

EXECUTIVE COACHING OUTCOME RESEARCH: THE CONTRIBUTION OF COMMON FACTORS SUCH AS RELATIONSHIP, PERSONALITY MATCH, AND SELF-EFFICACY

AQ: au

Erik de Haan, Anna Duckworth, David Birch, and Claire Jones
Ashridge Business School

AQ: 1

This article argues for a new way of studying executive-coaching outcomes, which is illustrated with a study based on data from 156 client–coach pairs. The argument accepts that we are unlikely to get robust data on coaching outcomes in the near future but assumes that we can expect similar effectiveness for coaching as that demonstrated in rigorous psychotherapy outcome research. Therefore, it is argued that it is more important now to (a) identify the “active ingredients” that predict the effectiveness of executive coaching, and (b) to determine the difference in predictive value of these active ingredients on coaching effectiveness. The outcome study examined some of these active ingredients, such as the working alliance between coach and client, the self-efficacy of the client, the personality of the client, and the “personality match” between client and coach. The results show that client perceptions of coaching outcome were significantly related to their perceptions of the working alliance, client self-efficacy, and to client perceptions of the range of techniques of the coach. The client–coach relationship mediated the impact of self-efficacy and range of techniques on coaching outcomes, suggesting that this relationship is the key factor in determining how clients perceive the outcome of coaching.

Keywords: executive coaching, coaching outcomes, leadership development, client–coach relationship

As our children teach us, often the most valuable questions in life are both very easy to ask and amazingly hard to answer: “Why did I do that?” “What did I mean by that?” or “What am I achieving here?” are some examples. In executive coaching, examples of these obvious and essential questions

Erik de Haan, Anna Duckworth, David Birch, and Claire Jones, Ashridge Centre for Coaching, Ashridge Business School, Berkhamsted, Hertfordshire, United Kingdom.

We thank our colleague Philippa Hardman for her contribution at the start of this research. Many thanks to Judy Curd, Paul Bassett, and Nadine Page for help with the processing of the data and the statistical analyses. Thanks also to Rebecca Jones for her thorough and generous contribution to our literature survey. Special thanks to Rob Kaiser, the Associate Editor of this journal, who made countless helpful comments and offered direct improvements for certain passages.

AQ: 11

Correspondence concerning this article should be addressed to Erik de Haan, Ashridge Centre for Coaching, Ashridge Business School, Berkhamsted, Hertfordshire, HP4 1NS United Kingdom. E-mail: Erik.Dehaan@ashridge.org.uk

that are, at the same time, profoundly difficult to answer, include the following: “Does our coaching work? Does it help clients with their critical objectives?” “What aspects of coaching work? What are the ‘active ingredients’? Under what circumstances do they work best?” and “What intervention would work best here and now, with this client at this moment?”

Thousands of coaches have asked these questions, and all of us are curious about effectiveness or outcome. These questions also occur frequently in the coaching literature; however, it is rare to encounter serious attempts at answering them with anything more than a coach’s opinion or a few carefully selected case studies. We estimate that there are probably fewer than 20 robust quantitative outcome studies throughout the coaching literature. One reason for this is the costly and cumbersome requirements of a rigorous outcome study. Another is that rather than studying, with detachment, their own effectiveness, a coach’s priority is usually to meet their clients’ needs and their own coaching commitments. However, if we do not address these questions, we may find it difficult to justify our fees, difficult to assert unequivocally that coaching conversations are indeed beneficial and difficult to avoid the potential risks of executive coaching, such as misjudging the situation, aggravating the status quo, or abusing our influence (Berglas, 2002).

This article has three aims. First, we briefly summarize the extensive and convincing outcome research findings in the adjacent field of individual psychotherapy and draw lessons from this literature for the study of coaching outcomes. Next, we provide a brief overview of the existing coaching outcome literature, including a few articles that approximate a properly controlled research design to provide for causal inference about the effects of coaching. Finally, we present the results of a coaching outcome study that focuses on the factors that are likely to be the most influential in achieving coaching effectiveness, the so-called common factors, that is, those factors that are common to all approaches to coaching and mentoring.

The Executive Coaching Landscape

We define executive coaching as a form of leadership development that takes place through a series of contracted one-to-one conversations with a qualified “coach.” Executive coaching aspires to be a form of organization and leadership development that results in a high occurrence of relevant, actionable, and timely outcomes for clients. Coaching is tailored to individuals so that they learn and develop through a reflective conversation within an exclusive relationship that is trusting, safe, and supportive. Coaching is therefore substantially different from the more conventional training and development that is characterized by the imparting of actionable information, instruction, and advice. A 2004 survey conducted by the Chartered Institute of Personnel and Development (CIPD) in the United Kingdom reported that 64% of organizations surveyed use external coaches, with 92% of survey participants judging coaching to be “effective” or “very effective” and 96% saying that coaching is an effective way to promote learning in organizations (Jarvis, 2004). In the same year, the *Harvard Business Review* reported that business coaching—including mentoring—was a \$1 billion industry in the United States and \$2 billion industry worldwide (Sherman & Freas, 2004).

When we take a step back and look at the nature of this industry, it is apparent that the coaching field is in a state of flux and is only just beginning to be regulated as a profession. Individuals are entering from very diverse backgrounds, such as senior management, organization development, and counseling. They may retrain as psychologists, psychotherapists, or sports coaches (Peltier, 2001), bringing influences as far apart as the GROW model, solution-focused brief therapy, and person-centered counseling (de Haan & Burger, 2005). This wide range of backgrounds and the plethora of models and approaches mean that individual professionals are practicing in vastly different ways. Given this, an executive-coaching intervention will be tailored not only to the needs and interests of the individual client or “coachee,” but also to those of the individual coach, given his or her particular background, theoretical orientation, and interests. It is also of note that, contrary to other helping professions such as counseling and psychotherapy, executive coaching is commissioned and paid for by a wide range of individual contractors within corporations: sometimes at board level (e.g., as part of a change program), sometimes from within the human resource function (e.g., as part

of a leadership-development program), and sometimes at a more local level, on an individual basis. In terms of Porter's well-known five forces analysis (Porter, 2008), the bargaining power of customers is therefore extremely weak and the bargaining power of suppliers is correspondingly strong. This increases the freedom of executive coaches to approach the coaching sessions as they see fit.

These features of the industry have clear repercussions for research. Whereas most of the services in psychotherapy are centrally commissioned by large health insurance companies or national health services, it is entirely different in executive coaching. As executive coaches, we are finding ourselves in a situation where there is very little pressure from clients to conduct rigorous outcome research and a dearth of funding for this type of research. At the same time, we know from psychotherapy outcome research (see the historical overview in Wampold, 2001) that we are likely to need very high *N*s, possibly well above 10,000, and a rigorous design with randomized controlled trials, to demonstrate beyond doubt that executive coaching is effective—with even greater statistical power needed to differentially explore active ingredients in effectiveness. For the same reasons as outlined here—no pressure from customers and no funding for research—there are, as yet, no rigorous randomized controlled trial studies available in the coaching literature.

In other words, at present, *all* coaching outcome studies are weak by the standards of psychotherapy and general medicine, and there are good, understandable reasons for this. Moreover, there is no likelihood of research funding by large and centrally coordinated bodies in the foreseeable future. It is therefore to be expected that this situation will continue and that we will keep seeing interesting individual studies of effectiveness but no firm conclusions.

