A SYSTEMATIC REVIEW OF QUALITATIVE STUDIES IN WORKPLACE AND EXECUTIVE COACHING: THE EMERGENCE OF A BODY OF RESEARCH
Erik de Haan

CITATION
A SYSTEMATIC REVIEW OF QUALITATIVE STUDIES IN WORKPLACE AND EXECUTIVE COACHING: THE EMERGENCE OF A BODY OF RESEARCH

Erik de Haan
Hult International Business School

There is substantial evidence that qualitative research in executive coaching has come of age in the previous decade. Two large research programs have yielded consistent and quantifiable results, and a range of case studies, field studies, and process research is inspiring newer quantitative-research designs. This study contains the first rigorous, systematic review of this qualitative-research base with preliminary conclusions in terms of what this body of work might be telling us. Comprehensive data gathering and screening categorized 101 publications (peer-reviewed articles, book chapters, and dissertations) containing original qualitative research into workplace and executive coaching. This seemed a sufficiently large number of original publications to analyze and then synthesize in terms of its comprehensive findings. Three research questions were formulated in terms of what the qualitative research may offer over and above standard quantitative outcome research, and they are systematically answered with the help of an interpretative synthesis of the findings in four domains. The qualitative-research body of workplace and executive coaching seems to warrant the following tentative findings. Success criteria are coachee-related, involving the development of trust in, acceptance of, and commitment to coaching and the coachee’s respect for the coaching contract. Another success criterion for both coaches and coachees is the ability for both to achieve agreement on tasks and goals, plus a deep level of shared psychological understanding and new insight.

What’s It Mean? Implications for Consulting Psychology

After some 25 years the qualitative-research body in workplace and executive coaching merits a systematic review. The manifold studies and investigations are beginning to offer a coherent and testable image of what works in coaching, as well as a rich palette of applications in different cultures and industries.

Keywords: executive coaching, systematic review, qualitative research, case studies, action research

Correspondence concerning this article should be addressed to Erik de Haan, Ashridge Centre for Coaching, Hult International Business School, Ashridge House, Berkhamsted, Hertfordshire HP4 1NS, United Kingdom. E-mail: erik.dehaan@ashridge.org.uk
Qualitative research in coaching arguably starts with the coaching session itself. During all coaching sessions, coach and coachee investigate what the coachee puts forward and wants to accomplish. Together they explore what this session is about, what else might be relevant, and how to achieve the coaching objectives. One could even argue that this exploration is *all* they do together; after all, the implementation of findings inside the organization is entirely left to the coachee. In other words, executive coaching can be seen as a form of (collaborative) qualitative research, which takes place in a privileged, confidential environment. The outside world will have no knowledge of it, and it does not explicitly inform other coaching sessions by other coaches or coachees.

A logical next step for coaching researchers, therefore, is to conduct research into their own coaching, which they can communicate, for example, through the preparation of case studies based on reflective notes. This form of qualitative research stays as close as possible to the events in the sessions themselves. Qualitative research does not work in the same way as the contrasting quantitative method, which uses a myriad of “data points” (usually numbers, more broadly discrete, quantifiable information reported retrospectively) to make statistically reliable statements. Qualitative research instead starts from “rich data” (emotions as experienced by coach or coachee in the moment, stories, notes, recordings, and transcripts) and tries to achieve a coherent understanding of the underlying richness.

Quantitative research, in essence, takes sequences of numbers originating from coaching assignments (such as rated coaching outcomes according to the various parties involved) and generalizes to produce statistically reliable statements applicable to all coaching assignments in the data set. For example, a common finding in quantitative studies is that the coaching sessions showed an average coachee-perceived effectiveness of 7.2 on a 10-point scale and that the spread of these outcomes was 1.6 (de Haan, Culpin, & Curd, 2011), in which case we can be reasonably certain that the sessions were fairly helpful in the view of most coachees. Any statement that quantitative researchers can make—and such statements are thin on the ground because of the statistical requirements that they need to meet—immediately raises follow-up questions that cannot be answered. As soon as quantitative research has established a significant degree of effectiveness, questions arise, such as:

- Exactly how is that effectiveness achieved?
- What makes some coaching sessions more effective than others?
- How does this effectiveness manifest itself in the coachee’s organization?
- How can you tell within the organization that coaching has been effective?
- What does this result mean for my own specific coaching conversations?

