

EXECUTIVE COACHING: NEW FRAMEWORK FOR EVALUATION

Katerine Osatuke, Boris Yanovsky, and Dee Ramsel

*Veterans Health Administration National Center for Organization Development,
Cincinnati, Ohio*

Evaluation research has struggled to keep up with the popularity of coaching, as measures of its effectiveness are challenging to standardize, particularly when coaching executives. Similar to interpersonally based interventions in other fields such as counseling and psychotherapy, coaching takes the form of a fluid, humanistic process, whereas coaching-evaluation standards strive to be consistent with a standardized, scientifically based method. This study describes our experience in facing these program-evaluation challenges while conducting a randomized, quasi-experimental investigation to explore effects of a developmental coaching intervention provided to senior leaders from different organizations within 1 large integrated health-care system. In the context of these challenges, we propose a conceptually new framework to the field of coaching research based on the assimilation model, an empirically grounded theory that originates within psychotherapy research and describes how people overcome issues they find problematic or challenging, whether in clinical or in broader development and growth contexts. We discuss how this framework—with its associated tool: the Assimilation of Problematic Experiences Scale (APES)—offers working solutions to the common and vexing problems faced by research into executive-coaching outcomes, and how it can specifically inform evaluation-planning strategy within studies of coaching effectiveness.

Keywords: executive coaching, leadership development, public sector, health care

Organizations invest substantial amounts of time and money in executive coaching, which reflects its substantial perceived value. According to a 2009 survey by the Harvard Business School, the duration of a coaching engagement can last anywhere between 7 and 12 months, with median costs to the organization estimated at \$500 per hour (Coutu & Kauffman, 2009). Coaching requires major commitments from multiple stakeholders—for example, time invested by coached executives, time and costs required to compensate for their scheduled absences, and costs paid for coaching. As the frequency of executive coaching, in general, is increasing, it is incumbent upon researchers to more fully develop an improved evaluation model for its impact.

At this time, however, systematic evidence to objectively support the benefits of executive coaching is lacking or inconclusive. Little empirical research exists to evaluate the impact of

This article was published Online First December 29, 2016.

Katerine Osatuke, Boris Yanovsky, and Dee Ramsel, Veterans Health Administration National Center for Organization Development, Cincinnati, Ohio.

Correspondence concerning this article should be addressed to Katerine Osatuke, VHA National Center for Organization Development, 11500 Northlake Drive, Suite 230, Cincinnati, OH 45249. E-mail: Katerine.Osatuke@va.gov

executive coaching, whether on individual outcomes (e.g., personal confidence; on-the-job behavior; perceptions by peers, bosses, and subordinates; leadership skills; career development) or short- and long-term organizational outcomes (e.g., workplace climate in supervised units, turnover and retention, organizational performance and operational costs).

The challenges of studying executive coaching reflect many specific features of this population (executives) and this process (coaching), which work together to create considerable complexity, both for scientific study and for practical application of insights gained from studying coaching. Several disciplines relevant or adjacent to the field of executive coaching each specialize in studying these features, individually as well as in their interaction. Examples of these disciplines include psychotherapy and counseling research, leadership development, adult learning, and others. Methodologies within these disciplines offer models, tools, and techniques already designed and previously validated as successful in addressing complex and interacting features of the coaching process as well as of the leadership and development processes and of executives as a population.

Because the process of executive coaching is fluid and tailored to clients' needs rather than standardized and entirely objective in its delivery and consumption, this type of engagement (i.e., leadership development) is humanistic in nature, versus rigorously scientific. This makes it likely for classic measurement evaluation approaches, often rooted in the scientific method, as described, for instance in Cronbach et al. (1985) and Shadish, Cook, and Leviton (1991), to miss the mark due to misalignment between the humanistic process of intervention and scientific method of evaluation.

Executive coaching is not the only field to experience such a challenge. Practitioners of counseling and psychotherapy as well as organization development (OD) all face similar tensions; see Messer (2004) and Wampold (2007), in addition to Rodgers and Hunter (1996) and Osatuke, Moore, and Dyrenforth (2013), for accounts of this challenge in these respective fields. These tensions are a consequence of using behavior-explanatory models and investigative tools created within lab-based science, yet relying upon intervention strategies and change-facilitation paths that reflect how people grapple with meanings. These latter involve inherently subjective processes basically similar to those engaged by spiritual rites, culturally indigenous interventions, and societally based persuasion mechanisms (cf., Frank & Frank, 1993). Accurately capturing meaning-making processes that take place in a coaching context therefore requires tools of study that afford for a substantial fluidity and subjectivity of the evaluated interpersonally based development. These tools should account for scenarios where priorities may change throughout coaching, similar goals may be supported in different ways across coaches and coachees, and other manifestations of fluid process characteristics. In contrast, traditional measurement approaches assume invariance of experiences, problems, and contexts across subjects and interventions. Therefore, they do not capture the unique changes occurring within coached individuals that are key to evaluating the coaching impact.

We propose that a solution to this challenge lies in a new way of thinking about coaching-outcome evaluation that draws upon methods used in adjacent disciplines and thus leverages the interdisciplinary knowledge (e.g., models and tools) proven to work for addressing similar issues in other fields. Specifically, we suggest an innovative approach—using the assimilation model from psychotherapy research (Stiles, 1999, 2002)—that is focused on tracking developmental level of coping used by clients as they experience, define, and tackle their presenting problems and needs brought to coaching. The assimilation model has conceptual similarity to other well-researched models of individual change (e.g., Prochaska, Wright, & Velicer, 2008; Velicer, Brick, Fava, & Prochaska, 2013), in that personal change is understood as a sequence of developmental stages progressing from less to more adaptive coping, and knowing the current stage helps define how to assist clients' progress to the next stage. However, other individual-change models (e.g., Prochaska's) focus on problems such as smoking, which are cognitively clear to clients and posit narrowly defined behaviors as change targets. In contrast, the assimilation concepts are rooted in the classic theory of psychosocial learning and cognitive development (e.g., Piaget, 1953; see Stiles, 1997, 2015, for an in-depth discussion); therefore the assimilation model has been formulated to apply to any—broad or narrow—problem definitions. In other words, the model was designed to describe the many various kinds of changes that all people go through as they develop and grow; this model has

no prerequisite requirement that the individual should define the problem clearly and operationalize it behaviorally. Instead, the way itself that individuals experience their presenting challenges serves to identify the assimilation stage of their problem. For this reason, we suggest this model is particularly well-suited to complex dilemmas brought to coaching, as it places fewer constricting assumptions on the nature of challenges that executive clients may face. We come back to these points later in the paper.