In our view, the way forward for quantitative researchers in this field is now to *assume* that the general effectiveness of helping conversations as convincingly demonstrated in psychotherapy (see, e.g., Wampold, 2001, or Cooper, 2008) will also be true in executive coaching. If we then also assume that clients' perceptions of outcome are indeed a meaningful measure of effectiveness (which is also supported by research—see, e.g., Stiles, Barkham, Mellor-Clark, & Connell, 2008), we can proceed by studying the active ingredients in coaching. Interestingly and significantly for our field, within the much more advanced and rigorous psychotherapy outcome literature, there is also a separate place for measuring active ingredients, and this is researched in similar ways (see Chapters 4 and 5 of Wampold's, 2001, authoritative overview).

The study reported here builds on several studies in the field of executive coaching that have tried to identify active ingredients on the basis of client perceptions. Although building on a burgeoning tradition in coaching outcome research, the study extends this trend with a larger sample, a broader range of "ingredients" or "factors" under study, and within real executive-coaching assignments taking place in a cross-section of the marketplace.

Lessons From Psychotherapy Outcome Research

As McKenna and Davis (2009) argue, the coaching field can learn from the fact that in the older and more established profession of psychotherapy, these same questions of effectiveness have been rigorously studied since at least the 1930s (Rosenzweig, 1936). In this tradition, research findings that seemed initially unclear and contradictory have begun to yield convincing results (starting with Smith & Glass, 1977) and are now almost universally accepted among professional practitioners. In summary, the answers to our initial questions, when applied to psychotherapy, are as follows:

- Does psychotherapy work? Yes, in fact, it has been demonstrated that the average psychotherapy client has an outcome higher than 80% of the people in the control group (Wampold, 2001).
- What aspects of psychotherapy work? Different interventions, approaches, models and protocols do not appear to make any difference in effectiveness. The aspects that dominate are *common* to all approaches—for example, client context (what happens outside the therapeutic relationship), therapist characteristics (including empathy, understanding, respect, warmth, and authenticity), and the quality of the relationship between client and

therapist during the session (Cooper, 2008). Common factors¹ are therefore central to effectiveness in psychotherapy.

Fn1

- Under what circumstances do we find differential effects? Not a lot is known yet, but there are strong indications that the therapist's *allegiance* to their approach and the client's *expectations* are more important than was previously thought (Wampold, 2001). These are also common factors.

One could argue that these intriguing findings from psychotherapy are not relevant for coaching because the studies were conducted with professional therapists working clinically with clients suffering from mental health problems such as depression and anxiety, and that this is different from the needs and issues typically addressed in executive coaching. On the other hand, coaching and psychotherapy both belong to a general class of personal interventions which are based on helping relationships. Moreover, the psychotherapy outcome results are very convincing as they are based on meta-analyses of multiple rigorous studies.

AQ: 2

Exemplary Executive Coaching Outcome Research

In our view, the strongest case for the effectiveness of executive coaching comes from two general classes of outcome research, one that includes a control group and one that does not. Although they do not include a control group, the following studies explore the effectiveness of coaching by looking at behavior and performance as dependent variables, which are more tangible and objective than client, coach, or manager attitudes such as satisfaction. Peterson (1993) studied 370 leaders from various organizations at three points in time (precoaching, postcoaching, and follow-up), with outcome defined by their achieving their own coaching objectives and five standard "control" items, rated by at least themselves, their manager, and their coach (multisource ratings). The coaching program was intensive and long-term, with typically 50+ hours of individual coaching with a professional coach over a period of at least a year. Peterson found that clients, on average, achieved significant improvement on all measures of outcome related to coaching objectives (effect sizes $d > 1.5$). Olivero, Bane, and Kopelman (1997) studied managers who had taken part in a 3-day educational training course followed by 8 weeks of coaching. They found that both the training and the coaching increased productivity considerably, with most of the increase attributable to the coaching (increase of 22.4% with training alone and of 88.0% with training and coaching, i.e., almost fourfold; a difference that was significant at the $p < .05$ level). In another study by Thach (2002), 281 managers participated in four 1-hr sessions of coaching over 5 months, with a 360° feedback before and after the coaching. They found an average increase in "leadership effectiveness" both as rated by the coaches and their coworkers (average increase 60% but no significance reported). Bowles, Cunningham, De La Rosa, and Picano (2007) looked at effectiveness in terms of increased productivity in Army recruitment managers ($n = 30$) and executives ($n = 29$) who received coaching compared with productivity changes in a nonrandom group of experienced recruitment managers over a similar, but not contemporaneous, time interval. The individuals who were coached showed greater productivity gains ($d = 0.43$, with $p < .05$, for the middle managers; and $d = 0.75$, with $p < .01$, for the executives). Finally, Perkins (2009) studied the effectiveness of executive coaching on improving leadership behaviors in meetings, as rated by the coach. Using quantitative and qualitative methods with a small sample ($N = 21$), pre- and postmeasurement of meeting behaviors were scored by the

¹ The idea of *common factors* was introduced by Rosenzweig (1936). He argued that if all professional therapies are equally effective, there is a good chance that the ingredients they have in common will determine the effectiveness of therapy—and not the specific interventions of an individual school of therapy. The active ingredients of therapy must therefore be common to all approaches. Examples are the relationship, the setting, the expectations, the personalities of coach and client, the presence of an ideology or approach, and so on.

coach and author, with a clear improvement of behaviors reported (effect sizes $d > 0.95$ for 9 out of 11 behaviors measured, and $p < .01$). Of course, there may have been researcher bias in these scores, as coaches might understandably want their clients to do well.

Even greater confidence in the impact of coaching comes from three studies that used behavioral and performance outcome measures and also employed a control group. Sue-Chan and Latham (2004) compared the impact of peer and faculty coaches and “self-coaching” on two master of business administration (MBA) programs with a wide difference in reputation in terms of (perceived) expertise and credibility. This outcome study involved MBA students in two countries (total $N = 53$) and compared performance in terms of team playing and exam grades and found small, but statistically significant, differences at $p < .05$, between faculty, peer, and self-coaching, with the first the most impactful. As in Perkins (2009), this study may suffer from researcher bias, as the external coaches/tutors did the scoring of team playing effectiveness.

Evers, Brouwers, and Tomic (2006) measured self-efficacy beliefs and outcome expectancies with three items for each. Their study compared a preintervention and postintervention measurement and also involved a (nonrandomized) control group. The intervention was short, with an average of only four coaching sessions. Although the sample was not very large (30 managers in both the experimental and the control group), they did find some objective evidence for a positive outcome of the coaching intervention. There was a significant increment for the coached group over the control group for one of the three items in both self-efficacy beliefs (“setting one’s own goals”) and outcome expectancies (“acting in a balanced way”), $d \approx 0.5$, with $p < .05$.

Smither, London, Flautt, Vargas, and Kucine (2003) conducted one of the most thorough studies on the impact of executive coaching to date. This study included a (nonrandomized) control group, and conclusions were based on multiple criteria that included evaluations by independent researchers together with clients’ superiors, peers, and subordinates. This research involved 1,202 senior managers in one multinational organization and two consecutive years of 360° feedback. However, there were no more than “two or three” coaching sessions per client (Smither et al., 2003, p. 29). The researchers found that managers who worked with an executive coach were significantly more likely than managers who did not work with a coach to (a) set specific goals ($d = 0.16$; $p < .01$), (b) solicit ideas for improvements from their superiors ($d = 0.36$; $p < .01$), and (c) obtain higher ratings from direct reports and superiors in the second year ($d = 0.17$; $p < .05$).

