Schwarzer & Jerusalem, 1995) and mindfulness (e.g. Mindful Attention Awareness Scale; Brown & Ryan,

Table 1. Study characteristics and outcome overview of studies included in the meta-analysis.

Study	*n*	Intervention	Nr. sessions	Outcomes	Design
Bozer and Sarros (2012)	96	Cognitive behavioral coaching	11	Self-reports • Career satisfaction • Job commitment • Job Performance • Self-Awareness Supervisory ratings • Job performance • Self-awareness • Task performance	RCT
Cerni, Curtis, and Colmar (2010)	14	Epstein constructive thinking intervention	10	Staff-ratings • Transformational leadership	QEF
Egan and Song (2005)	103	Coaching on goal-setting and goal achievement	Unknown	Self-reports • Job satisfaction • Organizational commitment • Performance goal orientation • Performance rating Supervisory ratings • Performance rating	QEF
Finn (2007)	32	CB-SFC	6	Self-reports • Developmental planning • Developmental support • Openness to new behaviors • Positive affect • Self-efficacy	QEF
Grant (2003)	20	CB-SFC group coaching	10	Self-reports • Anxiety • Depression • GAS • Insight • Self-reflection • Stress • Quality of life	WSD[6]
Grant (2008)	29	CB-SFC peer coaching	5	Self-reports • Anxiety • Cognitive hardiness • Depression • GAS • Learning • Personal insight • Well-being	WSD

(*Continued*)

Table 1. (*Continued*).

Study	*n*	Intervention	Nr. sessions	Outcomes	Design
Grant et al. (2009)	41	CB-SFC	10	Self-reports • Anxiety • Depression • Gas • Resilience • Stress • Well-being • Work	RCT
Grant, Green, and Rynsaardt (2010)	65	CB-SFC	10	Self-reports • Agency • Anxiety • Cognitive hardiness • Depression • Hope • Interpersonal problems • General health questionnaire • GAS personal • GAS work • Resilience • Stress • Workplace well-being	RCT
Green et al. (2007)	44	CB-SFC	10	Self-reports • Anxiety • Cognitive hardiness • Hope • Depression • Stress	RCT
Green et al. (2006)	50	CB-SFC in groups	9	Self-reports • Agency • Autonomy • Environmental mastery • Goal striving • Negative affect • Personal growth • Positive affect • Positive relations with others • Purpose in life • Satisfaction with life • Self-acceptance • Total hope	RCT
Kochanowski, Seifert, and Yukl (2010)	84	Coaches provided Feedback on leadership style	10	Self-reports • Leadership behavior	QEF
Luthans and Peterson (2004)	67	Coaches discussed discrepancy between executives' own performance ratings and ratings of employees	1	Self-reports • Anxiety • Depression • GAS[3] • Goal difficulty	WSD

(*Continued*)

Table 1. (*Continued*).

Study	*n*	Intervention	Nr. sessions	Outcomes	Design
				• Organizational commitment • Satisfaction with coworkers • Satisfaction with supervision • Satisfaction with work • Self-insight scale • Self-reflection scale • Stress • Turnover intentions • Quality of Life Inventory	
Moen and Skaalvik (2009)	19	Coaching aimed at self-actualization and promotion of resources	4	Self-reports • Attribution of failure to ability • Attribution failure to circumstances • Attribution failure to effort • Attribution failure to strategy • Attribution success to ability • Attribution success to circumstances • Attribution success to effort • Attribution success to strategy • Autonomy • Competence • Goal clarity • Goal commitment • Goal difficulty • Goal feedback • Goal Strategy • Need satisfaction at work • Relatedness • Self-efficacy	QEF
Peterson (1993)	100	Individual coaching for effectiveness program	50	Self-reports • Effectiveness Supervisory ratings • Effectiveness Ratings by coach • Effectiveness	WSD
Poepsel (2011)	28	Online CB-SFC	4	Self-reports • Goal attainment • Hope • Subjective well-being	RCT post-test only
Smither, London, Flautt, Vargas, and Kucine (2003)	1243	Coaching based on 360-degree feedback	3	Self-reports • Elaboration of self-reports • Elaboration of peer ratings	QEF

(*Continued*)

Table 1. (*Continued*).

Study	*n*	Intervention	Nr. sessions	Outcomes	Design
				• Elaboration of supervisory ratings • Performance Rating by peers • Performance Supervisory ratings • Performance Rating by authors • Goal specificity	
Spence et al. (2008)	15[a]	Coaching on goal-setting and goal achievement	4	Self-reports • Anxiety • Depression • Environmental mastery • Mindfulness • Reflection • Rumination • Satisfaction with life • Self-Acceptance • Stress	RCT
Spence and Grant (2007)	40	CB-SFC	10	Self-reports • Autonomy • Environmental mastery • Goal commitment • Goal progression • Negative affect • Personal Growth • Positive affect • Positive relations with others • Purpose in life • Self-acceptance • Subjective well being	RCT

Notes: RCT = randomized control trial; QEF = quasi-experimental field study in which participants were non randomly allocated to experimental and control groups; WSD = within-subjects design without control group, which includes both pre- and post-intervention measures; CB-SFC = cognitive behavioral solution-focused coaching; and GAS = goal-attainment scaling.
[a]Only the C-MT condition (see original) was included for the calculation of effect sizes.

2003). The work attitudes category includes outcome measures related to cognitive, affective, and behavioral responses toward work and career, such as job satisfaction (e.g. Job Description Index; Smith, Kendall, & Hulin, 1969), organizational commitment (e.g. Organizational Commitment Scale; Porter, Steers, Mowday & Boulian, 1974), and career satisfaction (e.g. Career Satisfaction Scale; Greenhaus, Parasuraman & Wormley, 1990). Finally, the goal-directed self-regulation category includes all outcome measures related to goal-setting, goal-attainment, and goal-evaluation. This category also includes goal-attainment scaling (GAS) measures, which are increasingly popular in coaching settings (see Spence [2008] and Peterson [1993] for overviews of the use of GAS in coaching research).