All of these straightforward questions require very extensive follow-up research. No matter how easy it is to ask such questions, the quantitative researcher is usually unable to answer them.

Qualitative research, on the other hand, starts with extensive descriptions (or recordings) of the events themselves and thus with a detailed picture of what happens within coaching. This material can supply an answer to follow-up questions, albeit a highly personal and provisional answer. The difference between quantitative and qualitative research is substantial: Quantitative research has to reduce an entire coaching assignment to a single number (or row of numbers), whereas qualitative research takes time to listen very carefully to a whole story flowing from, say, a single moment in a session. Qualitative research can provide a detailed, nuanced, coherent, and well-founded answer yet does not know the extent to which this can be generalized. Nevertheless, based on this detailed account and descriptive information, qualitative research can ultimately lead back to quantitative research, provided enough corresponding “data points” can be generated from demonstrably similar populations. In most cases, this would require an unprocurable amount of qualitative data. Nevertheless, qualitative research can still produce informed suggestions for narrow hypotheses that can subsequently be investigated by means of quantitative research, and it can even yield new theoretical variables or verifiable coaching models.

Coaching is still an emerging profession, so there are few systematic overviews of qualitative research (one exception being Athanasopoulou & Dopson, 2018, which however included articles
of both quantitative and qualitative research). In this contribution, I would like to present a thorough overview of all qualitative research conducted to date in the field of executive and workplace coaching, including what this research has demonstrated.

A vast amount of qualitative research has already taken place within coaching. Work has included:

- Case studies based on coachee or coach experiences (see, e.g., Levenson, 2009).
- Process research conducted by studying reports or direct recordings of sessions (see, e.g., Ianiro, Schermuly, & Kauffeld, 2013; and Myers & Bachkirova, 2018).
- Process research via interviews and surveys (see, e.g., Wasylyshyn, 2003).
- Action research (see Reason & Bradbury, 2001) as a deeper exploration of coaching practice and of new actions within the coaching relationship (see de Haan, 2016, for a book brimming with research articles on coaching based on action-research methodology).
- Field research, often a form of participatory research in practice—for example, research via evaluation forms, Q-sort techniques, and one-to-one or panel interviews (see Perkins, 2009).
- Descriptive research into coaching interventions and coaching situations, as reflected in numerous manuals but also more rigorous studies (see, e.g., Graf, 2012).

Often it is the coaches themselves who conduct qualitative research, in the role of researchers studying their own practice or that of other coaches. The coach’s perspective on coaching sessions is therefore overrepresented in the literature, and some studies may suffer from partisanship or bias on the part of the researchers, as they may embrace or highlight their own ideas, concepts, and theoretical influences in the “research material.” With the exception of action research, which is a disciplined form of self-study, results are usually more generalizable and objective when the research is independent, that is, when the researchers themselves are not part of the coaching practice being studied.

A helpful way to map out the different types of qualitative research in coaching is in the form of concentric circles around coaching (see Figure 1). In the first type of qualitative research, action research, the researcher is the same person as the coach or coachee and participates in the sessions studied. The second type, case-study research, looks at the session(s) more retrospectively and from the outside. The third type, field research, is again contemporaneous with the studied coaching relationships, but it involves many practitioners and sessions. The final and fourth type of qualitative research, process research, stands truly outside of the sessions themselves and asks questions about recordings of sessions or data from post hoc interviews. Most quantitative research stands on a par with this fourth type of qualitative research; it quantifies the processes studied or the answers obtained through qualitative research. This is why quantitative research cannot really develop or test a hypothesis about how coaching works within sessions. Somewhere in between the unique and complex reality of the sessions and the universally valid conclusions from sound quantitative research, new perspectives and models of coaching are born and important heuristics are abstracted from direct experience.