In summary, we note that outcome research in coaching is still in its infancy and that the holy grail of executive coaching—proof that executive coaching is an effective intervention from a controlled study with random assignment and multiple behavioral and performance outcome measures—is yet to be found. In fact, no clear and agreed sense of what “outcomes” should be included or how they should be measured has yet emerged. Nor is there an agreed research standard like the randomized controlled trials used in psychotherapy outcome research (Wampold, 2001). Furthermore, the studies include a variety of processes that might themselves affect outcomes, such as explicit goal setting, written development objectives, 360° feedback, and other assessment tools, manager involvement, and even training programs and a presentation to senior executives to summarize achievements (e.g., Olivero et al., 1997). That said, what is striking is that the first five research articles discussed here (Peterson, 1993; Olivero et al., 1997; Thach, 2002; Bowles et al., 2007; Perkins, 2009), which did not make use of a contemporary control group, found large effects ($d > 0.75$), generally larger than those found in psychotherapy. On the other hand, the more rigorous studies involving control groups (e.g., Evers et al., 2006; Smither et al., 2003) only found small effects, generally smaller than those found in psychotherapy ($d < 0.5$; compare with average $d \approx 0.80$ in psychotherapy—see Wampold, 2001). However, these were studies with internal coaches and coaching programs of unspecified rigor and structure, whereas many of the studies without control groups involved more significant coaching programs with qualified professional coaches, and this is also a possible contributing factor. It appears that if the client is the sole source of data, the outcome tends to be very positive. However, when such same-source bias is controlled for, the effect is much smaller, although still positive.

A New Form of Coaching Outcome Study

This overview of outcome research in coaching provides some indication that executive coaching is an effective intervention. However, there is also another body of coaching outcome research to which the present study belongs. This newer body of research assumes the general effectiveness of coaching and then compares conditions to determine the degree to which various aspects of coaching, the coach, or the client have an effect on outcomes. If one accepts the assumption of general effectiveness (e.g., as demonstrated by the studies cited here) the experimental conditions of this type of research can be a lot less stringent. First, one does not need to employ randomized controlled groups because the various conditions create proper comparison samples within the study. Second, one can use self-reports because, in psychotherapy, studies using self-report data in realistic settings have consistently corroborated results from more rigorous studies with randomized controlled trials (Shadish, Navarro, Matt, & Phillips, 2000; Stiles et al., 2008). However, these self-reports generally prove a more reactive dependent variable, that is, we need to bear in mind that outcome effects can be overestimated in self-report studies.

We have found six studies that explored the question of *what coaching factors* are effective; in other words, *which* coaching models, personality matches, or coaching behaviors make a significant difference to clients?

Scoular and Linley (2006) looked at how both (a) a “goal-setting” intervention at the beginning of the conversation, and (b) personality (dis)similarities between coach and client in terms of Myers-Briggs Type Inventory (MBTI; Myers, McCaulley, Quenk, & Hammer, 1998) profiles, affected perceived effectiveness. The sample included 117 clients and 14 coaches. No difference resulted for outcome measurements at 2 and 8 weeks after the session between “goal-setting” and “no goal-setting”; but when the coach and client differed more on aspects of their MBTI profiles, the outcome scores were significantly higher.

Stewart, Palmer, Wilkin, and Kerrin (2008) looked at how both client personality and client self-efficacy predicted coaching outcome. They measured the Big Five personality traits (Digman, 1990) and general self-efficacy (Schwarzer, Mueller, & Greenglass, 1999) for 110 clients and correlated these with coaching outcome. They found moderate positive effects for conscientiousness, openness, emotional stability, and general self-efficacy, but cautioned that other factors are likely to play a role as well.

Boyce, Jackson, and Neal (2010) studied 74 client-coach relationships in a U.S. military academy where the clients were cadets and the coaches were senior military leaders who had some previous training in executive coaching. The study analyzed the impact of the client-coach relationship (in terms of rapport, trust, and commitment) and matching criteria (demographic commonality, behavioral compatibility, and coach credibility) on coaching outcome. Their main findings were that matching was not related to outcome, although outcome was significantly related to the relationship as assessed by both the client and the coach. The client ratings were more highly related to outcome than the coach ratings (explained proportion of variance around 50% vs. around 25%). Moreover, the client ratings fully mediated the other dependencies, suggesting that variables such as credibility and compatibility positively impact the coaching outcome through supporting the development of the coaching relationship.

Baron and Morin (2009; 2012) studied 30 internal coach/client pairs involved in a leadership development program at a manufacturing company. They showed that the coaching relationship, as measured by clients’ ratings of the Working Alliance Inventory (Horvath & Greenberg, 1986), predicted coaching outcome in terms of changes in client self-efficacy (with explained proportion of variance around 25%). However, coaches’ ratings of the working alliance did not predict this outcome significantly.

De Haan, Culpin, and Curd (2011) examine how various executive coaching interventions make a difference to clients. Seventy-one coaching clients, from as many organizations, reported on the various coaching interventions, and these ratings were compared with their ratings of the effectiveness of the conversations. There were no effects for specific coach interventions, leading to the conclusion that coaching effectiveness is not a function of specific techniques or interventions so

much as it is a function of factors common to all coaching, such as the quality of the coaching relationship, empathic understanding, positive expectations, and so forth.

Present Study

The present study examines the relative impact and importance of various common factors for 156 executive coaching clients and 34 experienced coaches. The purpose of this research is to study elements common to *all* coaching approaches (i.e., “common factors”) that are likely to have the highest positive impact on clients. If clear results are found, executive coaches should be able to use the findings to improve their own clients’ outcomes.

There are various taxonomies of “common factors,” focusing on relationship-, client-, coach-, change-, and structure-related factors (e.g., de Haan, 2008; Grencavage & Norcross, 1990). Of the various factors proposed, we have chosen as many as possible that can be measured and that have been shown in other outcome studies, across the helping professions, to make a difference. These include coach personality, client personality, a range of coach techniques common to most coaching engagements, relationship quality, and self-efficacy.

We based our predictions on the coaching outcome studies reviewed here and the prevailing idea that common factors are likely to have a differential, high impact on coaching outcomes. The most consistent effect among the common factors studied in coaching outcome research has been for the quality of the client–coach relationship, or the working alliance.²

Fn2

Hypothesis 1: The strength of the coaching relationship will predict coaching outcomes, both (a) as measured by the client, and (b) as measured by their coach.

Scoular and Linley (2006) found that personality dissimilarity as measured by the MBTI was related to coaching outcome, thereby supporting the intuition that a different perspective might help most in coaching. We acknowledge that the MBTI has only “sufficient” levels of reliability and validity and is a poor predictor of managerial effectiveness, job performance, and employee commitment (Gardner & Martinko, 1996). On the other hand, the MBTI is still the most frequently used psychometric instrument and has been used by previous researchers. For example Bell’s (2006) doctoral dissertation uses the MBTI to demonstrate for 67 clients that their profiles of perceptions of effectiveness of coaching across different dimensions vary significantly with MBTI type.

Hypothesis 2: Client–coach personality dissimilarity (as characterized by the MBTI profile) will predict coaching outcomes.