Calculating the effect sizes

One of the most challenging steps in a meta-analysis is combining the effect sizes of different studies in one

analysis (McGaw & Glass, 1980). Since effect sizes based on means are easily interpretable and the studies in our analysis employed a large variety of outcome measures to assess the impact of coaching interventions, the use of an effect size index based on standardized means was the logical choice (Borenstein, Hedges, Higgins, & Rothstein 2009; Cohen, 1988). Since the most popular index, Cohen's *d,* tends to overestimate the population effect size when small samples are included in the analysis, we chose to use Hedges *g* which can still be interpreted as the mean difference expressed in standard deviation units and applies a simple correction to overcome this bias (Hedges, 1981).

Effect sizes can be defined in relation to pre-intervention scores, post-intervention scores, or difference scores. In theory, it is possible to choose either definition because effect sizes can be transformed into a common effect size index by using the correlation between pre- and post-intervention scores to estimate the sampling bias (Morris & DeShon, 1997). Unfortunately, the correlations between pre- and post-intervention scores were seldom provided and the amount of studies that used a mixed design outnumbered the amount of studies that used a within-subject design. Therefore, effect sizes based on the pooled (experimental and control groups) standard deviations of the post-intervention scores were chosen as the referent effect size index. By doing so, the estimation of parameters was minimized since only the pre-intervention/post-intervention correlations for the (minority of) studies that used a within-subject design needed to be estimated[2]. Finally, effect sizes based on post-intervention standard deviations are likely to be slightly biased downward, and are thus the most conservative choice (Carlson & Schmidt, 1999).

Meta-analytic procedure and statistical analyses

The Hedges and Olkin (1985) approach to meta-analysis was used to calculate the effect sizes. Comparisons with other commonly applied methods such as by Hunter and Schmidt (1990) and Rosenthal (1991) suggest that differences between the Hedges and Olkin methods and the other methods are relatively small and only apply under very specific circumstances. If anything, the Hedges and Olkin method can be considered to be the most conservative approach because it does not allow for the statistical corrections for artifactual sources of variance (e.g. measurement error and restriction of range) that tend to result in an inflation of effect size estimates (Borenstein et al., 2009). Additionally, the Hedges and Olkin approach seems to provide the most conservative estimate of the (lower limit of) confidence intervals (Johnson, Mullen, & Salas, 1995), which can be used for determining the statistical significance of effect sizes.

After selecting the general approach for the meta-analysis, the statistical model for the meta-analysis has to be designated. In terms of the model for the meta-analysis, the (conservative) random-effect model was adopted as recommended by the National Research Council (1992). As opposed to the fixed effect model, the random-effect model allows that the true effect size varies from study to study based on both the variability of the independent variable (e.g. intensity or duration of intervention) and differences in the samples of the research population such as age, educational background, and type of job of the coaches (Borenstein et al., 2009; Hedges & Cooper, 1994).

Heterogeneity between studies was quantified by an assessment of both the classical Cochran Q statistic (1954) and the I^2 statistic as proposed by Higgins and Thompson (2002), see also Higgins, Thompson, Deeks, & Altman, 2003). While the Q statistic serves as a test of significance for between-study heterogeneity, the value for I^2 represents the proportion of between-study variance in effect sizes that can be attributed to between-study heterogeneity rather than within-study variability (Borenstein et al., 2009). When the value for I^2 is large (see Higgins & Thompson, 2002 for some guidelines on interpretation), one of the possible explanations for this is the existence of moderating variables. In what should be considered exploratory analyses due to the relatively small amount of studies, we explored the influence of two of these potential (methodological) moderating variables in order to provide guidance for the future methodological approach of coaching research. First, following the example of Fusar-Poli et al. (2012), we conducted subgroup analyses for sets of studies that were different in terms of study design (mixed designs vs. within-subject designs). Second, we examined the influence of the number of coaching sessions using meta-regression analysis. Finally, following recent recommendations by Sterne et al. (2011), we assessed the risk for publication bias (or small study bias) by visually inspecting funnel plots and by applying the regression intercept of Egger, Smith, Schneider, and Minder (1997).

Software for the analysis and statistical corrections

The software that was used for the analysis was comprehensive meta-analysis. CMA was developed by Borenstein, Hedges, Higgins, and Rothstein (2005) and is based on the Hedges and Olkin's (1985) approach to meta-analysis. CMA offers advantages over other software in terms of its flexibility to handle multiple data entry formats (e.g. data from within-subject designs and mixed designs) and its intuitive approach to sensitivity analysis and the detection of between-study heterogeneity (Borenstein et al., 2009).

Results

Aggregated effect sizes and overall between-study heterogeneity

Table 2 contains the weighted effect sizes (aggregated over outcomes) per study.

The point estimate of the overall weighted effect size (aggregated over all studies and outcomes) was significant (g = 0.66, 95% CI, 0.39–0.93, p = 0.000), suggesting that coaching, in general, has a significant positive effect across the range of outcome measures we examined. The relatively large point estimate of g in the study by Peterson (1993) encouraged us to perform a sensitivity analysis in which the analysis was repeated while excluding the results of this study. Although the overall weighted point estimate of g dropped, it remained significant (g = 0.51, 95% CI, 0.34–0.69, p < 0.000). Thus, excluding the study by Peterson (1993) did not alter our conclusions regarding the point estimate of the overall weighted effect size. According to the tentative criteria set by Higgins and Thompson (2002), the heterogeneity in effect sizes was statistically significant and large in magnitude (Q = 130.05; p < 0.000; I^2 = 86.93). As mentioned, this substantial variance in effect sizes encourages the consideration of moderating variables. Therefore, we further explored the heterogeneity in effect sizes.

Effect sizes per outcome category

The main goal of this meta-analysis was to provide insight into the effects of coaching on various individual-level psychological outcomes. Table 3 contains the results for all outcome categories: performance and skills, well-being, coping, work attitudes, and goal-directed self-regulation.