Other than the qualitative research reviewed here, there is a large group of empirical studies using “mixed methods,” in which a number of open questions or interviews are combined with an otherwise quantitative research design (see, e.g., de Haan et al., 2011; Grant, 2014; Kombarakaran, Yang, Baker, & Fernandes, 2008; or Bachkirova, Arthur, & Reading, 2015). Like purely qualitative paradigms, such open questions can generate new variables for future quantitative research. However, I have chosen not to review them here because (a) the qualitative questions are usually only used to illuminate, compare, and expand on the quantitative findings that are more central to their argument and (b) it would yield a very large and unwieldy data set for analysis for a single study. In sum, this systematic review sets out to be a timely celebration and rigorous synthesis of purely qualitative findings in workplace and executive coaching.

I approached the extant qualitative coaching literature with the following research questions in mind:
Research Question 1: What variables of relevance (e.g., for effectiveness) are being put forward convincingly by qualitative studies? Qualitative research can be particularly helpful in putting forward new variables for effectiveness (such as trust, see Alvey & Barclay, 2007; or commitment to the coaching contract, see Audet & Couteret, 2012). Qualitative research can study the meaning and dynamics of such variables even if it cannot demonstrate any predictive power of, or significant relationships among, such variables.

Research Question 2: What particular applications of coaching in specific industries suggest measurable differences within the profession? Qualitative studies can limit themselves to particular instances such as coaching for organizational change, coaching as a part of leadership-development programs, coaching in small and medium-sized businesses, or coaching in particular business and national cultures. A good qualitative study needs only 1 to 10 cases, which can be studied in detail, and can therefore be set up in contexts where quantitative researchers would not find sufficient data.

Research Question 3: What has the analysis in qualitative studies yielded in terms of significant, robust, and quantitative results? As argued above, any qualitative-research design can be turned into a robust quantitative analysis provided it can be expanded to enough data points for significance (i.e., around 100 data points if only measuring a handful of variables, and a multiple of this if richer, multidimensional dependencies are studied). This is not to say that quantitative results are preferable, just that some qualitative results can be quantified with the usual limitations of quantitative research.
Method

Data Collection and Screening Process

The contribution of this paper consists of a systematic review of the full qualitative coaching-research literature—from selection and analysis of data to synthesis of outcomes and variables (Hoon, 2013). I started with a PsycINFO search on the search terms coaching and qualitative research in 2018, which yielded 790 sources. This could be combed down to 627 sources by taking out the articles about sports coaching and the articles about coaching as a leadership style by line managers. After checking for original empirical studies with new data and taking out all mixed-methods research where the quantitative study dominates, I was left with 105 sources between 1996 and 2018. I then checked the master’s theses for originality and left most out because they were usually small-scale, while retaining doctoral studies. This led to 77 studies in total. I accepted other languages than English as long as I could read them, so a few German and Dutch sources made it through to this list. By searching for quotations within all sources, as well as the direct citation of the sources through Google Scholar and ResearchGate, and finally by checking Grant’s (2011) annotated bibliography of coaching research, I could extend the amount of relevant original studies to 101. The final set of 101 studies from 1995 to 2019 includes one paper (Winum, 1995) that does not use the modern term coaching at all, yet it is investigating an individual consultation contract that we would now call coaching. The search incorporated dissertations, edited and nonedited books, and 25 professional journals with different rankings stemming from various psychology and business areas to reflect the wide-ranging interest in workplace and executive-coaching studies.

Analysis of Qualitative-Research Literature

As argued in the introduction, I was particularly interested in the following questions:

RQ1: What variables of relevance (e.g., for effectiveness) are suggested by qualitative studies?

RQ2: What particular applications of coaching in specific applications or industries suggest measurable differences within the profession?

RQ3: What quantitative findings are produced based on the coding of qualitative data sets?