Researchers in the fields of personality, clinical psychology, and, more recently, OD (e.g., Moore, Osatuke, & Howe, 2014) have used the assimilation model to examine change process in a way that captured its key impacts on individuals and groups. Of note, these key impacts identified within the assimilation framework had been missed by more traditional evaluation models that defined the impact and progress in more narrow and standardized ways (see Moore et al., 2014, for an illustration of the framework and method as applied to evaluating an OD intervention and for additional references). In the assimilation model, change is viewed in the context of an individual's experience and identification of his or her own problems; these are seen as specific to this person's developmental stage with respect to addressing this particular issue. Progress or change (in the model's term, *assimilation*) of presenting problems is thus defined as evolution in the client's perception of a problem, which passes through predictable, empirically derived, developmental stages. For example, the problem moves from being initially experienced as external or imposed from the outside, to the coachee more clearly defining the core conflict and seeing his or her own contribution to maintaining it, to ultimately taking a proactive, improvement-focused stance that allows the client first to achieve partial and then full success in resolving the presenting issue. At the top stages of assimilation, the initial presenting issue is no longer seen as a problem but is instead experienced as an opportunity—hence, *assimilated*.

For executive coaching, any relevant assessment results notwithstanding (e.g., see <http://www.hoganassessments.com/content/assessments> for frequently used instruments), initial problems are typically defined mainly based upon clients' personal experience as leaders and upon the organizational context, which the coached executives typically understand in greater depth than their coaches. Progress with respect to problems discussed in coaching also reflects changes in personal experience and context; further, the experiences and contexts both vary widely from person to person. Thus, coaching participants are typically capable of assessing their progress (i.e., intermediate and ultimate coaching outcomes) with more reliability and validity than their coaches. Evaluating effectiveness of executive-coaching interventions, therefore, presents challenges because of the subjective nature of the outcome construct, or, in other words, difficulties of objectively evaluating progress in subjectively experienced problems.

We believe the assimilation model, as borrowed from psychotherapy and counseling research, offers a solution to these challenges currently faced by executive coaching researchers. The model approaches the dilemma (of objectively evaluating the change in clients' subjective experience) through focusing on the clients' coping stance vis-à-vis specific experiences that are difficult or problematic for them. This coping stance is called *an assimilation stage* of the problem. The model has an associated tool—the Assimilation of Problematic Experiences Scale (APES; Stiles et al., 1991)—used for evaluating clients' assimilation stages of problems (higher is better) and for tracking progress (from lower to higher stages) through psychosocially based interventions such as therapy, counseling, integration of individuals into a different culture, organizational-development programs, and, we suggest, also executive coaching. We propose this approach as useful for studying executive coaching because it provides a new, empirically grounded and validated strategy for defining and tracking those developments in coaching that, based upon extensive research in adjacent fields, underlie successful outcomes.

Within this paper we will focus on three main goals: (a) introducing some of the challenges faced by researchers in evaluating the effectiveness of executive-coaching interventions, while drawing parallels to similar challenges in interdisciplinary research where applicable; (b) sharing our experience in conducting a coaching-evaluation study in the context of these challenges; and (c) proposing a new model for conducting a study of executive-coaching

evaluation, based on the tenets of the assimilation model, in order to address the current gaps in the field and literature.

Current Challenges

Evaluating the effectiveness of coaching interventions has proven to be an elusive task because of the presence of some critical features within the coaching interventions themselves. That is, the same features that define a coaching intervention also preclude us from measuring the intervention's effectiveness.

Challenge 1: Defining the Focus of Coaching Intervention

One valuable feature of executive coaching is that coaches work with executives collaboratively to choose a focus for the engagements. This approach maintains respect for the variation in executive clients' needs and preferences and results in coaching targets being customized for specific coachees, as well as making it possible for targets to be redefined as needed throughout the coaching engagement (Kilburg, 1996; MacKie, 2007). This approach is not unique to executive coaching and is frequently found in such adjacent disciplines as growth-oriented psychotherapy and supportive counseling (e.g., Rogers, 1959), as well as in action research (e.g., Bartunek, Rousseau, Rudolph, & DePalma, 2006), where interventions are flexibly tailored to the clients' needs. However, while in these other fields treatment approaches that express a deficit-conflict view of clients' needs still outnumber collaborative, growth-oriented approaches, the deficit-conflict view and its associated prescriptive stance toward an intervention focus has been recognized as clearly inappropriate for coaching executive leaders (see Kauffman & Scoular, 2004, for a discussion). In essence, acknowledging the need in an unequivocally and comprehensively collaborative approach toward defining the content of coaching for executive clients creates a challenge with "the what" of executive coaching. The outcomes for coaching targets must be defined in ways that allow substantial customization to clients' needs. The developmental areas, skills, or competencies on which the intervention should focus are thus identified in ways that are flexible and open to change, rather than in standard and permanent ways.

Challenge 2: Defining How to Deliver Intervention

In addition to "the what" challenge, where the focus of an intervention may vary across clients and across time, the process by which a coach conducts the intervention also depends on context. Presented with different contextual cues, a coach will approach coaching sessions uniquely across clients (and even within clients across sessions) even if the goals of the sessions are similar. In contrast to "the what" challenge, this challenge is more akin to "the how," or the means of achieving results in a coaching intervention. Contextual features such as the client's level of experience, setting, baseline levels of a skill or competency, coach's own areas of best expertise, organizational (or social, personal, political) climate, and other factors influence how a coach approaches a similar developmental goal across different clients. These same considerations, and the resulting challenges for process-outcome research, also apply to interpersonally based intervention fields adjacent to coaching—for example, to traditional psychotherapy (Stiles, Honos-Webb, & Surko, 1998) and OD (Golembiewski, Billingsley, & Yeager, 1976; Coughlan, Suri, & Canales, 2007)—where contextual influences are important determinants of the process.

Challenge 3: Choosing the Measures

Yet another challenge in the study of executive-coaching evaluation reflects (mis)alignment of measures with the constructs being measured. Unlike the first two challenges, this one reflects considerations of valid measurement and evaluation more so than the demands of the coaching process itself. Evaluation studies tend to follow a format where measures are constructed to match the intervention. Since an intervention is expected to have some predictable sort of impact on clients,

measures are typically chosen to tap into this type of impact—for examples, see [Bluckert \(2005\)](#); [Thach \(2002\)](#); [Theeboom, Beersma, and van Vianen \(2014\)](#). However, while in traditional program evaluations it is logical, even necessary, to choose the measures and create an evaluation plan before your study begins ([Cronbach et al., 1985](#); [Fink & Kosecoff, 1978](#); [Rossi, Lipsey, & Freeman, 2003](#)), this is often unrealistic in executive-coaching evaluations because the goals (and processes) of coaching sessions are context-dependent and subject to shifting (cf., [Kilburg, 2000](#)). It should not, however, be assumed that coaching has no tangible impact or that the benefits of coaching do not extend further into the organization, beyond the immediate coachee. Two recent meta-analyses have demonstrated that coaching has a positive impact on individual outcomes, such as performance, job attitudes, and increased self-efficacy ([Jones, Woods, & Guillaume, 2016](#); [Theeboom, Beersma, & van Vianen, 2014](#)). The challenge here instead lies in identifying the specific areas that should be expected to change, given the individual, organizational, and other contextual considerations of the coaching intervention.