The results indicate that coaching interventions have significant positive effects on all outcome categories: performance and skills (g = 0.60, 95% CI, 0.04–0.60, p = 0.036), well-being (g = 0.46, 95% CI, 0.28–0.62, p < 0.001), coping (g = 0.43, 95% CI, 0.25–0.61, p < 0.001), work attitudes (g = 0.54, 95% CI, 0.34–0.73, p < 0.001), and goal-directed self-regulation (g = 0.74, 95% CI, 0.42 – 1.06, p < 0.001). Based on our earlier sensitivity analysis (see above), we repeated the analysis for the performance and skills outcome while excluding the results of the study by Peterson (1993). Although the weighted point estimate of the effect size *for performance and skills* dropped considerably, it remained significant (g = 0.19, 95% CI, 0.04–0.32, p = 0.013). We will discuss the implications of this finding in the discussion section. Finally, we note that the significance of the Q statistics and the moderate-to-high values of I^2 for both the performance and skills outcomes and the goal-directed self-regulation outcomes indicate that the influence of between-study heterogeneity is especially apparent for these outcome categories.

The influence of study design

We explored the differences in effect size patterns between studies that used a mixed design vs. studies that used a within-subject design. The results of this analysis are shown in Table 4.

Table 2. Weighted effect sizes per study aggregated over outcomes.

Study	n	Sessions	Study design	g	CI (95%) Lower	CI (95%) Upper	p-value
Bozer and Sarros (2012)	96	11	RCT	0.36	−0.12	0.83	0.145
Cerni et al. (2010)	14	10	QEF	0.08	−0.15	0.30	0.501
Egan and Song (2005)	103	Unknown	QEF	0.69	0.29	1.08	0.001
Finn (2007)	32	6	QEF	1.17	0.34	1.99	0.006
Grant (2003)	20	10	WSD	0.84	0.31	1.38	0.002
Grant (2008)	29	5	WSD	0.71	0.23	1.19	0.004
Grant et al. (2009)	41	10	RCT	0.63	0.00	1.25	0.049
Grant et al. (2010)	65	10	RCT	0.25	−0.34	0.83	0.406
Green, Grant, and Rynsaardt (2007)	44	10	RCT	0.61	0.02	1.20	0.042
Green et al. (2006)	50	9	RCT	0.70	0.13	1.28	0.016
Kochanowski et al. (2010)	84	10	QEF	0.25	−0.16	0.66	0.239
Luthans and Peterson (2004)	67	1	WSD	0.74	0.47	1.01	0.000
Moen and Skaalvik (2009)	19	4	QEF	0.82	−0.10	1.75	0.080
Peterson (1993)	100	50	WSD	2.33	1.93	2.72	0.000
Poepsel (2011)	28	4	RCT	0.92	0.16	1.70	0.018
Smither et al. (2003)	1243	3	QEF	0.13	0.00	0.27	0.049
Spence et al. (2008)	15[a]	4	RCT	0.10	−0.39	0.58	0.695
Spence and Grant (2007)	40	10	RCT	0.29	−0.33	0.91	0.355
Total	*2090*	–	–	*0.66*	*0.39*	*0.93*	*0.000*

Notes: N = analyzed sample size; g = Hedges' g; CI (95%) = 95% confidence interval for g; RCT = randomized control trial; QEF = quasi-experimental field-study in which participants were non randomly allocated to experimental and control groups; and WSD = within-subjects design without control group, which includes both pre- and post-intervention measures.

[a]Only the C-MT condition (see original) was included for the calculation of effect sizes.

Table 3. Weighted effect sizes of coaching interventions on all outcome categories.

	k	N	g	CI (95%) Lower	CI (95%) Upper	p	Q	I^2
Performance/skills	6	2007	0.60	0.04	1.16	0.036	112.24*	95.55
Well-being	10	564	0.46	0.28	0.62	0.000	7.72	0.00[a]
Coping	10	1703	0.43	0.25	0.61	0.000	6.36	43.01
Attitudes	7	507	0.54	0.34	0.73	0.000	8.64	30.51
Self-regulation	11	789	0.74	0.42	1.06	0.000	81.27*	53.38

Notes: k = number of studies included in the analysis; N = total sample size in k studies; g = Hedges' g; CI = 95% random effects confidence intervals; and Q = Cochran Q statistic. I^2 = the proportion of total variation in the estimates of treatment effect that is due to heterogeneity between studies.
*indicates that between-study heterogeneity significant at $\alpha = 0.000$.
[a]The I^2 was truncated to zero because the Q statistic used for the computation of I^2 was smaller than its degrees of freedom.

The weighted point estimates of the effect sizes seem to indicate that the effect sizes of studies that used a within-subject design are larger than the effect sizes of the studies that used an independent-group design. Subgroup analysis indicated that the overall effect size (aggregated over outcomes) of studies that used a within-subjects design (g = 1.15, 95% CI, 0.46–1.83) was significantly larger than the effect size of studies that used a mixed design (g = 0.39, 95% CI, 0.22–0.56, p = 0.036). These results imply that study design has a considerable influence on the relation between coaching interventions and individual-level outcomes (Hunter & Schmidt, 2004). We will further reflect on these results in the discussion section.

The influence of the number of coaching sessions

Table 5 displays the differences in effect sizes for studies that differ in terms of the number of coaching sessions. The choice for the comparison of studies with less or equal to five sessions vs. more than five sessions emerged from the available study data with some studies reporting a maximum of five coaching sessions whereas other studies reporting more than this maximum.

The weighted point estimates of the effect sizes show a mixed picture. Although a larger number of coaching sessions seems to be beneficial for both coping and goal-directed self-regulation outcomes, the reversed pattern is observed for work/career attitudes and performance/skills (higher effect sizes for a smaller number of sessions). A meta-regression in which the number of coaching sessions was entered as a predictor of the weighted effect sizes (aggregated over outcomes) revealed no significant effects. These results indicate that the number of coaching sessions is not related to the effectiveness of the interventions.