To analyze the data in the articles, a 150- to 200-word summary was produced for each of the 101 articles and the articles were divided into the four most commonly used subdivisions of qualitative research in organization studies: action research (19 sources), case studies (26 sources), field studies (19 sources), and process research (37 sources). From there on, a first global synthesis was achieved using the eight-step procedure proposed by Hoon (2013) for organizational case studies, by initially grouping articles together and following themes within the four qualitative-research domains. After all the articles had been subsumed into the four categories, a preliminary answer was obtained on each of the main research questions within that category, which was further checked with the original sources until the most coherent answer was obtained. Table 1 contains the classification of the different types of qualitative research that I used.

Results: An Overview of Research Findings

Qualitative-research methodologies have been summarized in Figure 1 and in Table 1. Of the four different broad methodologies, I found that action research contributes mostly to RQ1 about relevant variables; case studies and field research contribute mostly to RQ2 about specific applications; and process research contributes mostly to RQ3, yielding testable hypotheses and evidence.
Action-research (Lewin, 1946; Reason & Bradbury, 2001) is an extremely rich and creative exploration of coaching in real time, working from session to session. The process asks coaches (or coachees) to formulate a research question and to study (i.e., explore, respond to, refine, or transform) this question in their own practice. Action research allows experimentation with the research question because the coach can, for example, try out a different type of intervention and then see what the result is in his or her own practice. This type of research further allows other parties to be involved, although of course they do need to be informed in advance and give their informed consent. For coachees, taking part in research can be problematic: Coaching is about them and serves their issues, so objectives can be undermined if the coachee is distracted by considering and responding to action-research questions. The emphasis might then be placed too much on the coach at the expense of benefits for the coachee. Nevertheless, action research can yield much that is positive for the coach’s practice: both generalizable and testable observations, as well as a deep study of active ingredients and emotions during coaching, which future coaching relationships can benefit from. Moreover, action research opens up a cycle of research and of influencing the coaching process itself (via the learning process of the action researcher, i.e., the coach), so it can also benefit the coaching itself.

Here are some examples of action research in coaching, collected from final theses published by graduates of the Ashridge program for a master of science in executive coaching: Jane Cox (2012) studied the importance of idiom in coaching and the experience of coaching in a second or third language. She concluded that the experience and results are less dependent on such language barriers than was previously thought and that time lags resulting from working in a foreign language can even benefit the process because they allow more reflection before a response. David Skinner (2012) studied the political influences of various stakeholders in multiparty coaching contracts at the top of large multinational organizations. Rob Watling (2012) studied his own learning process en route to a different approach to coaching. Alison Zarecky (2014) inquired into her own strengths-based coaching intervention to see how it helps with the exploration of identity.

The book Behind Closed Doors (de Haan, 2016) summarized action research by 15 coaches each focusing on a different action-research question related to coaching, such as their own identity as coaches, the influences of self-criticism, self-disclosure, humor, safety and trust, and somatic observations on coaching relationships.

Relevance of Action Research to RQ1, RQ2, and RQ3

Action research helps to demonstrate the complex nature of reality, which is much more complex than the simple categories we use to understand it. It mainly contributes to RQ1, about relevant variables for coaching, namely by deepening our understanding of such variables (rather than providing empirical findings about them). Variables that are shown to be much more complex than normally understood are, for example, power (Skinner, 2012), self-disclosure, and somatic responses (Chapters 6 and 12 in de Haan, 2016).
Case-Study Analysis

Case studies also bring the researcher very close to the live practice of coaching, just like action research; however they are usually undertaken retrospectively on the basis of process notes. Case studies attempt to recount the whole story of a coaching assignment, or a single coaching session, usually in chronological order. Choices made by coach and coachee from moment to moment are subjected to retrospective investigation, and the long-term effects of interventions are monitored.

Here is a synthesized review of original case-study research in coaching to date.