Research by Anderson and Betz (2001) showed that a person’s self-efficacy expectations have a direct bearing on their personal and career development. Sherer and Maddux (1982) also found that personal self-efficacy expectations are often regarded as primary determinants of behavioral change. Within the coaching outcome literature, Stewart et al. (2008) showed that general self-efficacy predicts coaching outcome, and Evers et al. (2006) demonstrated that small improvements in self-efficacy can occur as a result of coaching. For the purpose of our study, we have assumed that general self-efficacy changes very little over the duration of coaching assignments.

Hypothesis 3: General self-efficacy of the client will predict coaching outcome.

De Haan et al. (2011) found that specific coaching techniques and interventions were not related to coaching outcomes, concluding that the full range of techniques that are common to different

² Working alliance, as originally defined by Greenson (1965), is a measure for the strength of the coaching relationship. Bordin (1979) suggested that the working alliance can be thought of as a combination of agreement on tasks, agreement on goals, and strength of bonds. Based on Bordin’s (1979) model, Horvath and Greenberg (1986) designed the Working Alliance Inventory with three subscales: “tasks,” “goals,” and “bonds,” which is now the most widely used of many well-validated tools to measure the working alliance.

AQ:12

methods, theoretical approaches, and specific interventions are more reliably associated with the effectiveness of coaching.

Hypothesis 4: The full range of coach techniques—all coaching behaviors as reported by clients—will predict coaching outcomes, and will do so approximately equally.

AQ: 3

We propose that the client’s perception of the relationship may be the key active ingredient in coaching effectiveness and a determining factor of the influence of the other active ingredients. In other words, if the relationship is bad, none of the other factors can make up for it. On the other hand, if the relationship is strong, it facilitates the effects of the other factors. The results of [Boyce et al. \(2010\)](#) indeed show that, in their data, the strength of the coaching relationship as seen by the client mediated the significant effects found for other variables.

Hypothesis 5: The strength of the coaching relationship mediates (a) the impact of client–coach personality match on coaching outcome, as stated by Hypothesis 2; (b) the impact of general self-efficacy of the client on coaching outcome, as stated by Hypothesis 3; and (c) the impact of perceived coach techniques on coaching outcome, as stated by Hypothesis 4.

Figure 1 shows the various common factors hypothesized to have a positive impact on the outcomes of coaching. This figure also demonstrates how the impact of these common factors may be mediated through the relationship, as predicted by Hypothesis 5.

F1

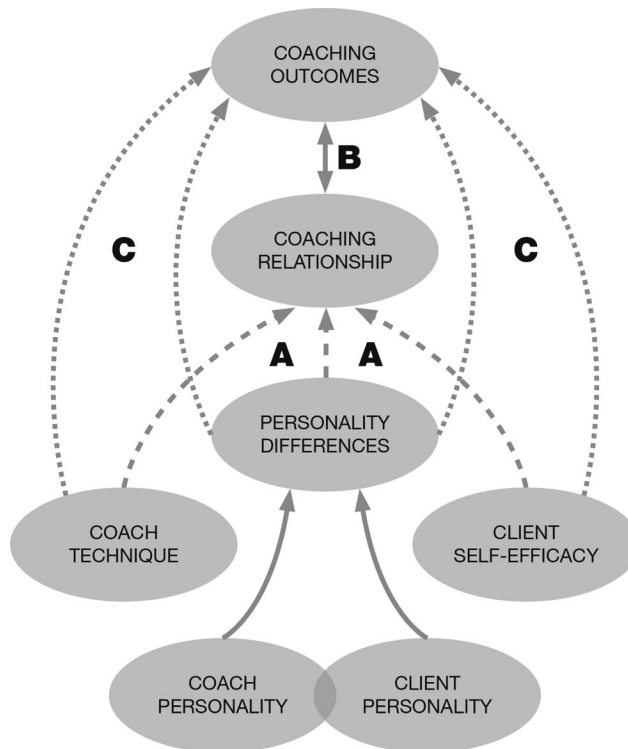


Figure 1. Graphical depiction of the various common factors studied as independent variables. **Note:** In our analysis, we investigate both direct influences of the independent variables on coaching outcomes (Dependencies B and C) and the probability of mediation of this influence through the strongest dependency—the coaching relationship (Dependency A plus B compared with C).

Method

Participants

The participating client–coach pairs were selected through our own networks of experienced and qualified executive coaches. The coaches were employed by different institutions, such as Ashridge Business School and The Oxford School for Coaching and Mentoring OCM. About one third of the coach sample was self-employed. Coaches were told that we were interested in improving the effectiveness of coaching for different personality types (characterized by MBTI profiles). Each coach completed an online “coach survey” and then invited their clients to complete an online “client survey.” It was made clear to both coach and client that their responses would be treated in confidence and that no individual respondent would be identified.

The coach response rate to the survey was 78.6% and the client response rate was 58.4%, a high rate that probably reflects the personal approach used to select respondents. With 34 coaches and 156 clients participating in the survey, it was possible to study 156 coaching relationships. The average experience of the coach was 10.3 years, with a minimum of 3 years and a maximum of 20 years. Seventeen (50%) of the coaches were accredited through the Ashridge Accreditation Process. This process requires the coach to submit a transcript report of a taped segment of their work, a written case study, and a live demonstration of a coaching session (de Haan, 2008). The majority of remaining coaches were accredited elsewhere, with 66% of the total coach sample qualified in the use of MBTI.

Clients were mostly senior and middle managers in large organizations, with a very small minority being coaches or consultants themselves. The number of sessions that coach and client had undertaken at the point that data was collected ranged from 1 to around 100, with a mean of 8.6, with session lengths ranging from around 75 to 120 min. Data collection took place over a 12-month period from August 2008 to August 2009. Approximately 60% of clients and approximately 80% of coaches were based in the United Kingdom, although all coaches and clients were based in (wider) Europe. Most clients were male (60%), whereas the coaches were evenly split.

Coaches were given a small incentive to participate in the research: a free group supervision session at our professional coaching establishment for every 10 clients they recruited for this research (value around £175).

Surveys

The coach survey could be completed in 5 min and the client survey in no more than 20 min. After establishing the relevant background information on gender, MBTI type (administered previously), coaching credentials, and some additional items about their style, the coach survey asked coaches to rate the quality of their relationship with each of their different clients by scoring the statement “I have a very good relationship with this client” on a 7-point response scale from *strongly disagree* to *strongly agree*.