Table 4. Weighted effect sizes of coaching interventions on all outcome categories for different study designs.

	k	N	g	CI (95%) Lower	CI (95%) Upper	p	Q	I^2
Performance/skills								
IDG	5	1372	0.18	0.04	0.32	0.013	5.51	27.43
Within-subjects	1	100	2.33	1.93	2.72	0.000	–[b]	–[b]
Well-being								
IDG	7	394	0.39	0.18	0.60	0.000	3.91	0.00[a]
Within-subjects	3	116	0.54	0.26	0.82	0.002	3.12	35.94
Coping								
IDG	7	325	0.26	0.15	0.37	0.000	11.67	48.57
Within-subjects	3	116	0.47	0.20	0.74	0.001	2.16	7.45
Attitudes								
IDG	6	169	0.48	0.29	0.67	0.000	6.23	19.77
Within-subjects	1	67	0.74	0.47	1.01	0.000	–[b]	–[b]
Self-regulation								
IDG	9	1565	0.32	0.21	0.43	0.000	37.37*	78.59
Within-subjects	2	116	1.33	0.85	1.81	0.035	–[b]	0.00[a]

Notes: k = number of studies included in the analysis; N = total sample size in k studies; g = Hedges' g; CI = 95% random effects confidence intervals; Q = Cochran Q statistic; and I^2 = the proportion of total variation in the estimates of treatment effect that is due to heterogeneity between studies.
*Indicates that between-study heterogeneity significant at $\alpha = 0.000$.
[a]The I^2 was truncated to zero because the Q statistic used for the computation of I^2 was smaller then it's degrees of freedom.
[b]The statistics could not be computed because not enough data was available.

Table 5. Weighted effect sizes of coaching interventions for studies that differ in terms of the number of coaching sessions.

	k	N	g	CI (95%) Lower	CI (95%) Upper	Q	I^2
Performance/skills							
≤5 sessions	3	1329	0.26	0.02	0.53	0.00	58.57
>5 sessions	2	110	0.11[b]	−0.09	0.32	0.00	0.00[a]
Well-being							
≤5 sessions	4	195	0.47	0.15	0.79	5.83	38.36
>5 sessions	7	296	0.46	0.22	0.69	0.00	0.00[a]
Coping							
≤5 sessions	5	1429	0.35	0.11	0.59	0.00	50.33
>5 sessions	5	182	0.54	0.30	0.79	0.00	0.00[a]
Attitudes							
≤5 sessions	3	126	0.67	0.41	0.94	19.14	44.39
>5 sessions	4	150	0.35	0.08	0.61	0.75	0.00[a]
Self-regulation							
≤5 sessions	6	1457	0.52	0.15	0.88	6.962	79.94
>5 sessions	5	215	1.02	0.67	1.36	0.00	32.18

Notes: k = number of studies included in the analysis; g = Hedges' g; CI = 95% random effects confidence intervals; Q = Cochran Q statistic; and I^2 = the proportion of total variation in the estimates of treatment effect that is due to heterogeneity between studies.
[a]The I^2 was truncated to zero because the Q statistic used for the computation of I^2 was smaller then it's degrees of freedom.
[b]One study reported an average of 50 coaching sessions and was therefore considered an outlier and excluded from the meta-regression analysis.

Publication bias

A visual evaluation of the funnel plot did not reveal obvious evidence of publication bias. Additionally, the Egger intercept was non-significant (p = 0.08). However, a visual inspection of funnel plots as well as tests for funnel asymmetry may produce unreliable results when a small number of studies is included in the analysis, especially when heterogeneity is substantial (Sterne et al., 2011). Therefore, we will further address the issue of potential publication bias in the discussion section.

Discussion

Summary of findings

This meta-analysis aimed to provide insight into the possible beneficial effects of coaching within an organizational context. We examined relationships between coaching interventions and several individual-level outcomes that are relevant for both individuals and organizations. The results show that coaching has significant positive effects on performance and skills, well-being, coping, work attitudes, and goal-directed self-regulation. In general, our meta-analytic findings indicate that coaching is an effective tool for improving the functioning of individuals in organizations.

We should note, however, that an examination of the between-study heterogeneity showed that the effects of coaching interventions varied considerably between studies (especially in the performance and skills and goal-directed self-regulation outcome categories). This heterogeneity could be – at least partially – attributed to the relatively small number of studies in the analysis. Alternatively, heterogeneity could also signal the presence of moderating factors (Higgins & Thompson, 2002). The findings of our exploratory analyses indeed suggest that research design could be one of these moderating factors. Studies using a within-subject design generally displayed stronger positive effects of coaching interventions than studies using an independent-group (only one study) or mixed design.

A possible explanation for this finding is that studies using a mixed-design controlled for additional sources of bias in comparison with studies using a within-subjects design (Morris & DeShon, 2002). More specifically, the addition of a control group allows the researcher to control for the natural maturation of participants over time and for selection effects (Cook & Campbell, 1979). At the very least, these results show that the choice for a specific study design has considerable implications for the conclusions that can be drawn with regard to the effectiveness of coaching interventions.

An examination of the results regarding the intensity of the coaching intervention suggests that a greater number of coaching sessions does not necessarily result in stronger positive effects. This, somewhat counter-intuitive, pattern of results might be caused by a spurious correlation. That is, people with less serious or complex problems may need fewer sessions and experience more positive effects of coaching than people with serious and/or complex problems. Alternatively, it could be explained by type of coaching interventions that were applied in the majority of studies in which the number of coaching sessions was small, namely solution-focused coaching. Solution-focused coaching originates in *brief family therapy* (de Shazer, 1988) and differs from other

forms of therapy and coaching by its premise that there is no need for an extensive analysis and understanding of problems in order to create solutions (Berg & Szabo, 2005; Grant & O'Connor, 2010). Therefore, it is possible to jump directly to the ultimate aim of coaching, namely the identification of solutions, potentially resulting in a smaller number of sessions needed to make progression (Kim, 2008). Future research could investigate whether solution-focused coaching is indeed more effective than other coaching approaches and whether specific coaching effects also depend on significance and/or complexity of coaches' problems.