A single case with depth. Older studies such as those by Winum (1995); Diedrich (1996); Kiel, Rimmer, Williams, and Doyle (1996); and Tobias (1996) described cases as illustrations of a particular approach, as did Day (2010) and Kets de Vries (2013). The cases from the 1990s, that is, the early days of executive-coaching research, show how the profession has changed and has become less dependent on psychometric instruments, regular feedback with management and colleagues, and long-term involvement of the executive coach. Orenstein (2002) explored three case descriptions from her own practice as an illustration of the psychological depth of coaching conversations. Mansi (2007) described an individual case in which a psychometric instrument played an important role. Blattner (2005) wrote about a case of executive coaching that lasted for 2 years. Peterson and Millier (2005) detailed their collaboration as coach and coachee in quite a lot of detail, as was also done by de Haan and Nieß (2012). Freedman and Perry (2010) described a critical, intensive, and lengthy case also from their own perspectives as coach and coachee.

Schnell (2005), Fahy (2007), Gorringe (2011), Lawrence (2015), Wilson and Lawton-Smith (2016), and Nanduri (2018) presented case studies showing how coaching can support organizational change and growth. Schnell (2005) reported on a 5-year internal coaching assignment for two “coleaders” in a university institute evolving from a pioneering organization toward vigorous growth and professionalization. Fahy (2007) described a senior team of 12 executives, each coached by himself as part of a cultural shift in a technology company. Lawrence (2015) researched through longitudinal interviewing his own 2-year struggle to further a coaching culture in a small Australian multinational with little experience of coaching or reflective dialogue, involving 15 coaching relationships (nine different coaches). Wilson and Lawton-Smith (2016) analyzed through a case study how “on the spot” coaching (one-off sessions on request with a range of coaches) can help in a rapidly changing business environment. Nanduri (2018) coached six managers and a human-resources business partner over a period of 3 months (five sessions each) during a radical reorganization and conducted a phenomenological analysis after the end of the assignment.

Kauffman and Hodgetts (2016) took the transcript of a single, short coaching session as the starting point for detailed analysis of the conceptual and practical choices made by a coach. By introducing several perspectives on the same case, they explored the benefits for the coach of consciously holding multiple perspectives at the same time.

Comparison of multiple cases. Huggler (2007) interviewed six CEOs recently coached by herself. Levenson (2009) provided detailed information on the positive effect of coaching within a company using 12 of his own successful case studies and analyzed factors contributing to successful outcomes; he concluded that transfer to the workplace is the most important factor. Ben-Hador (2016) conducted a multiple case study of 23 of her own coaching contracts between 2011 and 2013 in eight organizations in Israel, based on 79 interviews with all parties involved. She described her struggle with a hidden agenda often present in the organization: that of assessing and influencing the managers through their coach. Hurd (2009) also described four short cases of his own. Foster and Lendl (1996) illustrated through four case studies how EMDR (eye movement desensitization and reprocessing) can help when coaching executives who are struggling with a traumatic experience such as a serious accident, the need to face high-level job-interview panels after decades in a tenured position in a different sector, or, in the case of a CEO, a humiliating layoff. Similarly, Anderson (2002) showed how rational-emotive behavioral therapy can be used in coaching with help of seven short cases, and Gray, Burls, and Kogan (2014) did the same for positive psychology, with five short cases.
More good case studies are described in coaching handbooks for specialized applications, such as coaching in education (e.g., Van Nieuwerburgh, 2012) or coaching in specific cultures (e.g., Gallo, 2015, regarding coaching in China).

**Relevance of Case-Study Analysis to RQ1, RQ2, and RQ3**

Like action research, case studies help to illustrate complexity, and therefore they deepen our understanding of variables for RQ1. Transfer to the workplace and no interference from the workplace come out as important factors to make coaching more successful (Levenson, 2009) and to overcome hidden agenda’s (Ben-Hador, 2016).

The current extent of case studies in the literature help a lot with RQ2, as the case studies are always in specific industries and geographies. Many of the cases show how coaching can support a wider organizational agenda, such as radical transformation, although the one study that is geared at supporting a culture change toward a more coaching culture seems largely a failure (Lawrence, 2015). Also, there is some support for specific approaches, for example, positive psychology in a wide range of coaching settings (Gray et al., 2014) and EMDR for work-related trauma (Foster & Lendl, 1996).