In other words, the first two challenges have a consequence of limiting the ability to find appropriate measures for capturing the right outcomes through time. In order to be useful for coaching-outcome research (and not just for assessing progress of particular isolated clients), the collaboratively defined, contextually customized outcomes that are a must for success of executive coaching also need to be tracked in ways that allow comparing progress across various areas of focus, which differ across coachees and across organizational contexts.

Further adding to the measurement challenge, individuals have different expected trajectories of progress, reflecting the variation both in the range of demands they face (e.g., [Cavanagh & Grant, 2006](#); [Laske, 2007](#)) and in the range of personal and behavioral issues they experience (e.g., [Berman & Bradt, 2006](#); [Grant, 2007](#)). It is therefore unrealistic to expect to see progress on the same measure at the same point in time, even for coaching clients working on similar goals appropriately reflected by the measure. The measurement challenge thus follows from the previous two challenges (“the what” and “the how”) described earlier, as they establish the conditions for the existence of these measurement difficulties. The measurement challenge, however, creates concerns that particularly affect coaching research, more so than practice. Also, unlike the first two challenges that stem from inherent and unalterable aspects of executive coaching, the measurement challenge has solutions that, we suggest, can be found within the field of research and measurement—and specifically, solutions that may be offered by the assimilation model and measure that we describe later in this paper.

Current Challenges as Experienced in a Field Study

We now turn to describing our experience with conducting a study of coaching evaluation, presenting the barriers we faced and the lessons we learned. Specifically, we will provide examples of the challenges described above as they appeared in our own field study.

The field study described here examined effectiveness of leadership coaching by comparing several types of outcomes across three groups of executive leaders in the Department of Veterans Affairs (VA). VA is the second-largest federal employer in the United States of America and one of the world's largest providers of integrated health-care services. This system includes a wide variety in organizational complexity and geography and also substantial diversity in demographic and professional characteristics of executive participants. VA routinely uses executive-coaching services provided by doctoral-level clinical or counseling psychologists with postdoctoral training in organizational development who are employees of an internal National Center for OD (NCOD) within VA. The coaching services provided by NCOD are based on a process-consultation model where clients choose their priorities and direction while coaches collaboratively support a self-reflective process in clients and provide ongoing feedback grounded in data, observation, and general knowledge of interpersonal aspects and dynamics within an organizational context ([Reddy, 1994](#); [Schein, 1992, 1999](#)). The coaching contract signed in the first session mentions that the focus will be on organizationally relevant challenges, albeit understood in the context of the coachee's personal experience. This expectation protects the value of coaching from the standpoint of management (the paying party) and sets a clear boundary between the coaching and therapy. The

contract also spells out the expectation of confidentiality, unless the coach perceives imminent potential of grave personal or organizational harm about to be done by the coachee. Consistent with the client-centered recognition of a possible shift in the core issue targeted for coaching, the contract allows for more specific goals to be articulated after a more thorough assessment, as opposed to in the first session.

The purpose of the coaching evaluation was to establish objective, systematic evidence of impact, using newly collected measures as well as relevant data that was already available, and share the conclusions with prospective coaching participants and their organizations as a business case for documenting the value of this type of support to VA leaders. The study was conducted using a quasi-experimental design where we randomly assigned our participants to coaching conditions but, when some of them insisted on being transferred (i.e., to the active coaching condition), we accommodated their request. To maintain scientific integrity of the study, these transferred participants were excluded from further investigation. Participants were randomly selected into one of three conditions; coaching, treatment as usual (TAU), or comparison conditions. Participants in both the coaching and TAU conditions took part in a Health Care Leadership Development Program (HCLDP) that is offered by the Veterans Health Administration (VHA). The purpose of the HCLDP is to prepare high-potential VHA leaders for careers in the executive ranks of the department. This leadership program lasts approximately one year and includes diverse modes of training, ranging from self-guided activities, to teleconference trainings, to three separate 1-week onsite sessions. This program is designed to be comprehensive and intensive as it provides the participants with approximately a hundred accredited hours of formal face-to-face training in addition to an equal amount of preparation and participation outside of the sessions. Accreditation from the program is approved by the American College of Healthcare Executives, the Accreditation Council for Continuing Medical Education, and the American Nurses Credentialing Center. In addition to their standard leadership-development curriculum, those in the coaching condition received ten 1-hour coaching sessions. Participants in the TAU condition only received two 360-degree feedback sessions in addition to the HCLDP program curriculum. Individuals in the comparison condition (i.e., those who had applied to the HCLDP program but were not accepted) did not receive any training but were provided with two coaching sessions around their 360-degree feedback results and were sent e-mails with links to web-based surveys during the study. The final sample included 56 coaching, 46 TAU, and 14 comparison condition participants, yielding a total number of 116 subjects. This sample was smaller than we hoped for, as 184 participants were initially selected (81 coaching, 77 TAU, and 26 comparison), yet the final sample was sufficient to yield some usable data. The following sections describe the common challenges in coaching evaluation, how they presented themselves in our study, and the barriers they created for our evaluation efforts.

Challenge 1: Defining the Focus of Coaching Intervention

Coaching interventions, from goals to the processes, are dynamic and this was no exception in our study. The process-consultation nature of our approach to coaching enables the client to collaboratively choose areas of development with their coach. This feature of coaching carries with it the benefit of empowering the client to take charge of his or her own development. But this type of approach undermines an evaluation design by compromising study control. Furthermore, when the focus of the coaching intervention changes from client to client, the method of coaching (i.e., the treatment itself) may also change to realign with the goals. This results in suboptimal program-evaluation conditions. Whereas in an ideal setting the treatment is uniformly applied to all subjects (i.e., the gold standard in randomized clinical trials and a classic recommendation for scientifically based evaluation research—Cronbach et al., 1985; Rossi, Lipsey, & Freeman, 2003; Sechrest & Figueredo 1993), this does not truly work with psychosocially based interventions, to the extent that they are collaborative rather than prescriptive. For a detailed discussion of these issues for psychotherapy, see Wampold (2007); Osatuke, Moore, and Dyrenforth (2013) examined these issues in the context of an OD program.

This lack of consistency in the purpose of coaching was evident in the qualitative comments obtained within our study where we asked the clients to list the goals of the coaching engagement.

The comments revealed a variety of goals, including such things as time management and public speaking. As an example, one client stated a goal to “increase Psychological Safety when dealing with people and avoid being too ‘assertive,’” whereas another client wanted to “develop expertise in Healthcare Management.” The disparity in the nature of these goals had implications for our coaches in choosing how to conduct the coaching intervention, and on us as researchers in choosing the outcome(s) of the intervention.