Although the results should be interpreted with caution because of the exploratory nature of the analysis, the finding that coaching can be effective even when the number of coaching sessions is relatively small is encouraging for organizations and individuals in need of coaching. However, while the difference in the number of sessions does not seem to impact the mean effect size, the examination of the heterogeneity statistics does show that there is less variability in the effect sizes for studies using a larger amount of sessions. In other words, the robustness of the effects of coaching seems to increase with the number of coaching sessions. This finding corroborates research on adult learning which suggests that deeper levels of learning (e.g. transformative learning; Mezirow, 1991) only occur when there are sufficient opportunities for critical reflection and active experimentation.

Future research: a need for theoretical enrichment

It is our hope that future research will not only continue to examine whether coaching is effective, but also respond to the need for more theoretical development in coaching psychology. A strong theoretical framework is expedient from both an empirical and a practical perspectives (Grant, 2010; Spence & Oaedes, 2011). For scholars working in the field of coaching psychology, a strong theoretical foundation could purposefully guide the construction of cumulative knowledge. For practitioners, insight into how (rather than if) coaching works can provide guidelines for the improvement of extant coaching interventions and the development of new interventions.

One way in which theoretical enrichment of the coaching literature could be facilitated is by incorporating theoretical perspectives from several sub-disciplines of psychology (Grant, 2010), particularly from research into related fields of developmental interactions such as therapy, mentoring, and training (D'Abate, Eddy, & Tannenbaum, 2003). More specifically, the relative theoretical richness of these fields may serve as a source of inspiration for theoretical enrichment in four interrelated areas of coaching research: the design of coaching interventions, individual characteristics of the coach and the coachee, and the relationship between the coach and the coachee. We provide some specific suggestions for each of these areas below.

Research concerning the design of coaching interventions may benefit from the literature on training and mentoring which draws heavily on educational psychology and theories on (adult) learning (e.g. theory on transformative learning; Mezirow, 1991). Theories on adult learning and its underlying mechanisms can provide insights that are relevant for increasing the 'transfer of coaching' (i.e. long-term effectiveness of coaching interventions). Furthermore, Spence and Oaedes (2011) have suggested that Deci and Ryans' (1985) self-determination theory (SDT) is a valuable theoretical framework for future research on the design of coaching interventions. Central constructs of SDT such as goal-setting, intrinsic motivation and the human needs of competence, relatedness, and autonomy, are crucial for facilitating durable change within coaches (Ryan, Lynch, Vansteenkiste, & Deci, 2011).

Research concerning the characteristics of coaches may find a valuable starting point in the therapy literature. For example, studies investigating the personal characteristics of effective therapists have shown that individual characteristics such as (perceived) empathy are important predictors of therapy outcomes (Burns & Nolen-Hoeksema, 1992; Elliott, Bohart, Watson, & Greenberg, 2011). Recent work in the field of executive coaching indeed suggests that non-specific factors such as understanding, encouraging, and listening behaviors of the coach may be better predictors of coaching effectiveness than specific factors such as the coaching methodology (de Haan, Culpin, & Curd, 2011). In this light, the influence of constructs related to coaches' ability to perceive and manage the emotional states of coaches, such as emotional intelligence (Salovey & Mayer, 1989), seems especially relevant to examine in future research.

Research concerning the characteristics of coaches may explore the concept of 'coachability' that originates in the sports psychology literature. Coachability is a multidimensional construct that reflects the combination of personality traits (e.g. agreeableness, openness to experience) and motivational components (e.g. achievement motivation) needed to improve functioning and performance (Giacobbi, 2000). Furthermore, therapy research has shown that outcome expectations and self-efficacy of clients (coaches) play an important role in the effectiveness of therapeutic interventions (e.g. Goldin et al., 2012). These constructs will be of similar importance in the context of coaching.

Finally, studies on coaching and therapy have shown that the relationship (working alliance) between a coach and a coachee (therapist and client) has considerable implications for the effectiveness of interventions (Baron

& Morin, 2009; Del Re, Horvath, Flückiger, Symonds, & Wampold, 2012). With this in mind, both the literature on similarity attraction (Byrne, 1971) and interpersonal trust (Mayer, Davis, & Schoorman, 1995) can be used as theoretical frameworks to examine how functional relationships between a coach and a coachee can be established and sustained.

Limitations

Five limitations of the current study should be mentioned. First, the majority of the studies included in this meta-analysis relied on self-reports of outcome measures. According to Peterson (1993), there is a considerable inconsistency between self-reports and other-reports (e.g. by the supervisor or coach) when evaluating change in the coachee: self-reports tend to overestimate the effects of coaching interventions. Hence, self-report measures of performance seem troublesome (Podsakoff & Organ, 1986). Therefore, future studies on coaching should rely less on self-reports and should include other sources for measuring coaching outcomes such as 360 feedback (see Smither, London, Flautt, & Fargas, 2003 for an example) as well as tangible results.

Another problem with self-reports is that it is difficult to establish actual change on the outcome measure (alpha change) rather than respondents' redefinition of the rating scale (beta change) and/or the concept that is measured (gamma change). Both beta and gamma changes are due to a shifting conceptualization of the outcome as a result of coaching (Peterson, 1993). It should be noted, however, that also beta and gamma changes can be conceived as relevant outcomes of coaching. Transformative learning theory states that existing belief systems and frames of reference need to be challenged before deep-level changes will occur (Mezirow, 1991). The ultimate aim of coaching is to facilitate deep-level changes and learning (de Haan et al., 2011). Therefore, more insight into alpha, beta, and gamma changes and their underlying cognitive structures (Thompson & Hunt, 1996) is needed because this may help researchers and practitioners to better design a coaching intervention and measure its impact.

A second limitation is that most studies in our meta-analysis did not measure coaching effectiveness over time (at multiple time-points), making it difficult to assess the long-term impact of coaching interventions. Third, the focus on individual-level benefits of coaching in the studies included in our analysis neglected possible 'spillover' effects that coaching could have on other people within an organization. For example, if the coachee is an executive and his or her coaching results in improved leadership skills (the functioning of) subordinates and co-workers will benefit as well. Future research on the effectiveness of coaching could include subordinate and co-worker perceptions so as to assess the indirect effects of coaching.