**Phenomenological Field Research**

Real-time field studies involve surveying coaches and coachees in action, to find out more about the intervention and about their evolving experiences with coaching. Researchers get close to the existing coaching practice and interview participants while they are still in process. Sometimes the researcher plays a dual role, that is, including that of coach (e.g., in Perkins, 2009). There is a degree of overlap with action research; it has been described as “second- or third-person action research” (see Reason & Bradbury, 2001). The big difference is that the action researcher of coaching is at the same time also a participant of the coaching conversations, either a coach or a coachee, doing first-person inquiry into coaching practice.

Here are some examples of field research in coaching: Alvey and Barclay (2007) studied the development of trust in coaching relationships by interviewing 27 senior managers who had received coaching. They ranked their identified factors and showed how trust can build up incrementally over time for many of these coachees. Gray and Goregaokar (2010) studied how coachees in small to medium-sized enterprises (SMEs) choose their coaches. They interviewed them during coaching relationships. Their coaches were also interviewed in forum groups. The authors showed that gender played a major role in the choice of a coach and that sexism was invariably at play when male coachees chose their female coaches. Brauer (2005) used interviews to examine the effects of greater or lesser degrees of voluntariness on the part of coachees. Skinner (2014) looked specifically at coaching for women in senior roles and shows how coaching can support female leaders develop a distinct role identity as distinct from traditional male norms of leadership. Rohmert and Schmid (2003) interviewed coachees several times during their coaching processes.

Peel (2008) compared the coaching in 10 SMEs in Wales and highlighted factors that may differ compared with coaching in larger organizations, such as the dominant role played by senior executives in such companies (where ownership and senior management coincide). During a 3-year systematic field study Gray, Ekinci, and Goregaokar (2011) interviewed up to 46 coachees at the top of SMEs in Surrey, each just after the end of 10 hr of coaching, and then reflected on the interview transcripts with two focus groups of altogether 9 coaches. They discovered that the coaching had been more “personal development” than “performance improvement,” even though the coachees attached more importance to the latter. Gray, Gabriel, and Goregaokar (2015) undertook longitudinal field research during 24 months among 13 managers who had recently become unemployed, starting the interviews after the completion of 10 hr of coaching. All but one of the managers remained unemployed during the period of study, and they generally reported strong negative feelings about their unemployment and very mixed feelings about their coaching. Gray et al. (2015) found a positive relationship between the managers’ attitudes toward the coaching and their ability to learn from their trauma.
Buckle (2012) looked specifically at the use of psychometric instruments by interviewing three coaches and coachees with relevant experience. Yedreshteyn (2009) examined an internal coaching program for 18 coachees by conducting a case study of the entire program. Machin (2010) studied internal coaching programs from the coach’s and coachee’s perspective. He discovered that, even during brief internal coaching, a great deal of (psychological) depth was observed and appreciated by both sides, so it seems crucial for coaches to be prepared for this in internal coaching programs as well. Swart and Harcup (2013) studied organizational change (or collective learning) through individual coaching in two law firms, by interviewing 23 stakeholders, namely coachees, their team members (as observers of the coachees), their coaches, and representatives of management.

In a comprehensive study, Perkins (2009) described a particular form of field research in which he first rated 21 of his own coachees on “team leadership” during meetings, then coached them for as many months as were deemed necessary; finally, he observed and scored them again on their leadership behavior. This type of pre- and postmeasurement clearly contributes to the coaching process itself, and to research in the field of coaching, and makes the study at the same time less convincing as “evidence” for overall effectiveness of coaching if only because of the coach doing the scoring.

There is also ample field research into coachee and coach factors for success in coaching; for example, Wasylyshyn (2003) surveyed 87 of her own coachees and Blackman (2006) surveyed 114 current and recent coachees, to study factors within the coach, coachee, coaching process, and organization that seem important to coachees for successful coaching. Liljenstrand and Nebeke (2008) surveyed 2,231 coaches to detect differences in their practices and differences in the coaches’ backgrounds, namely psychological, clinical, educational, or business-school training or still “other” qualifications.