The client survey asked for name, gender, MBTI type, and questions about their coach and coaching journey, including prior expectations. The main body of the questionnaire contained four sections. The first section contained four questions requiring answers on a 7-point response scale about perceived outcomes: “your overall coaching experience,” “coaching adding value,” “impact of coaching on your performance at work,” and “coaching enables you to achieve what you want to achieve”; the average client’s perceived client coaching effectiveness was calculated as the averaged score across these four ratings (four items). The second section explored the extent to which clients perceived that they experienced a range of different interventions from their coach. This section was based upon the well-known Heron (1975) model, comprising six categories of intervention styles, and five items added by the current researchers to widen the scale from just “behaviors” to include more general “interventions.” The interventions included the following: (a) to be advised or told what to do by my coach, (b) for my coach to provide me with information, (c) for my coach to challenge my thoughts or actions, (d) for my coach to help me to make discoveries, (e) for my coach to support me, and (f) for my coach to help me to release emotions. The additional five items were (a) significant progress on my issues through step-by-step change, (b) significant progress on my

issues through critical moments of insight or realization, (c) significant growth relating to outcomes/doing, (d) significant growth relating to behaviors/being, and (e) explicit focus on my most important goals. The final developed scale of a “range of coach techniques” therefore consisted of 11 items. A reliability estimate was calculated for the scale and the Cronbach’s alpha was .86, suggesting a good degree of coherence. Only the removal of the first scale item (“to be advised or told what to do by my coach”) increased the Cronbach’s alpha—to .87. Interestingly, coaches would intuitively agree that advisory interventions such as Item 1 are least present and least helpful in executive coaching. Overall, the full 11-item scale was used in subsequent analyses to represent the range of techniques of coaching. The third section contained an adapted version of the well-established Working Alliance Inventory (Horvath & Greenberg, 1986; 36 items). This instrument is used widely in therapy for measuring the quality of the relationship between therapist and client, and was adapted here to measure the client–coach relationship. Sample items from the 36-item WAI scale include “I believe my coach is genuinely concerned for my welfare,” “I am clear as to what my coach wants me to do in these sessions,” “My coach and I respect each other,” and “I feel my coach is not totally honest about his or her feelings toward me.” Some of the items are reverse-scored, for example, the last one above. In the fourth section, another well-founded instrument was used to establish the client’s self-efficacy (Schwarzer et al., 1999; 10 items on 4-point Likert scale, range 10 to 40). Sample items from the self-efficacy scale are “I can always manage to solve difficult problems if I try hard enough,” “If someone opposes me, I can find the means and ways to get what I want,” and “It is easy for me to stick to my aims and accomplish my goals.” All sections included space for comments or clarification, and there was a final “Comments” section, a free text box for respondents to report their perception of the effectiveness of the coaching they were engaged in.

Reliability estimates were calculated for all of the scales. The Cronbach’s alphas were as follows: .90 for the four items in “average effectiveness”; .87 for the Total Working Alliance Inventory of 36 items (and for the three 12-item subscales: .86 for Task, .88 for Bond, .83 for Goal); and .83 for self-efficacy. Not all clients and coaches completed all sections of the questionnaires, although 90% of them did. This is why the *ns* in our summaries are below 156 for some variables.

To summarize, the main dependent variable was “average effectiveness” (the average of four different outcome items rated by the client), and the independent variables were the coaching relationship (as assessed by coach and client independently), the MBTI-preferred “types” of coach and client, the self-efficacy of the client, and the range of coaching techniques. All scale scores were computed as the sum of all item responses. Descriptive statistics, reliability estimates, and inter-correlations are reported in Table 1. Table 2 shows descriptive statistics for client and coach MBTI scores.

T1-2, AQ:5

Results

The model specified in Figure 1 ideally would be tested using one overarching model through the use of structural equation modeling. However, given the somewhat exploratory nature of this research, and the degree of uncertainty surrounding the precise nature of the relationships, we have analyzed these data using several bivariate correlations. We acknowledge that proceeding in this way inflates the probability of the Type II error rate, and thus in order to compensate for this, we will use a Bonferroni-corrected significance level of .007 to assess the results. As shown in Table 1, six out of the seven correlations with perceived effectiveness remained significant at this more stringent alpha level.

We found strong positive correlation, $r = .61$ ($n = 155$, $p < .001$), between the client’s experienced overall working alliance and the effectiveness of the coaching as measured by the client, and similar correlations for the three working-alliance subscales, but no correlation between the coach’s measure of relationship with the client’s estimate of the relationship, $r = .13$ ($n = 130$, $p = .13$). Hypothesis 1, therefore, was confirmed with regard to the client’s rating of the relationship, for all aspects of the standard Working Alliance Inventory (for task, $r = .63$, $n = 155$, $p < .001$; for goal, $r = .57$, $n = 155$, $p < .001$; for bond, $r = .47$, $n = 155$, $p < .001$). This is consistent with the studies by Boyce et al. (2010) and Baron and Morin (2009), although these

Table 1
Descriptive Statistics, Reliability Estimates, and Intercorrelations Among Study Variables

Variable	Items	<i>N</i>	<i>M</i>	<i>SD</i>	Range	<i>r</i>											
						1	2	3	4	5	6	7	8				
1. Client's perceived coaching effectiveness	4	155	23.89	3.15	10–28	.90											
2. Coach's perceived strength of relationship	1	130	6.07	0.92	3–7	.13	—										
3. Total WAI	36	155	214.70	25.19	112–252	.61**	.16	.87									
4. Client's WAI – Task	12	155	71.79	8.41	42–84	.63**	.17	.92**	.86								
5. Client's WAI – Goal	12	155	71.92	10.88	20–84	.57**	.16	.90**	.85**	.88							
6. Client's WAI – Bond	12	155	70.99	8.71	39–84	.47**	.01	.88**	.68**	.63**	.83						
7. Client's self-efficacy	10	151	22.86	3.86	11–30	.25**	.12	.38**	.36**	.39**	.30**	.83					
8. Client's perceived range of coach techniques	11	146	63.94	8.69	22–77	.52**	.12	.67**	.65**	.55**	.61**	.23**	.88				

Note. Cronbach alphas of every multi-item variable are on the diagonal in the intercorrelations section of the table. WAI = Working Alliance Inventory.

** $p < .01$.

studies also found a significant correlation between coach rating of the relationship and client rating of coaching outcome. Our result is similar to what has been found in psychotherapy (see, e.g., Horvath & Symonds, 1991, where client ratings—and not therapist ratings—of the alliance are the best predictor of outcome). It appears from our results that the correlation with client-rated outcome is slightly higher for the task aspects of the relationship (clarity and mutual agreement on the tasks, strength of collaboration, etc.). Computation of Fisher's r -to- z transformation showed that the relationship between the task aspects of the relationship was significantly stronger compared with the bond aspects ($Z = 2.01, p = .04$), but the differences between task and goal, and between goal and bond, were not significantly different ($Z = 0.82, p = .41$; $Z = 1.19, p = .23$, respectively). The overall correlation with the working alliance scale of $r = .61$ ($n = 154, p < .001$) counts as a strong relationship (Cohen, 1988). However the coach measure of relationship strength does not correlate with self-efficacy ($r = .13, n = 126, p = .17$).