A fourth limitation of this study is that the findings are based on a relatively small number of studies. Although we did not find any evidence for publication bias, Sterne et al. (2011) noted that analyses for publication bias could produce unreliable results when the number of studies is small and heterogeneity across studies is substantial. For this reason, our findings should be interpreted with caution. At the same time, however, our findings consistently showed effectiveness of coaching across a broad spectrum of outcome measures. Also, our sensitivity analysis indicated that the removal of the study by Peterson (1993) did not alter our conclusions.

Fifth, the 1243 participants in the study by Smither et al. (2003) account for a large proportion of the total number of participants in the studies we examined. However, since the effect sizes in this study were much smaller than the average effect sizes over all studies, the inclusion of the study of Smither et al. has resulted in a conservative rather than optimistic estimation of the effectiveness of coaching.

Finally, the general lack of empirical work on coaching and its weak theoretical foundation has resulted in a large variety of coaching interventions and outcomes. As a consequence, the number of comparable studies suitable for a meta-analytic synthesis was relatively limited.

Conclusion

Despite its limitations, the current meta-analysis indicates that coaching can be effectively used as an intervention in organizations. Furthermore, this study has pointed out several methodological issues that need to be addressed in future studies on coaching effectiveness. The biggest overall limitation of the coaching literature is the lack of rigorous examinations showing the causal mechanisms by which coaching interventions are effective. Thus, we agree with Fillery-Travis and Lane (2006) that it is the time to shift attention from the question 'does it work?' to 'how does it work?'. This second question can only be answered by building a firm theoretical framework that can be used to identify the underlying mechanisms and processes.

Notes

1. While some scholars explicitly distinguish personal coaching from organizational coaching[2] (e.g. Grant, 2010), we take the position that this distinction will be less clear in practice. For example, when coaching concerns stress issues, it often taps into the domain of work–life balance and the intended changes and outcomes will thus affect both the professional and the personal functioning of the coached. Therefore, we do not make this distinction in our endeavor to answer the question whether coaching is effective.

2. For the transformation of effect sizes based on change scores standard deviations (repeated measure designs) into the referent effect size index (based on post intervention score standard deviations), the comprehensive meta-analysis (CMA) software's' default option (r = 0.5) was used. As recommended by Morris (2000) and Borenstein et al. (2009), additional analyses based on different values for r (i.e. r = 0.1 and r = 0.9) were conducted and these demonstrated similar results. The interested reader can contact the first author for more information.

References

Baron, L., & Morin, L. (2009). The coach–coachee relationship in executive coaching: A field study. *Human Resource Development Quarterly, 20*, 85–106.

Berg, I. K., & Szabo, P. (2005). *Brief coaching for lasting solutions*. New York, NY: W. W. Norton & Company, Inc.

Bond, F. W., & Bunce, D. (2003). The role of acceptance and job control in mental health, job satisfaction, and work performance. *Journal of Applied Psychology, 88*, 1057–1067.

Bono, J. E., Purvanova, R. K., Towler, A. J., & Peterson, D. B. (2009). A survey of executive coaching practices. *Personnel Psychology, 62*, 361–404.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2005). *Comprehensive meta-analysis version 2* [*Computer software*]. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2006). Comprehensive meta-analysis (Version 2.2.027) [Computer software]. *Organizational Research Methods, 11*, 188–191.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Cornwall: Wiley.

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

Brock, V. G. (2008). *Grounded theory of the roots and emergence of coaching* (Unpublished Doctoral dissertation). International University of Professional Studies. Retrieved from http://libraryofprofessionalcoaching.com/wp-content/uploads/2011/10/dissertation.pdf

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*, 822–842.

Brunn, E., & Milczarek, M. (2007). *Expert forecast on emerging psychosocial risks related to occupational safety and health* (European Risk Observatory Report). Retrieved from http://osha.europa.eu/en/publications/reports/7807118/

Brunning, H. (2006). *Executive coaching: Systems-psychodynamic perspective*. London: Karnac.

Burns, D. D., & Nolen-Hoeksema, S. (1992). Therapeutic empathy and recovery from depression in cognitive-behavioral therapy: A structural equation model. *Journal of Consulting and Clinical Psychology, 60*, 441–449.

Burton, W. N., Conti, D. J., Chen, C. Y., Schultz, A. B., & Edington, D. W. (1999). The role of health risk factors and disease on worker productivity. *Journal of Occupational and Environmental Medicine, 41*, 863–877.

Byrne, D. E. (1971). *The attraction paradigm*. New York, NY: Academic Press.

Carlson, K. D., & Schmidt, F. L. (1999). Impact of experimental design on effect size: Findings from the research literature on training. *Journal of Applied Psychology, 84*, 851–862.

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

Cochran, W. G. (1954). Some methods for strengthening the common χ^2 tests. *Biometrics, 10*, 417–451.

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

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field studies*. Boston, MA: Houghton Mifflin Company.

Couto, D., & Kauffman, C. (2009). What can coaches do for you? *Harvard Business Review*, 1–8. Retrieved from http://hbr.org/2009/01/what-can-coaches-do-for-you/ar/1

Cox, E., Bachkirova, T., & Clutterbuck, D. (2009). *The complete handbook of coaching*. London: Sage.

D'Abate, C., Eddy, E., & Tannenbaum, S. T. (2003). What's in a name? A literature-based approach to understanding mentoring, coaching, and other constructs that describe developmental interactions. *Human Resource Development Review, 2*, 360–384.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum.

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*, 227–268.

De Haan, E., Culpin, V., & Curd, J. (2011). Executive coaching in practice: What determines helpfulness for clients of coaching? *Personnel Review, 40*, 24–44.

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

Del Re, A. C., Horvath, A. O., Flückiger, C., Symonds, D., & Wampold, B. E. (2012). Therapist effects in the therapeutic alliance-outcome relationship: A restricted-maximum likelihood meta-analysis. *Clinical Psychology Review, 32*, 642–649.

De Shazer, S. (1988). *Clues: Investigating solutions in brief therapy*. New York, NY: W.W. Norton & Company Inc.

Duijts, S. F. A. P., Kant, I. P., van den Brandt, P. A. P., & Swaen, G. M. H. P. (2008). Effectiveness of a preventive coaching intervention for employees at risk for sickness absence due to psychosocial health complaints: Results of a randomized controlled trial. *Journal of Occupational & Environmental Medicine, 50*, 765–776.