Challenge 2: Defining How to Deliver Intervention

To continue the previous example, the disparity in the nature of clients’ goals influences how coaches conduct the interventions. Even if the focus of the coaching engagement is held constant, coaches may—and should—choose different approaches to facilitate progress on the client’s priorities, reflecting the different context of the engagement (e.g., different organizational roles and job positions, which characterized our sample of executive coachees). Furthermore, different coaches may have nonidentical approaches to coaching (e.g., tools, methods, and processes), or even a different balance between the methods and processes used (e.g., relative proportions of probing and questioning vs. supportive listening, which may uniquely characterize different coaches). Also, with two clients both focused on interpersonal skills as their main priority in coaching, one client might be ready to move past the self-reflecting phase, while the other is not yet ready, even with the same coach using similar coaching methods. With the two clients reaching coaching milestones at different times, the coach must then adjust his or her process to the needs of the client—a manifestation of responsiveness that is required for successful coaching interventions but creates noise for process-outcome evaluation purposes. Having two separate coaches in this same scenario will add more noise or process inconsistency.

Our study was susceptible to all these pitfalls, as we had variety in types of goals identified by coaching clients (including a focus on interpersonal effectiveness [IE], technical skills, health-care-system expertise, etc.), and we utilized 14 coaches. The coaches within our study were all highly skilled doctoral-level psychologists (PhDs or PsyDs in clinical or counseling psychology, with postdoctoral OD training), all trained in the same intervention model (Reddy, 1994; Schein, 1969) in an attempt to increase consistency in their coaching-delivery methods. We also used the best design possible for a coaching study occurring in a real organizational context (a quasi-experimental design, with random participant selection). Still, the variety in clients’ and coaches’ characteristics and in the resulting aspects of coaching-session process certainly limited our ability to compare process and outcomes across the study conditions.

Challenge 3: Choosing the Measures

Coaching evaluation is a special case of program evaluation, thus special considerations must be taken when addressing the fundamental components of selecting the outcomes to track, measures to capture these outcomes, and analytic methods to summarize the measures.

First, the impact (or outcome) of a coaching intervention may work less directly than one might hope. For example, coaching changes mindsets, which eventually impact outcomes, rather than coaching directly impacting outcomes immediately. Indirect outcome measures such as multisource feedback, workplace climate perceptions, workgroup performance, and productivity may be attractive—and we were attracted to them in our design—but are likely inappropriate, as they might be too far removed from the most relevant changes that happen through coaching. In other words, it may not be fair to expect a coaching intervention to have an impact on these types of outcomes immediately, and it may be more reasonable to look for an effect on more proximal outcomes—such as personal skill building, self-efficacy, motivation, and self-perceptions of coaching effectiveness. Although these are not bottom-line outcomes that organizational leaders may traditionally seek to establish the business case for coaching, there is evidence to believe that developing leaders on these areas can, in turn, impact the bottom line through improved individual performance and organizational outcomes (Locke & Latham, 2002; Stajkovic & Luthans, 1998).

In selecting measures for our study, we balanced the need for capturing the impact of the intervention, organizational relevance, and the dynamic, tailored nature of coaching engagements. With these considerations in mind, we included measures of goal attainment, perceived organizational support, multisource developmental feedback ratings, and readiness for coaching. Our data show that differences between coaching and TAU clients on indirect measures, such as a 360 multisource feedback, were small and not significant. The measure used in this study was a 360 multisource-assessment instrument, routinely available to any interested executives within VA and offered for developmental (as opposed to performance-evaluative) purposes. The effects of coaching condition on multisource ratings of IE—the part of the 360 assessment relevant to our research focus—were examined. The differences between coaching and noncoaching TAU and comparison conditions were all small and nonsignificant. Although all mean differences were in the expected direction for the coaching to the TAU comparisons, with coaching clients receiving better multisource feedback ratings, the estimates of effect sizes were small and the differences were not significant, suggesting that these differences may possibly be due to chance. Results between the coaching and comparison conditions were mixed, with staff providing more favorable ratings and peers providing more critical ratings of the coaching clients. Effect sizes for the differences were small to medium (Cohen, 1988), ranging from $d = .03$ to $d = .30$.

Our study also included more proximal measures in the form of coaching-client self-perceptions of goal attainment and organizational support. Coaching and noncoaching participants were compared on their goal attainment self-perception ratings at Time 3 of the study, 1 year after the onset of coaching. Additionally, coaching and noncoaching participants were compared on their perceived organizational support at Time 3 of the study. Descriptive statistics indicated higher means for coaching participants on some items of goal attainment and on all but one item on perceived organizational support, compared with TAU or comparison condition participants. Although descriptive in nature and not necessarily representative of the population, coaching subjects reported greatest mean differences on “I have strengthened my leadership capability” and “I have a significantly better understanding of my strengths and challenges” with respect to goal attainment. These differences were not statistically significant but yielded modest effect sizes, ranging from $d = .19$ to $d = .66$, according to Cohen’s (1988) guidelines. Similar to ratings of self-assigned goals, coaching subjects also reported higher means on perceived organizational support on all items when compared with comparison condition and on all but one item (“This organization provides frequent feedback to people about their performance”) when compared with TAU condition. The difference between coaching and TAU subjects on the item “This organization believes a person’s development is a joint responsibility of the individual and the organization” was statistically significant at the $\alpha = .05$ level, $t(62) = 2.32$, $p < .05$. Effect sizes for the group differences ranged between $d = .12$ and $d = .48$, indicating small to medium effects (Cohen, 1988). To summarize, in our study, the clearest findings resulted from the proximal outcome measures used, while findings from more distal outcome measures were mixed and inconclusive.

Second, besides being realistic in projecting the expected impact and selecting its optimal measures, our experience with conducting this study taught us that the process of change itself must be taken into account when administering measures, so that expectations of change can be aligned with the actual, empirically grounded change trajectory in individuals. The process of change takes on many forms depending on the individual experiencing the change. For example, what specifically is involved in reaching the aspired level of interpersonal skill (a frequent self-identified goal within our sample) varies widely across coachees. With the different scope of the targeted accomplishment, it may be unrealistic to expect change from individuals within a standard (uniform) time period. Further, to continue the example, “improved interpersonal skill” might be a category that lacks sensitivity to particular participants’ targeted goals (e.g., one coachee might work on better listening skills while another works on developing more assertiveness). Different changes likely require different time for completion, both within the same category (e.g., interpersonal skills) and across categories (i.e., it may take less time to improve technical skills than interpersonal skills). In sum, much like one would not expect to measure a return on investment after the first session, care must be taken in tailoring measurements throughout the coaching engagement to realistically reflect the pace of progress for different individuals, contents of focus, and so forth.