Hypothesis 2 was not borne out in any way. First, we used t tests to evaluate differences in the reported outcome of coaching for each of the four MBTI client personality dichotomies, the extraversion-introversion (E/I) dichotomy ($n = 77, 65$), the sensing-intuiting (S/N) dichotomy ($n = 64, 78$), the thinking-feeling (T/F) dichotomy ($n = 86, 53$), and the judging-perceiving (J/P) dichotomy ($n = 82, 57$; Myers et al., 1998) and found none. We then conducted t tests to evaluate differences in the reported outcome of coaching for the coaches' MBTI dichotomies, E/I ($n = 70, 76$), S/N ($n = 16, 130$), T/F ($n = 24, 122$), J/P ($n = 41, 105$) and found no significant differences at $p = .05$. The T/F comparison was verging on significance, $t(143) = 2.83, p = .08$, and showed a more favorable client-rated outcome for coaches who had an F ($M = 24.41, SD = 2.41$) as opposed to T ($M = 22.57, SD = 4.65$) preference. Finally, we used correlation to look for evidence of differences in client-rated outcome depending on degrees of separation between coach and client types on the MBTI (as a measure of personality mismatch) and found none. We also used a t test to investigate the impact of matching and mismatching the coach and client MBTI "temperaments" (Myers et al., 1998; $n = 104$) and found no difference. This discrepancy with the results found by Scoular and Linley (2006) may be due to differences in the design of the two studies: Our data comes from longer term coaching relationships, whereas the data from Scoular and Linley (2006) came from one-off 30-min sessions, where it is possible that the impact of matching might be more significant as they are essentially first meetings. In addition, in Scoular and Linley's sample of just 14 coaches, the relatively high numbers of N types compared with

AQ: 6

Table 2
Client and Coach MBTI Preferences and Profiles

	Client		Coach	
	Count	Percentage	Count	Percentage
ENFJ	6	3.8	18	11.5
ENFP	11	7.1	25	16.0
ENTJ	15	9.6		
ENTP	9	5.8	12	7.7
ESFJ	10	6.4	10	6.4
ESFP	5	3.2		
ESTJ	15	9.6	6	3.8
ESTP	5	3.2		
INFJ	5	3.2		
INFP	7	4.5	69	44.2
INTJ	15	9.6	6	3.8
INTP	9	5.8		
ISFJ	5	3.2		
ISFP	5	3.2		
ISTJ	12	7.7		
ISTP	7	4.5		
Extraversion versus introversion (E vs. I)				
E	77	49.4	70	44.9
I	65	41.7	76	48.7
Sensing versus intuiting (S vs. N)				
S	64	41.0	16	10.3
N	78	50.0	130	83.3
Thinking versus feeling (T vs. F)				
T	86	55.1	24	15.4
F	53	34.0	122	78.2
Judging versus perceiving (J vs. P)				
J	82	52.6	41	26.3
P	57	36.5	105	67.3

Note. MBTI = Myers-Briggs Type Inventory.

S types may play a role in confounding the results (Anne Scoular, personal communication, January 27, 2010).

Although these MBTI findings are significant for our sample, the combination of a lack of robustness in using MBTI as a psychometric measure of personality, and the required sample size to give appropriate statistical power for the likely effect sizes, makes us recognize the need for further work to be carried out in order for results to be conclusive. We are already working to increase the sample size in our next research program.

Hypothesis 3 was supported by a significant correlation of $r = .61$ ($n = 150, p = .002$) between the clients' self-efficacy and the client outcome measure. Again, this confirms well-established results in a related field: a significant correlation between self-efficacy and perceived outcome in self-regulated learning (see, e.g., Schunk, 1990).

Hypothesis 4 was also supported in that the range of techniques that clients perceive they get in their coaching affects their perceived outcome. The Pearson's correlation value of $r = .52$ ($n = 145, p < .001$) between the scale representing a range of coach techniques and working alliance was moderate, positive, and significant.

Hypothesis 5 states that the strength of the coaching relationship mediates the influence of (a) personality differences, (b) client self-efficacy, and (c) coach techniques on coaching outcomes.

AQ: 7

Personality differences were shown not to predict outcome, so hypothesis 5(a) is not relevant and cannot be supported. We used mediation analysis (Baron & Kenny, 1986) to regress self-efficacy and working alliance on client-rated coaching outcome. Working alliance, $\beta = .60$, $t(146) = 8.51$, $p < .01$, was significantly related to coaching outcome, and self-efficacy became nonsignificant, $\beta = .02$, $t(146) = .033$, $p = .74$. The results of Sobel's test showed that the parameter estimate for the relationship between self-efficacy and client-rated coaching outcome became nonsignificant in the mediated condition, $Z = 4.46$, $p < .01$, indicating that the working alliance fully mediated the relationship between self-efficacy and coaching outcome (Baron & Kenny, 1986), providing support for Hypothesis 5b (see Table 3).

T3,AQ:8

Similarly, when the perceived range of coach techniques and working alliance were regressed on client-rated coaching outcome, working alliance, $\beta = .51$, $t(142) = 6.04$, $p < .01$, was significantly related to coaching outcome and the range of coach techniques remained significant but was weaker, $\beta = 0.21$, $t(142) = 2.49$, $p < .05$. The results of the Sobel's test showed that the parameter estimate for the relationship between range of coach techniques and coaching outcome was significantly lower in the mediated condition than in the nonmediated condition, $Z = 7.35$, $p < .01$, indicating that the working alliance partially mediated the relationship between perceived range of coach techniques and coaching outcome, providing partial support for Hypothesis 5c (see Table 4). This result is similar to those of Baron and Morin (2009) and Boyce et al. (2010), who also found that the relationship mediates significantly the other independent variables that correlate with perceived coaching outcome.

T4

Discussion

The research confirms Hypothesis 1(a), Hypothesis 3, Hypothesis 4, and Hypothesis 5 (b-c) on the impact of common factors on coaching outcome. We have found strong indications that the coaching relationship (or, to be more precise, the working alliance), as rated by the client, correlates with client-rated coaching outcome to a considerable degree. As always with statistical correlation, we need to realize that this does not imply causality; that is, coaching outcome may predict a strong relationship just as much as the relationship may predict good coaching outcome (see the double-headed arrow "B" in Figure 1).

Table 3
Mediated Regression Results Testing Hypothesis 5b With Standardized Regression Coefficients

Predictor (IV)	Mediator working alliance	Coaching outcome (DV)
Step 1		
Self-efficacy		.25**
R^2		0.06**
Step 2		
Self-efficacy	.38*	
R^2	.15*	
Step 3		
Self-efficacy		.02
Working alliance		.60**
R^2		.37**
F		43.83**
df		2,146
Sobel test		4.46*

Note. DV = dependent variable; IV = independent variable.
* $p < .05$ (two-tailed). ** $p < .01$ (two-tailed).

Table 4
Mediated Regression Analyses Testing Hypothesis 5c With Standardized Regression Coefficients

Predictor (IV)	Mediator working alliance	Coaching outcome (DV)
Step 1		
Coaching behaviors		.57**
R^2		.33**
Step 2		
Coaching behaviors	.69**	
R^2	.48**	
Step 3		
Coaching behaviors		.21*
Working alliance		.51**
R^2		.46**
F		61.33
df		2,142
Sobel test		7.35**

Note. DV = dependent variable; IV = independent variable.

* $p < .05$ (two-tailed). ** $p < .01$ (two-tailed).

We have also found indications that client self-efficacy and perceived range of coach techniques predict client-rated coaching outcome. The broader the range of coaching techniques used by the coach, the better the perceived outcome. This may indicate that those coaches who have a more pronounced or broader range of expertise and employ more techniques with their clients achieve more favorable outcomes in their clients' eyes.

We have found no evidence for a differential impact of either client personality or client-coach personality matching. This means our results confirm those by Boyce et al. (2010) and Baron and Morin (2009) regarding the coaching relationship, and those by Stewart et al. (2008) in the area of self-efficacy. We have not been able to confirm the findings in the area of personality matching by Scoular and Linley (2006).

All our significant results have been from predictor variables that were rated by the client. The dependent variable, coaching outcome, was also rated by the client. This means we cannot be sure of independence. In other words, what we have found may be entirely ascribable to "common methods" bias (Meade, Watson, & Kroustalis, 2007), that is, more positively or optimistically inclined clients may have registered more optimistic answers on *all* questionnaires, and may thus have mapped out an effect that is entirely unrelated to the executive-coaching assignments that we are studying. This may equally hold for other studies that use client or coach measures as outcome and predictor variables, as noted in the introduction of this article. However, even if some of the results are due to common-methods bias, then they are still important to notice as a "placebo" effect: More hopeful or more optimistic clients will find they achieve better results as clients of coaching.