*Egan, T., & Song, Z. (2005). *A longitudinal quasi-experiment on the impact of executive coaching*. Paper presented at the 20th Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629–634.

Elliott, R., Bohart, A. C., Watson, J. C., & Greenberg, L. S. (2011). Empathy. *Psychotherapy, 48*, 43–49.

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

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

Fillery-Travis, A., & Lane, D. (2006). Does coaching work or are we asking the wrong question? *International Coaching Psychology Review, 1*, 23–35.

*Finn, F. A. (2007). *Leadership development through executive coaching: The effects on leaders' psychological states and transformational leadership behavior* (Unpublished Doctoral dissertation). Queensland University of Technology. Retrieved from http://eprints.qut.edu.au/17001/

Freire, T. (2013). Positive psychology approaches. In J. Passmore, D. Peterson, & T. Freire (Eds.), *The Wiley-Blackwell handbook of the psychology of coaching and mentoring* (pp. 426–442). West Sussex: Wiley-Blackwell.

Fusar-Poli, P., Bonoldi, I., Yung, A. R., Borgwardt, S., Kempton, M. J., Valmaggia, L., & McGuire, P. (2012). Predicting psychosis: Meta-analysis of transition outcomes in individuals at high clinical risk. *Archives of General Psychiatry, 69*, 220.

Gallwey, T. W. (1974). *The inner game of tennis* (1st ed.). New York, NY: Random House.

Giacobbi, P. R. (2000). *The athletic coachability scale: Construct conceptualization and psychometric analyses* (Doctoral dissertation). University of Tennessee, Knoxville. Retrieved from http://sunzi.lib.hku.hk/ER/detail/hkul/2688806

Goldin, P. R., Ziv, M., Jazaieri, H., Werner, K., Kraemer, H., Heimberg, R. G., & Gross, J. J. (2012). Cognitive reappraisal self-efficacy mediates the effects of individual cognitive-behavioral therapy for social anxiety disorder. *Journal of Consulting and Clinical Psychology, 80*, 1034–1043.

Grant, A. M. (2001). *Towards a psychology of coaching: The impact of coaching on metacognition, mental health and goal attainment* (Unpublished Doctoral dissertation). Macquarie University, Sydney. Retrieved from http://www.eric.ed.gov/PDFS/ED478147.pdf

*Grant, A. M. (2003). The impact of life coaching on goal-attainment, metacognition and mental health. *Social Behavior and Personality, 31*, 253–264.

Grant, A. M. (2008). Personal life coaching for coaches-in-training enhances goal attainment, insight and learning. *Coaching: An International Journal of Theory, Research and Practice, 1*, 54–70.

*Grant, A. M. (2010). The development of coaching psychology. *International Coaching Psychology Review, 1*, 12–22.

Grant, A. M. (2013). The efficacy of coaching. In J. Passmore, D. Peterson, & T. Freire (Eds.), *Handbook of the psychology of coaching and mentoring* (pp. 15–39). West Sussex: Wiley-Blackwell.

Grant, A. M., & Cavanagh, M. J. (2004). Toward a profession of coaching: Sixty-five years of progress and challenges for the future. *International Journal of Evidence-based Coaching and Mentoring, 2*, 1–16.

Grant, A. M., & Cavanagh, M. J. (2007). Evidence-based coaching: Flourishing or languishing? *Australian Psychologist, 42*, 239–254.

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

*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, 3*, 151–168.

Grant, A. M., & O'Connor, S. A. (2010). The differential effects of solution-focused and problem-focused coaching questions: A pilot study with implications for practice. *Industrial and Commercial Training, 42*, 102–111.

Grant, A. M., Passmore, J., Cavanagh, M., & Parker, H. (2010). The state of play in coaching. *International Review of Industrial & Organizational Psychology, 25*, 125–168.

*Green, L. S., Grant, A. M., & Rynsaardt, J. (2007). Evidence-based life coaching for senior high school students: Building hardiness and hope. *International Coaching Psychology Review, 2*, 24–32.

*Green, L. S., Oades, L. G., & Grant, A. M. (2006). Cognitive-behavioural, solution-focused life coaching: Enhancing goal striving, well-being and hope. *Journal of Positive Psychology, 1*, 142–149.

Greenhaus, J. H., Parasuraman, S., & Wormley, W. M. (1990). Effects of race on organizational experience, job performance evaluations, and career outcomes. *Academy of Management Journal, 33*, 64–86.

Hart, V., Blattner, J., & Leipsic, S. (2001). Coaching versus therapy: A perspective. *Consulting Psychology Journal: Practice & Research, 53*, 229–237.

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

Hedges, L. V., & Cooper, H. (1994). *The handbook of research synthesis*. New York, NY: Sage.

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*, 1539–1558.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*, 557–560.

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

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.

International Coach Federation. (2012). Retrieved May, 16, 2012 from, http://www.coachfederation.org/about-icf/overview/

Johnson, B. T., Mullen, B., & Salas, E. (1995). Comparison of three major meta-analytic approaches. *Journal of Applied Psychology, 80*, 94–106.

Judge, T. A., Thoresen, C. J., Bono, J. E., & Patton, G. K. (2001). The job satisfaction-job performance relationship: A qualitative and quantitative review. *Psychological Bulletin, 127*, 376–407.

Kampa-Kokesch, S., & Anderson, M. Z. (2001). Executive coaching: A comprehensive review of the literature. *Consulting Psychology Journal: Practice and Research, 53*, 205–228.

Kilburg, R. R. (1996). Toward a conceptual understanding and definition of executive Coaching. *Consulting Psychology Journal: Practice and Research, 48*, 134–144.

Kim, J. S. (2008). Examining the effectiveness of solution-focused brief therapy: A meta-analysis. *Research on Social Work Practice, 18*, 107–116.

*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.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.

Latham, G. P. (2007). Theory and research on coaching practices. *Australian Psychologist, 42*, 268–270.