Our Recommendations: Toward a New Model of Coaching Evaluation

Any proposal of a new evaluation model for executive coaching must face the reality of balancing the dynamic nature of coaching with the rigorous demands of objective program evaluation. That is, whereas it is desirable for a program-evaluation plan to remain static throughout the study to reduce bias and maximize validity, the coaching process may shift and change to meet the demands of the clients. Altering the coaching process to make it more static for the purpose of program evaluation would change the fundamental nature of coaching, resulting in a process that differs from how coaching is actually conducted in the field. At the same time, altering a program-evaluation approach to make it more dynamic could prove to be a fatal flaw to the methodology of a study by undermining validity and reliability. The model we propose resolves this dilemma by systematically using a stable, empirically grounded taxonomy for summarizing clients' stances toward their concerns discussed in coaching—while allowing for the subjective, context-dependent, and highly fluid nature of the concerns themselves.

In the assimilation model (Stiles, 1999, 2002), people are conceptualized as psychologically made of traces of their life experiences—including past and current beliefs, interpersonal encounters, motives, plans for the future, wishes, goals, commitments, and other elements of lived experience. These experiences may be cognitively processed to a different extent, but they are more than just cognitive representations as all of them contain motivational elements. The experiences are active “voices” within a person, “wanting” to be expressed, for instance, by saying or doing something or by taking a stance on current issues that the person deals with. This tenet means that, according to the model, people have no cognitive, motivational, or affective functions separate from the specific experiences these individuals are “made” of; in other words, there is no central processing capacity that acts upon experiences in order to process, evaluate, and manage them (such capacity is itself a “voice”; it may be central or peripheral and underdeveloped).

The implication of this view for understanding change is that changes within intervention participants should be tracked by following data one experience at a time by dissecting participants' experience horizontally. An example is following how the client's experience of “difficulty with interpersonal skills” changes through coaching. This contrasts with the typical approaches of tracking changes within participants vertically, where change is tracked separately across subjects by areas of functioning, such as motivation, attitudes, skills, or behaviors. The benefit of horizontal dissection is that it preserves the uniqueness of individual experiences and contexts, and the content tracked through process-outcome studies closely and specifically matches what is of key relevance to address in coaching interventions. As a consequence of this strategy, change processes (e.g., what was done in coaching to facilitate progress on interpersonal skills) can be concretely and specifically connected to outcomes (e.g., what changed in this participant's interpersonal skills from before to after coaching). Furthermore, the outcome evaluation then matches the initial presenting problems, which makes evaluation results useful and easy to consume not only for researchers and evaluators but also for intervention participants and coaches.

Perhaps the most innovative and useful answer that the assimilation model provides to current dilemmas in the research on coaching evaluation concerns a strategy of comparing different tracked content across participants. According to the model, individuals' various experiences, or *voices*, are integrated (assimilated) to different extents. The already-assimilated voices become interlinked and form a congruent, dominant core of the person's self. Voices inconsistent with the already-assimilated ones are, precisely for this reason, experienced by the person as problematic or challenging. These voices bring distress because they are not (yet) integrated into the rest of the self. This tenet has three important implications. First, the model explains why certain experiences are subjectively challenging for people: It is because they are inconsistent with a person's core experiences. The distinction between the conscious and unconscious, which is important in psychodynamic approaches, is not relevant in this model (e.g., both conscious and unconscious goals represent particular experiences and are part of the dialogue between voices within the person). Second, the model clarifies the common denominator of helpful processes: It is those processes (whether within psychosocial interventions or generally within personal growth and development) that facilitate assimilation of problematic experiences. Third, and

most relevant for our focus here, the model suggests that all experiences, including those discussed in coaching, can be characterized in terms of their level of assimilation. This is the commonality that allows describing levels of coping across various issues that coachees bring to sessions; it creates a useful shared metric for process and progress in coaching, while not losing sight of the individually unique aspects of intervention targets.

Assimilation levels are captured in the APES (Stiles et al., 1991; Table 1). They extend from denial and avoidance (low end), through experiencing strong negative affect (e.g., discomfort) to the problematic content (middle ground), to labeling the issue and developing a detailed understanding into how it is maintained, which serves as the basis for potential solutions (higher end), and eventually reaching mastery over the problematic content so that it no longer is experienced as difficult but instead is used as a resource available for handling new situations (highest assimilation levels). For example, at this level, a coachee may not only effortlessly use the interpersonal skills acquired in coaching but also generalize them to new areas of life. The description and sequence of APES levels have been derived from extensive case studies, mainly using psychotherapy and counseling data but also data from nonclinical settings, and they have been also supported through

Table 1
Assimilation of Problematic Experiences Scale (APES)

0	Warded off/dissociated. Client seems unaware of the problem; the problematic voice is silent or dissociated. Affect may be minimal, reflecting successful avoidance. Alternatively, the problem appears as somatic symptoms, acting out, or state switches
1	Unwanted thoughts/ active avoidance. Client prefers not to think about the experience. Problematic voices emerge <i>or seemingly important concerns are acknowledged only in response to therapist interventions or external circumstances</i> , and are suppressed or actively avoided, e.g., by changing the topic or failing to follow up on a prior commitment without explaining or acknowledging the change. Affect involves unfocused negative feelings; their connection with the content may be unclear
2	Vague awareness/emergence. Client is aware of the problem but cannot formulate it clearly —can express it but cannot reflect on it. Problematic voice emerges into sustained awareness. Affect includes intense psychological pain— fear, sadness, anger, disgust, or intense discomfort associated with the problematic experience
3	Problem statement/clarification. Content includes a clear statement of a problem—something that can be worked on. Opposing voices are differentiated and can talk about each other; <i>the opposite sides or perspectives upon the key conflict are clearly articulated.</i> Affect is negative but manageable, not panicky
4	Understanding/insight. The problematic experience is formulated and understood in some way. Voices reach an understanding with each other (a meaning bridge). <i>There is insight into a bigger picture, or what connects the opposite sides of the conflict. In other words, clarity is gained upon what causes or maintains the problem.</i> Affect may be mixed, with some unpleasant recognition, e.g., from understanding one's own contribution to the problem, but also some pleasant surprise, e.g., from seeing a new perspective
5	Application/working through. The understanding is used to work on a problem. This is the implementation stage; insights are used to generate solutions. Voices work together to address problems of living. Affective tone is positive, optimistic
6	Resourcefulness/problem solution. The formerly problematic experience has become a resource, used for solving problems. Voices can be used flexibly. Affect is positive, satisfied
7	Integration/mastery. Client automatically generalizes solutions; voices are fully integrated, serving as resources in new situations. Affect is positive or neutral (i.e., this is no longer something to get excited about)

Note. Assimilation is a continuum, with intermediate levels allowed; for example, 2.5 assimilation level is half way between vague awareness/emergence (2.0) and problem statement/clarification (3.0). This table, authored by Stiles, appears in several of his papers. The first version was printed in “Longitudinal Study of Assimilation in Exploratory Psychotherapy,” by W. B. Stiles, L. A. Morrison, S. K. Haw, H. Harper, D. A. Shapiro, and J. Firth-Cozens, 1991, *Psychotherapy: Theory, Research, Practice, Training*, 28, pp. 198, 199. Copyright 1991 by American Psychological Association. Adapted with permission. The descriptions in bold italics were added by the current authors (i.e., not part of the original scale). We also added the bold to the APES levels descriptors in the original scale that apply well to the business context (as opposed to the clinical context).

several quantitatively based studies (see [Stiles, 2002](#), for an overview and references). This available evidence supports that progress in interventions is indeed associated with clients' movement from lower to higher assimilation, consistent with the theoretical tenet of the model that assimilative process is what underlies success in psychosocially based treatments. Importantly, the model and APES levels are formulated to track change in clients' experience of presenting problems, in their broadest sense—as opposed to tracking progress on particular challenges, symptoms, contents, or intervention techniques. This makes the model and its associated measure suitable for a variety of contents, client populations, intervention foci, interventionists' approaches, and so forth.