A thorough way to do field research is by rigorous application of multiple live case studies (Yin, 1994). Audet and Couteret (2012) demonstrated this by comparing six cases of intensive coaching (up to one day a week) by experienced entrepreneurs for the benefit of start-up entrepreneurs. The cases ranged from very successful to unsuccessful, and the authors showed that the main (“necessary”) factors determining success belonged entirely to the coachee: receptiveness to coaching, commitment to the coaching, and respect for the coaching contract. Styhre (2008) conducted another impressive longitudinal field study by (re)interviewing six site managers in the construction industry as they received individual and group coaching from the same coach. Styhre discovered a multi-layered contribution of executive coaching in terms of reflection and self-observation.

Relevance of Field Research to RQ1, RQ2, and RQ3

With respect to RQ1, trust was shown to be a variable that seems important and multilayered (Alvey & Barclay, 2007), and the contribution of reflection in coaching seemed multilayered as well (Styhre, 2008).

Field studies have had a lot to say about RQ2 as well, because case studies take place in specific industries/countries, such as local SME’s management (Peel, 2008; Gray & Goregaokar, 2010; Gray et al. 2011), law firms (Swart & Harcup, 2013), and construction (Styhre, 2008).

Audet and Couteret (2012) posed a convincing argument related to RQ3, when they showed that in their series of cases, client commitment and engagement with the coaching contract seemed to be a necessary condition for success.

Process Research

Process research seeks to map out and investigate aspects of coaching, including perspectives on the dynamics between coach and coachee, interaction patterns, and active ingredients according to various parties involved. This type of research can take many different forms, as we will see in this overview. Common methods of data collection are interviews or recordings of conversations, which are then documented and examined by means of image, audio, or textual analysis. Researchers can break down long texts with the help of “grounded theory” (Corbin & Strauss, 1990) to yield meaningful fragments that can be studied to find themes and categories that appear to be of interest.
to those involved in the coaching process. Once categories have been identified, interview or coaching transcripts can be coded, from which point quantitative analysis can take over if interrater reliability is strong enough (Boyatzis, 1998).

Here are some examples of process research in coaching: Hall, Otazo, and Hollenbeck (1999) interviewed 75 executives and 15 coaches asking about coach behaviors that they found most helpful in the coaching conversations. Both coaches and coachees emphasized listening, but coachees also mentioned the introduction of helpful ideas, feedback, and challenge as being very important.

David, Clutterbuck, and Mergginson (2014) based their study of the effectiveness of coaches in different countries on interviews. Rekalde, Landeta, and Albizu (2015) and Salomaa (2017) studied observations of success criteria (or factors in effectiveness) among coaches, coachees, and sponsors—Salomaa more specifically among expatriate coachees. Burger (2013) talked to top executives of large organizations, executive coaches, and nonexecutive directors, to compare experiences of executive coaching on the part of leaders, sponsors, and coaches. Bickerich, Michel, and O’Shea (2018) studied the views of 15 coachees and 18 coaches on the use of coaching during organizational change. They found clear overlaps in terms of perspectives but also differences. Coachees seemed more focused on practical help in terms of leading the changes and overcoming resistance, whereas coaches focused more on the coachees themselves and the coaching conversations, including emotions during the conversations. Terblanche, Albertyn, and Coller-Peter (2017) studied experiences of transition coaching among 16 recently promoted leaders who had been coached during their transition; they also looked at their coaches and sponsors.