Leonard-Cross, E. (2010). Developmental coaching: Business benefit – Fact or fad? An evaluative study to explore the impact of coaching in the workplace. *International Coaching Psychology Review, 5*, 36–47.

Linley, P. A., & Harrington, S. (2005). Positive psychology and coaching psychology: Perspectives on integration. *The Coaching Psychologist, 1*, 13–14.

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.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression Anxiety stress scales (DASS) with the beck depression and anxiety inventories. *Behaviour Research and Therapy, 33*, 335–343.

*Luthans, F., & Peterson, S. J. (2004). 360-degree feedback with systematic coaching: Empirical analysis suggests a winning combination. *Human Resource Management, 42*, 243–256.

Maslach, C., & Jackson, S. E. (1986). *MBI: Maslach Burnout Inventory manual* (research edition) (1st ed. 1981). Palo Alto, CA: Consulting Psychologists Press.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review, 20*, 709–734.

McGaw, B., & Glass, G. (1980). Choice of the metric for effect size in meta-analysis. *American Educational Research Journal, 17*, 325–337.

Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco, CA: Jossey-Bass.

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

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*, 264–269.

Morris, S. B. (2000). Distribution of the standardized mean change effect size for meta-analysis on repeated measures. *British Journal of Mathematical and Statistical Psychology, 53*(1), 17–29.

Morris, S. B., & DeShon, R. P. (1997). Correcting effect sizes computed from factorial ANOVA for use in meta-analysis. *Psychological Methods, 2*, 192–199.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*, 105–125.

Mullen, E. J. (1994). Framing the mentoring relationship as an information exchange. *Human Resource Management Review, 4*, 257–281.

National Research Council (1992). *Combining information: Statistical issues and opportunities for research*. Washington, WA: National Academy Press.

Organ, D. W., & Ryan, K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behavior. *Personnel Psychology, 48*, 775–802.

Passmore, J. (Ed.). (2010). *Excellence in coaching: The industry guide*. London: Kogan Page.

Passmore, J., & Fillery-Travis, A. (2011). A critical review of executive coaching research: A decade of progress and what's to come. *Coaching: An International Journal of Theory, Practice & Research, 4*, 70–88.

*Peterson, D. B. (1993). *Measuring change: A psychometric approach to evaluating individual training outcomes*. Symposium conducted at the Eighth Annual Conference of the Society for Industrial and Organizational Psychologists, San Francisco, CA.

Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management, 12*, 531–544.

*Poepsel, M. A. (2011). *The impact of an online evidence-based coaching program on goal striving, subjective well-being, and level of hope* (Unpublished Doctoral dissertation), Minneapolis, MN. Harold Abel School of Social and Behavioral Sciences. Retrieved from http://gradworks.umi.com/3456769.pdf

Porter, L. W., Steers, R. M., Mowday, R. T., & Boulian, P. V. (1974). Organizational commitment, job satisfaction, and turnover among psychiatric technicians. *Journal of Applied Psychology, 59*, 603–609.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. ed.). Newbury Park, CA: Sage.

Ryan, R. M., Lynch, M. F., Vansteenkiste, M., & Deci, E. L. (2011). Motivation and autonomy in counseling, psychotherapy, and behavior change: A look at theory and practice. *The Counseling Psychologist, 39*, 193–260.

Salovey, P., & Mayer, J. D. (1989). Emotional intelligence. *Imagination, Cognition and Personality, 9*, 185–211.

Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35–37). Windsor: NFER-NELSON.

Seligman, M. E. (2007). Coaching and positive psychology. *Australian Psychologist, 42*, 266–267.

Seligman, M. E., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55*(1), 5–14.

Smith, P. C., Kendall, L. M., & Hulin, C. L. (1969). *The measurement of satisfaction in work and retirement*. Chicago: Rand McNally.

*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*, 23–44.

Spence, G. (2008). *New directions in evidence-based coaching: Investigations into the impact of mindfulness training on goal attainment and well-being* (Unpublished Doctoral dissertation). University of Sydney, Sydney. Retrieved from http://ses.library.usyd.edu.au/bitstream/2123/2469/1/New%20Directions%20in%20the%20Psychology%20of%20Coaching%20(Spence,%202006).pdf

Spence, G. B., Cavanagh, M. J., & Grant, A. M. (2008). The integration of mindfulness training and health coaching: An exploratory study. *Coaching: An International Journal of Theory, Research and Practice, 1*, 145–163.

*Spence, G. B., & Grant, A. M. (2007). Professional and peer life coaching and the enhancement of goal striving and well-being: An exploratory study. *Journal of Positive Psychology, 2*, 185–194.

Spence, G. B., & Oaedes, L. G. (2011). Coaching with self-determination in mind: Using theory to advance evidence-

based coaching practice. *International Journal of Evidence Based Coaching and Mentoring, 9*, 37–55.

Sterne, J. A., Sutton, A. J., Ioannidis, J., Terrin, N., Jones, D. R., Lau, J., & Higgins, J. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal, 343*(7818), 302–307.

Sue-Chan, C., & Latham, G. P. (2004). The relative effectiveness of external, peer, and self-coaches. *Applied Psychology, 53*, 260–278.

Tepper, B. J. (1995). Upward maintenance tactics in supervisory mentoring and nonmentoring relationships. *Academy of Management Journal, 38*, 1191–1205.

Thompson, R. C., & Hunt, J. G. (1996). Inside the black box of alpha, beta, and gamma change: Using a cognitive-processing model to assess attitude structure. *Academy of Management Review, 4*, 655–690.

Waldron, V. R. (1991). Achieving communication goals in superior-subordinate relationships: The multi-functionality of upward maintenance tactics. *Communications Monographs, 58*, 289–306.

Whitmore, J. (1992). *Coaching for performance*. San Diego, CA: Pfeiffer.

Wright, T. A., & Cropanzano, R. (2000). Psychological well-being and job satisfaction as predictors of job performance. *Journal of Occupational Health Psychology, 5*, 84–94.

Wrzesniewski, A., & Dutton, J. (2001). Crafting a job: Employees as active crafters of their work. *Academy of Management Review, 26*, 179–201.