Although the language of the scale (see [Table 1](#)) may suggest its clinical origin, the concepts defining each APES level come from a broader context. They are grounded in the general psychology of development ([Piaget, 1953](#)) and reflect a client-centered focus ([Rogers, 1959](#))—an approach that has importantly shaped the roots of OD (e.g., [Herzberg, Mausner, & Snyderman, 1959](#); [Maslow, 1973](#)) and its current practice (for examples, see the description of OD at <http://www.odnetwork.org/?page=WhatIsOD>). Of note, client-centered approaches are the least clinically narrow of all psychological theories (see [Rogers, 1959](#), for a detailed argument), as they are strength-based; the focus is on clients' growth potentials and development processes, not on psychological handicaps or clinical symptoms.

Given the client-centered nature of the assimilation model and the broad relevance of the assimilation concepts beyond the clinical applications, the clinically slanted language of the APES can be adjusted for a business context while still maintaining conceptual integrity of the model. For example, while “somatic symptoms, acting out, or state switches” (APES Level 0, [Table 1](#)) will unlikely ever characterize executive clients, we believe the other Level 0 descriptors (see [Table 1](#)) accurately portray what is occasionally seen in coaching. Such applicable Level 0 descriptors include: being unaware of the problem; problematic issues silenced; and a visible lack of emotional concern, reflecting successful avoidance. This is because the key concepts of APES Level 0—unawareness or successful avoidance—do show up in executive coaching. For example, clients occasionally dismiss or minimize the relevance of grave organizational issues, or they give these concerns lip service when asked by the coach and then quickly change the topic. Adding to our own observations, [Moore et al. \(2014\)](#) also offered many examples of how early APES level concepts are manifested in the business context, in the experience of executives running a medical center.

The study by [Moore et al. \(2014\)](#) thus far constitutes the only available source of systematic quantitative evidence for applicability of the model to OD intervention settings. Taken together with evidence derived from psychotherapy and counseling research, we believe this shows excellent promise for applicability of the model and tool to the field of coaching evaluation. Below we summarize the main take-aways that we believe an assimilation-based framework suggests to process-outcome research in coaching effectiveness.

One of the biggest barriers that prevented us from optimally measuring the impact of executive coaching in our own study may have been our overreliance on traditional, indirect outcome measures such as multisource feedback, workplace climate measures, job-performance metrics, and even returns on investment. The impact of executive coaching may not reach beyond the client during the time that measures are administered. Moreover, it could be argued that the direct impact of coaching is not intended to reach beyond the client. Executive coaching is intended for personal and professional development of the individual being coached, and gains in this development can be expressed as shifts in mindsets, ultimately expected to make an impact on the company's bottom line. Measuring the direct impact on the bottom line itself may be unrealistic, because this impact is mediated by changes in mindsets and in stances toward organizational issues; therefore these latter need to be described and evaluated first. To be certain, the impact of coaching on the organizational bottom line is of high relevance and capturing it is important. Our point is that this impact is indirect rather than direct; therefore it should be conceptualized as such in order to be effectively measured.

The assimilation framework provides a working strategy for accomplishing this goal. For example, following [Moore et al. \(2014\)](#), who summarized their evaluation results in terms of increases in APES levels from pre- to postintervention separately for various types of intervention participants (i.e., employees, managers, union representatives), it should be possible to summarize

coaching participants' progress as change in their stance toward the initial issues they brought to coaching. To illustrate how this summary may look, a coaching participant who brought to coaching her difficulty of struggling with delegation may, with the coach's assistance, move from the initial active avoidance of the presenting problem (where she felt that her controlling stance, while negatively perceived by others, is necessary to ensure quality organizational performance) to labeling and clarifying the main conflict (difficulty transitioning from her previous top-performer role to her current executive role where she must ensure performance by supporting the work of others rather than by completing tasks herself). This way of expressing the outcome at the individual participant level also allows quantifying progress for purposes of summarizing it at a group level, such as for a coaching program or for all individuals who worked with a particular coach. Specifically, in the example just given, the outcome can be expressed numerically as movement from APES 1 (active avoidance) to APES 3 (problem statement).

This quantification allows comparing outcomes of individual participants across a range of content areas, different points in time, and organizational contexts. This is possible because, in the assimilation framework, any movement up the APES is considered progress, but more movement is more progress. To illustrate, while a change from APES 1 to APES 3 is an accomplishment, change from APES 1 (active avoidance) to APES 5 (the understanding gained is used to work on a problem) is a bigger accomplishment. Importantly, this measurement strategy does not impose standardization where not appropriate. Different and unique issues that participants bring to coaching are not forced into rigid yet nonspecific categories such as "technical issues," "interpersonal issues," and so forth. Yet the commonality between different participants and their experiences is articulated as a clear theoretical construct; assimilation reflects the extent to which participants are able to actively and adaptively cope with specific problems of relevance to them. Also importantly, the relationship of this construct to participant-level positive outcome is validated in previous research—although, as a limitation, this evidence mainly comes from settings that differ from executive coaching. Specifically, except one quantitative study (Moore et al., 2014) and one qualitative study (Osatuke, Moore, Wernke, Stiles, & Dyrenforth 2007), the bulk of the assimilation research involves psychotherapy and counseling data, plus two studies of adjustment of normal individuals to a different culture; see Stiles (2002) for a review of assimilation research.