Whether or not influenced by common methods, we have found convincing indicators of the importance of certain common factors in executive coaching—in particular, the coaching relationship as viewed by the client. On the other hand, there are no indications for a differential impact of matching on the MBTI personality instrument. This would indicate that client-coach matching in terms of personality or demographics might be overstated in the literature. It might be more important to focus on coach selection—in terms of qualifications, accreditation, and supervision records—than on client-coach matching, as Wycherley and Cox (2008) also suggest. The only form of matching between coach and client that this research has found indirect support for is that of engaging the client's first impression of the coaching relationship, from a preliminary meeting or trial session with the coach, after which the client determines whether to proceed with that coach or not.

By contrast with these findings related to personality characteristics, we have found no indications that the importance of the coaching relationship (as judged by the client) has been overstated in the coaching profession: The working-alliance scores by clients in this study predict 25% of total proportion of variance of coaching outcome (see Table 1). Although the quality of the experienced relationship seems to be crucial, the ability to self-motivate (“general self-efficacy”) also seems to be significant (this amounts to around 4% of proportion of variance, according to Table 1).

Although this has been found before in psychotherapy (Horvath & Symonds, 1991), we think it is fascinating that despite the high predictive value of the client estimate of the coaching relationship, the coach estimate of that same relationship neither correlates with the coaching outcomes nor with the strength of the relationship as measured by the client. However much we emphasize the importance of the coaching relationship for effectiveness, we must also emphasize that client and coach have their own unique experience of their relationship and may judge it differently, as the low correlation between their relationship estimates also indicates. This suggests that coaches may not be as “in tune” with their clients as they typically assume. It may also help coaches to appreciate the extent to which their experience in coaching can differ from that of their client and, as a result, to apply a more open-minded curiosity about how the client feels, both about the coaching relationship and about their progress in coaching. It is our experience that this discovery can be a huge insight for a practicing coach at any stage in their career. Finding out the client’s experience, by using the Working Alliance Inventory and encouraging frankness, appears to be a way of resolving that dilemma. However, for a coach to inquire into the client’s estimate of the relationship also has an impact on that relationship: Coach and client are both not only observers but also key players within the coaching relationship. Moreover, reflecting on this relationship, clients might be polite, defensive, avoidant, or otherwise unfocused in their answers to their coach. For convenience, the coaching literature speaks about “the coaching relationship,” which suggests that there are relational aspects that client and coach hold in common. However, the relationship between coach and client only exists in their respective minds (and in the minds of outside observers), causing them to represent “it” in a completely independent way and moreover evaluate “it” completely independently and according to highly personal criteria and expectations. Research into therapeutic relationships shows, time and time again, that there is no one thing called “the helping relationship” as it is perceived and evaluated independently by clients, therapists, and, indeed, observers (see, e.g., Horvath & Marx, 1990; Horvath & Symonds, 1991).

This research shows that although helping conversations including executive coaching are probably generally quite effective, nevertheless, the two partners in conversation, possibly due to their different roles, may be largely out of touch with each other’s experience of the relationship. This is a serious matter, particularly when that same relationship, as experienced by the client, is such a strong predictor of effectiveness. Although an effective coach can probably gauge quite well what the client’s feelings are regarding events *outside* the room, to gauge the client’s feelings regarding what expires *inside* the room is an altogether different matter. We believe it would be useful to find a way of measuring what client’s perceptions of the coaching relationship are in real time, that is, during the coaching intervention, in such a way that that same relationship is not negatively affected (as Miller, Duncan, Brown, Sorrell, & Chalk, 2006, have also tried to do for psychotherapy).

Limitations

Although our key findings seem fairly robust, there are certain limitations that lead to our recommendation of further research. They affect practically all coaching research, including ours. The fact remains that, in this emerging profession of executive coaching, researchers have not been able to achieve the “gold standard” of therapy outcome research, namely, randomized controlled trials with qualified professionals and independent outcome criteria (Wampold, 2001). In this particular study, this means that we have assumed general effectiveness of executive coaching, an assumption that was based on preliminary indications in coaching outcome studies and on the more

rigorous outcome studies in psychotherapy. Second, we have not been able to suggest objective criteria for outcome, such as the assessment by independent outsiders on a well-validated instrument, so that we run the risk of “common-methods” or “same-source” bias on some of our variables. Ely, Boyce, Zaccaro, Hernez-Broome, and Whyman (2010), in their overview study of coaching outcome research designs, also warn about common-methods bias. Third, a limitation of our design is that every coaching relationship studied was measured only once, and at a random stage of development, in terms of number of sessions. Fourth, the nested design meant that, on average, a little over four coaching relationships involved the same coach, and we had insufficient statistical power to do hierarchical linear modeling. This affects, in particular, the variables provided by the coaches: the coaches’ perceived strengths of the relationship and the coaches’ MBTI scores, and hence the conclusions on MBTI matching. Fifth, we employed different scales for the assessment of the working alliance for coaches and their clients. Even if this is the first larger scale study with reports from both client and coach, the sample is still limited and not all measures (particularly those regarding the coaching relationship) have been the same between client and coach.

In summary, we believe there is certainly more research needed into coaching outcome, in particular in the area of the personality of coach and client, and personality matching. As some of our calculations have shown, we need greater statistical power on the impact of personality and relationship, so that we can look more closely into key aspects such as the MBTI preferences of coaches, and relational aspects such as “agreement on task,” “agreement on goal,” and “bond” as seen by both clients and coaches. We are already well underway with a new project that will have well over 1,500 participants and that will use the same relationship measure for client and coach.

Conclusions

This is one of the first studies to systematically explore and compare the contribution of various factors—the so-called common factors—that are deemed to predict coaching effectiveness. The study has found evidence for the central importance of the quality of the working relationship (the “working alliance”) as seen from the perspective of the client, for the importance of the full range of coach techniques as seen from the perspective of the client, and of general self-efficacy of the client who comes to the coaching relationship. In addition, it suggests that personality factors and personality matching are likely to play a lesser role as a predictor of success in executive coaching. These are important findings that may guide the development of the profession and the choices that are made in the recruitment, development, deployment, and matching of executive coaches. Theoretically, these findings are also significant in that they indicate that the concrete skills, actions, or personalities of individuals may be less important to outcome than what these individuals create “in between” themselves, that is, the strength of their working alliances. This means we have to define and model the coaching *relationship* alongside coaching approaches, skills, and techniques, as has been done so extensively up to now.