Finally, the assimilation framework has roots in the broader developmental context that reflects a nonpathological, growth-focused perspective upon problems and challenges—a perspective that, together with Kauffman and Scoular (2004), we believe to be the only appropriate way of approaching executive coaching. Precisely because of this grounding in the context of individual psychological development in general, the assimilation model has been proven to be sensitive to capturing a broader range of progress than traditional outcome measures with a more narrow focus on problems, symptoms, or behaviors targeted for change. For example, in Moore et al.'s (2014) study, applying the model to qualitative interview data from an OD intervention and expressing pre- to postcomparisons in terms of APES ratings showed progress from low to medium stages of APES, for several specific problems experienced by several different types of stakeholders. When the pre- and postdata from this same intervention were previously assessed on two more narrowly focused measures (Maslach Burnout Inventory [MBI] and Moos Work Environment Scale [WES], both described and referenced in Moore et al., 2014), these measures failed to reflect the quantitative progress, which contradicted the stakeholders' strong endorsement of the intervention as helpful and left a puzzling inconsistency between conclusions derived from quantitative and qualitative portions of the assessment. In Moore et al.'s (2014) analysis, this inconsistency reflected the fact that MBI and WES, while both matched the relevant content of the intervention, operationalized progress as movement from *identifying* to *solving* the problems, whereas the impact of the intervention consisted of change from *avoiding* to *facing* the problems. Thus, the assimilation framework provided a better tool for capturing the organizational change in that intervention, because the change was under the floor of the two more common measures yet within the range of progress reflected on APES. We suggest that this consideration—an ability of measures to capture a broad range of progress and not just the subsegment falling between identifying and resolving problems—is also an asset for

coaching research, as we suspect that in our own study a lack of consistent and conclusive results was at least partly due to this limitation.

With the benefit of hindsight, perhaps the main limitation of our study and the best piece of advice that we wish we could have given ourselves would concern using a framework such as one offered by the assimilation model in planning our strategy for tracking progress throughout this coaching intervention. This also constitutes our main recommendation for future research in executive-coaching effectiveness. In line with this recommendation are our suggestions to consider more proximal rather than distal measures for evaluating coaching effectiveness as well as to customize based on clients' initial presenting needs and their subsequent progress how the measures are administered through time. One possible approach to evaluating executive-coaching research that incorporates our recommendations is represented in Figure 1, where measures are administered at staggered time intervals. In this model, Client B is experiencing a different trajectory than Client A and may not be ready to complete the first assessment until prior to session three. In this scenario, the first two coaching sessions have different goals for both of the clients; therefore the change trajectory should be expected to vary and the measures should remain sensitive to this possibility.

In addition, it may be useful to track process measures for specific sessions throughout the coaching engagements. For example, one appropriate tool extensively validated and used in different psychosocial intervention contexts but not in coaching is the Session Evaluation Questionnaire (Stiles, Gordon, & Lani, 2002).

With respect to the outcome content, any measure chosen by a researcher for evaluating the coaching impact should consider the client's own perception of progress with respect to the initial presenting difficulties. Following the tenets of the assimilation model, where changes in individuals occur as a function of evolution in the individual's mindset about and experience of the initial challenge, outcome measures should be chosen to account for this process.

Conclusion

Objectively measuring change in the inherently subjective client experience that is located in the context of human systems and facilitated by an interpersonally based coaching process is not an easy task. One solution suggested here offers the benefits of an already-existing model (the assimilation model), with substantial empirical evidence accumulated for its theoretical tenets and with an associated measure (APES) that is applicable to a variety of intervention foci and client characteristics. Drawing upon this framework offers an opportunity to bring in interdisciplinary knowledge—theoretical, empirical, and psychometric—in order to meet the challenges of coaching-outcome research that we experienced in our own study and that we reviewed in this paper.

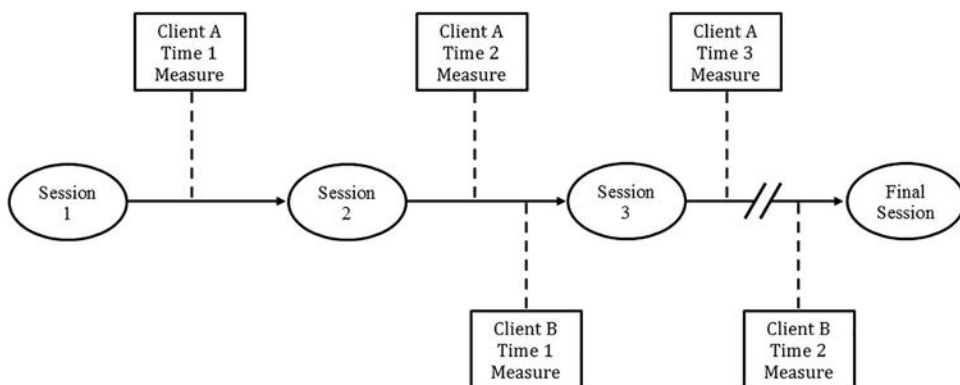


Figure 1. Timeline of coaching model with outcome measures administered at staggered, individually tailored time intervals.

References

- Bartunek, J. M., Rousseau, D. M., Rudolph, J. W., & DePalma, J. A. (2006). On the receiving end: Sensemaking, emotion, and assessments of an organizational change initiated by others. *Journal of Applied Behavioral Science*, 42, 182–206. <http://dx.doi.org/10.1177/0021886305285455>
- Berman, W. H., & Bradt, G. (2006). Executive coaching and consulting: “Different strokes for different folks.” *Professional Psychology: Research and Practice*, 37, 244–253. <http://dx.doi.org/10.1037/0735-7028.37.3.244>
- Bluckert, P. (2005). The foundations of a psychological approach to executive coaching. *Industrial and Commercial Training*, 37, 171–178. <http://dx.doi.org/10.1108/00197850510602060>
- Cavanagh, M., & Grant, A. M. (2006). Coaching psychology and the scientist-practitioner model. *The modern scientist-practitioner: A guide to practice in psychology* (pp. 146–157). New York, NY: Routledge.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Coughlan, P., Suri, J. F., & Canales, K. (2007). Prototypes as (design) tools for behavioral and organizational change: A design-based approach to help organizations change work behaviors. *The Journal of Applied Behavioral Science*, 43, 122–134. <http://dx.doi.org/10.1177/0021886306297722>
- Coutu, D., & Kauffman, C. (2009). What can coaches do for you? *Harvard Business Review*, 87, 91–97.
- Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., & Weiner, S. S. (1985). *Toward reform of program evaluation*. San Francisco, CA: Jossey-Bass.
- Fink, A., & Kosecoff, J. (1978). *An evaluation primer*. Beverly Hills, CA: Sage.
- Frank, J. D., & Frank, J. B. (1993). *Persuasion and healing: A comparative study of psychotherapy*. Baltimore, MD: John Hopkins University Press.
- Golembiewski, R., Billingsley, K., & Yeager, S. (1976). Measuring change and persistence in human affairs: Types of change generated by OD designs. *The Journal of Applied Behavioral Science*, 12, 133–157. <http://dx.doi.org/10.1177/002188637601200201>
- Grant, A. M. (2007). A languishing-flourishing model of goal striving and mental health for coaching populations. *International Coaching Psychology Review*, 2, 250–264.
- Herzberg, F., Mausner, B., & Snyderman, B. B. (1959). *The motivation to work*. New York, NY: John Wiley.
- Jones, R. J., Woods, S. A., & Guillaume, Y. R. F. (2016). The effectiveness of workplace coaching: A meta-analysis of learning and performance outcomes from coaching. *Journal of Occupational and Organizational Psychology*, 89, 249–277. <http://dx.doi.org/10.1111/joop.12119>
- Kauffman, C., & Scoular, A. (2004). Toward a positive psychology of executive coaching. In P. A. Linley & S. Joseph (Eds.), *Positive psychology in practice* (pp. 287–302). Hoboken, NJ: John Wiley & Sons, Inc. <http://dx.doi.org/10.1002/9780470939338.ch18>
- Kilburg, R. R. (1996). Toward a conceptual understanding and definition of executive coaching. *Consulting Psychology Journal: Practice and Research*, 48, 134–144. <http://dx.doi.org/10.1037/1061-4087.48.2.134>
- Kilburg, R. R. (2000). *Executive coaching: Developing managerial wisdom in a world of chaos*. Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10355-000>
- Laske, O. (2007). Contributions of evidence-based developmental coaching to coaching psychology and practice. *International Coaching Psychology Review*, 2, 202–212.
- 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. <http://dx.doi.org/10.1037/0003-066X.57.9.705>
- MacKie, D. (2007). Evaluating the effectiveness of executive coaching: Where are we now and where do we need to be? *Australian Psychologist*, 42, 310–318. <http://dx.doi.org/10.1080/00050060701648217>
- Maslow, A. H. (1973). *Dominance, self-esteem, self-actualization: Germinal papers of A. H. Maslow*. New York, NY: Brooks/Cole.
- Messer, S. B. (2004). Evidence-based practice: Beyond empirically supported treatments. *Professional Psychology: Research and Practice*, 35, 580–588. <http://dx.doi.org/10.1037/0735-7028.35.6.580>
- Moore, S. C., Osatuke, K., & Howe, S. R. (2014). Assimilation approach to measuring organizational change from pre- to post-intervention. *World Journal of Psychiatry*, 4, 13–29. <http://dx.doi.org/10.5498/wjp.v4.i1.13>
- Osatuke, K., Moore, S. C., & Dyrenforth, S. R. (2013). Civility, respect, and engagement (CREW) in the workplace at the Veterans Health Administration. In M. P. Leiter (Ed.), *Analyzing and theorizing the dynamics of the workplace incivility crisis* (pp. 55–68). Amsterdam, the Netherlands: Springer.
- Osatuke, K., Moore, S. C., Wernke, J., Stiles, W. B., & Dyrenforth, S. R. (2007, April). *Using employee interviews to describe a voice of the organization: An application of the assimilation model*. Paper presented at the annual conference of the North American Society for Psychotherapy Research, O. H. Division, Oxford, Ohio.