References

- Anderson, S. L., & Betz, N. E. (2001). Sources of social self-efficacy expectations: Their measurement and relation to career development. *Journal of Vocational Behavior*, *58*, 98–117. doi:10.1006/jvbe.2000.1753
- Baron, L., & Morin, L. (2009). The coach–coachee relationship in executive coaching: A field study. *Human Resource Development Quarterly*, *20*, 85–106. doi:10.1002/hrdq.20009
- Baron, L., & Morin, L. (2012). The working alliance in executive coaching: Its impact on outcomes and how coaches can influence it. In E. de Haan & C. Sills (Eds.), *Coaching relationships* (pp. 213–226). Farringdon, UK: Libri.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182. doi:10.1037/0022-3514.51.6.1173
- Bell, S. E. (2006). Myers-Briggs Type Indicator and executive coaching: Participants’ self-perceptions about the

- effectiveness of the two when used together. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 66, 3980B.
- Berglas, S. (2002). The very real dangers of executive coaching. *Harvard Business Review*, 80, 86–92.
- Bordin, H. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice*, 16, 252–260. doi:10.1037/h0085885
- Bowles, S. V., Cunningham, C. J. L., De La Rosa, G. M., & Picano, J. J. (2007). Coaching leaders in middle and executive management: Goals, performance, buy-in. *Leadership & Organization Development Journal*, 28, 388–408. doi:10.1108/01437730710761715
- Boyce, L. A., Jackson, R. J., & Neal, L. J. (2010). Building successful leadership coaching relationships: Examining impact of matching criteria in a leadership coaching program. *Journal of Management Development*, 29, 914–931. doi:10.1108/02621711011084231
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Cooper, M. (2008). *Essential research findings in counseling and psychotherapy – the facts are friendly*. London, UK: Sage.
- de Haan, E. (2008). *Relational coaching – journeys towards mastering one-to-one learning*. Chichester, UK: Wiley.
- de Haan, E., & Burger, Y. (2005). *Coaching with colleagues: An action guide for one-to-one learning*. Chichester, UK: Wiley.
- de Haan, E., Culpin, V., & Curd, J. (2011). Executive coaching in practice: What determines helpfulness for clients of coaching? *Personnel Review*, 40, 24–44. doi:10.1108/00483481111095500
- Digman, J. M. (1990). Personality structure: Emergence of the Five Factor Model. *Annual Review of Psychology*, 41, 417–440. doi:10.1146/annurev.ps.41.020190.002221
- Ely, K., Boyce, L. A., Zaccaro, S. J., Hernez-Broome, G., & Whyman, W. (2010). Evaluating leadership coaching: A review and integrated framework. *The Leadership Quarterly*, 21, 585–599. doi:10.1016/j.leaqua.2010.06.003
- 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
- Gardner, W. L., & Martinko, M. J. (1996). Using the Myers-Briggs Type Indicator to study managers: A literature review and research agenda. *Journal of Management*, 22, 45–83. doi:10.1177/014920639602200103
- Greenson, R. R. (1965). The working alliance and the transference neuroses. *Psychoanalysis Quarterly*, 34, 155–181.
- Grencavage, L. M., & Norcross, J. C. (1990). Where are the commonalities among the therapeutic common factors? *Professional Psychology: Research and Practice*, 21, 372–378. doi:10.1037/0735-7028.21.5.372
- Heron, J. (1975). *Helping the client*. London, UK: Sage.
- Horvath, A. O., & Greenberg, L. (1986). The development of the Working Alliance Inventory: A research handbook. In L. Greenberg and W. Pinsoff (Eds.), *Psychotherapeutic processes: A research handbook* (pp. 529–556). New York, NY: Guilford Press.
- Horvath, A. O., & Marx, R. W. (1990). The development and decay of the working alliance during time-limited counseling. *Canadian Journal of Counselling*, 24, 240–260.
- Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, 38, 139–149. doi:10.1037/0022-0167.38.2.139
- Jarvis, J. (2004). *Coaching and buying coaching services*. London, UK: CIPD.
- McKenna, D. D., & Davis, S. L. (2009). Hidden in plain sight: The active ingredients of executive coaching. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 2, 244–260. doi:10.1111/j.1754-9434.2009.01143.x
- Meade, A. W., Watson, A. M., & Kroustalis, C. M. (2007). Assessing common methods bias in organizational research. In *Proceedings of the 22nd Annual Meeting of the Society for Industrial and Organizational Psychology*.
- Miller, S. D., Duncan, B. L., Brown, J., Sorrell, R., & Chalk, M. B. (2006). Using formal client feedback to improve retention and outcome: Making ongoing, real-time assessment feasible. *Journal of Brief Therapy*, 5, 2, 5–22.
- Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI manual*. Palo Alto, CA: Consulting Psychologists Press.
- 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.

- Peltier, B. (2001). *The psychology of executive coaching: Theory and application*. New York, NY: Brunner and Routledge.
- Perkins, R. D. (2009). How executive coaching can change leader behaviour and improve meeting effectiveness: An exploratory study. *Consulting Psychology Journal: Practice and Research*, *61*, 298–318. doi:10.1037/a0017842
- Peterson, D. B. (1993). *Measuring change: A psychometric approach to evaluating individual coaching outcomes*. Paper presented at the annual conference of the Society for Industrial and Organizational Psychology, San Francisco, CA.
- Porter, M. E. (2008). The five competitive forces that shape strategy. *Harvard Business Review*, *86*, 78–93.
- Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy: "At last the Dodo said, 'Everybody has won and all must have prizes.'" *American Journal of Orthopsychiatry*, *6*, 412–415. doi:10.1111/j.1939-0025.1936.tb05248.x
- Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist*, *25*, 71–86. doi:10.1207/s15326985ep2501_6
- Schwarzer, R., Mueller, J., & Greenglass, E. (1999). Assessment of perceived self-efficacy on the internet: Data collection in cyberspace. *Anxiety, Stress & Coping: An International Journal*, *12*, 145–161. doi:10.1080/10615809908248327
- Scoular, A., & Linley, P. A. (2006). Coaching, goal-setting and personality type: What matters? *The Coaching Psychologist*, *2*, 9–11.
- Shadish, W. R., Navarro, A. M., Matt, G. E., & Phillips, G. (2000). The effects of psychological therapies under clinically representative conditions: A meta-analysis. *Psychological Bulletin*, *126*, 512–529. doi:10.1037/0033-2909.126.4.512
- Sherer, M., & Maddux, J. E. (1982). The self-efficacy scale: Construction and validation. *Psychological Reports*, *51*, 663–671. doi:10.2466/pr0.1982.51.2.663
- Sherman, S., & Freas, A. (2004). The wild west of executive coaching. *Harvard Business Review*, *82*, 82–90.
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, *32*, 752–760. doi:10.1037/0003-066X.32.9.752
- 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. doi:10.1111/j.1744-6570.2003.tb00142.x
- Stewart, L. J., Palmer, S., Wilkin, H., & Kerrin, M. (2008). The influence of character: Does personality impact coaching success? *International Journal of Evidence Based Coaching and Mentoring*, *6*, 32–42.
- Stiles, W. B., Barkham, M., Mellor-Clark, J., & Connell, J. (2008). Effectiveness of cognitive-behavioural, person-centered, and psychodynamic therapies in UK primary care routine practice: Replication in a larger sample. *Psychological Medicine*, *38*, 677–688. doi:10.1017/S0033291707001511
- 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
- Thach, E. C. (2002). The impact of executive coaching and 360° feedback on leadership effectiveness. *Leadership & Organization Development Journal*, *23*, 205–214. doi:10.1108/01437730210429070
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods and findings*. Mahwah, NJ: Lawrence Erlbaum.
- Wycherley, I. M., & Cox, E. (2008). Factors in the selection and matching of executive coaches in organizations. *Coaching: An International Journal of Theory, Research, and Practice*, *1*, 39–53.

Received May 20, 2011

Latest revision received December 6, 2012

Accepted December 10, 2012 ■