- Piaget, J. (1953). *The origins of intelligence in children*. London, UK: Routledge and Kegan Paul.
- Prochaska, J. O., Wright, J. A., & Velicer, W. F. (2008). Evaluating theories of health behavior change: A hierarchy of criteria applied to the transtheoretical model. *Applied Psychology: An International Review*, 57, 561–588. <http://dx.doi.org/10.1111/j.1464-0597.2008.00345.x>
- Reddy, W. B. (1994). *Intervention skills: Process consultation for small groups and teams*. San Diego, CA: Pfeiffer.
- Rodgers, R., & Hunter, J. E. (1996). The methodological war of the “hardheads” versus the “softheads.” *Journal of Applied Behavioral Science*, 32, 189–208. <http://dx.doi.org/10.1177/0021886396322006>
- Rogers, C. R. (1959). A theory of therapy, personality, and interpersonal relationships as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of science: Vol. 3. Formulation of the person and the social context* (pp. 184–256). New York, NY: McGraw-Hill.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2003). *Evaluation: A systematic approach*. Thousand Oaks, CA: Sage.
- Schein, E. H. (1969). *Process consultation: Its role in organization development*. Reading, MA: Addison Wesley.
- Schein, E. (1992). *Organizational culture and leadership*. San Francisco, CA: Jossey-Bass.
- Schein, E. (1999). *Process consultation revisited: Building the helping relationship*. Reading, MA: Addison-Wesley-Longman.
- Sechrest, L., & Figueroa, A. J. (1993). Program evaluation. *Annual Review of Psychology*, 44, 645–674. <http://dx.doi.org/10.1146/annurev.ps.44.020193.003241>
- Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.
- Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, 124, 240–261. <http://dx.doi.org/10.1037/0033-2909.124.2.240>
- Stiles, W. B. (1997). Signs and voices: Joining a conversation in progress. *British Journal of Medical Psychology*, 70, 169–176. <http://dx.doi.org/10.1111/j.2044-8341.1997.tb01896.x>
- Stiles, W. B. (1999). *Signs, voices, meaning bridges, and shared experience: How talking helps*. (Visiting Scholar Series No. 10). Palmerston North, New Zealand: School of Psychology, Massey University.
- Stiles, W. B. (2002). Assimilation of problematic experiences. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 357–365). New York, NY: Oxford University Press.
- Stiles, W. B. (2015). Signs and voices 16 years on. In B. Strauss, J. Barber, & L. G. Castonguay (Eds.), *Visions in psychotherapy research and practice: Reflections from the presidents of the Society for Psychotherapy Research* (pp. 153–159). London, UK: Routledge.
- Stiles, W. B., Gordon, L. E., & Lani, J. A. (2002). Session evaluation and the Session Evaluation Questionnaire. In G. S. Tryon (Ed.), *Counseling based on process research: Applying what we know* (pp. 325–343). Boston, MA: Allyn & Bacon.
- Stiles, W. B., Honos-Webb, L., & Surko, M. (1998). Responsiveness in psychotherapy. *Clinical Psychology: Science and Practice*, 5, 439–458. <http://dx.doi.org/10.1111/j.1468-2850.1998.tb00166.x>
- Stiles, W. B., Morrison, L. A., Haw, S. K., Harper, H., Shapiro, D. A., & Firth-Cozens, J. (1991). Longitudinal study of assimilation in exploratory psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, 28, 195–206. <http://dx.doi.org/10.1037/0033-3204.28.2.195>
- Thach, E. C. (2002). The impact of executive coaching and 360 feedback on leadership effectiveness. *Leadership & Organization Development Journal*, 23, 205–214. <http://dx.doi.org/10.1108/01437730210429070>
- Theeboom, T., Beersma, B., & van Vianen, A. E. (2014). Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context. *The Journal of Positive Psychology*, 9, 1–18. <http://dx.doi.org/10.1080/17439760.2013.837499>
- Velicer, W. F., Brick, L. A., Fava, J. L., & Prochaska, J. O. (2013). Testing 40 predictions from the transtheoretical model again, with confidence. *Multivariate Behavioral Research*, 48, 220–240. <http://dx.doi.org/10.1080/00273171.2012.760439>
- Wampold, B. E. (2007). Psychotherapy: The humanistic (and effective) treatment. *American Psychologist*, 62, 857–873. <http://dx.doi.org/10.1037/0003-066X.62.8.857>

Received June 25, 2015

Latest revision received October 18, 2016

Accepted November 10, 2016 ■