Stevens (2005), Bush (2005), Gyllensten and Palmer (2007), Mackenzie (2007), Passmore (2010), and Timson (2015) analyzed the coachee experience of the coaching relationship by means of retrospective interviews. In the first three studies of 7, 10, and 9 executives with positive memories of coaching, respectively, Stevens (2005) asked open questions, Bush (2005) asked about “effectiveness,” and Gyllensten and Palmer (2007) asked about the “coaching relationship.” The fourth study, by Mackenzie (2007), took place in the context of a major leadership program at the U.K. Royal College of Nursing. It involved interviews with eight coaches and asked a very open question about “your experience with coaching as part of the program.” Passmore (2010) interviewed six coachees about their experience and through grounded theory identified a variety of potentially helpful coaching behaviors. Timson (2015) asked six coachees about their experiences after a brief resilience intervention in the same organization.

Carter, Blackman, Hicks, Williams, and Hay (2017) surveyed 256 managers about the impact of possible supports and barriers on their coaching experience. Waning support from third parties did not have a significant influence on their reported effectiveness but the experienced barriers did. “Unclear development goals and lack of agreement with my coach on my goals” was the single biggest barrier identified by this sample of coachees.

Motsoaledi and Cilliers (2012) conducted research based on psychodynamic role analysis and extensive notes after each session, during 10 months of coaching with six culturally diverse executives in a South African context. They showed how psychodynamic coaching can help in understanding and learning to deal with conscious and unconscious intercultural dynamics.

Some field studies are designed to deepen conceptual understanding. O’Broin and Palmer (2010) conducted field research among six coachees and six coaches to identify the dimensions needed for a “good coaching relationship”; both groups turned out to share very similar ideas about that. Toogood (2012) interviewed six coaches about “positive psychology,” Bachkirova (2015) interviewed six supervisors about the concept of “self-deception,” and Noon (2018) interviewed six coachees and six coaches plus a focus group about the concept of “presence.”

A particularly fruitful program of process research via filmed or recorded sessions was pioneered by Greif (2010) and Graf (2012) and then continued by Adrian Myers (2014). In his PhD, Myers (2014) looked at the descriptions of six video-recorded coaching sessions involving six different coachees and coaches. Session recordings were observed by the coachees and coaches and by 35 observers (one recording each). On reviewing their own sessions with the help of a Q-sort procedure, coachees focused mainly on their interpersonal needs and on the space and structure to develop new perspectives, whereas coaches focused more on the application of their expertise and
on choices made during their interventions. Factor analysis showed that all Q sorts spanned only two factors: client-led coaching and process-led (directive) coaching (Myers & Bachkirova, 2018). Two sessions could be shown to be more client-centered and two more directive (with the other two in between).

Braunschweig University produced a series of quantitative studies based on a similar type of process research, including Ianiro et al. (2013); Ianiro and Kauffeld (2014); Gessnitzer and Kauffeld (2015); Ianiro, Lehmann-Willenbrock, and Kauffeld (2015); Will, Gessnitzer, and Kauffeld (2016); and Gessnitzer, Schulte, and Kauffeld (2016). Here is a summary of these researchers’ findings, based on their analysis of thousands of consecutive moments in sessions:

1. Ianiro et al. (2013) coded and analyzed the complete coaching interaction belonging to 33 initial coaching conversations, on the basis of video recordings. Trainee psychologists acted as coaches, with students from other faculties taking the role of the coachees. The authors observed how coachee and coach behaved toward each other, focusing specifically on two fundamental variables: affiliation and dominance (Leary, 1957). From their findings they concluded that both (a) the dominant behavior of the coach and (b) comparable (observers’) scores for dominance and affiliation in coach and coachee predicted that after five sessions the coachee would award more positive scores for (1) achieving goals and (2) quality of the relationship with the coach.

2. Ianiro and Kauffeld (2014) coded and analyzed video recordings of the first coaching session of 48 coaching dyads and, just before the session, asked all coaches and coachees to score their mood using the Multi-Dimensional Mood Questionnaire (with subscales: good/bad mood, awake/tired, and calm/nervous). The coachees also scored the working alliance after the fifth session, that is, 3 months later. With respect to coach behavior, once again coach and coachee dominant-friendly behaviors (as seen by the observers) were found to correlate clearly, and both also correlated with the working alliance as scored by the coachee; in fact, coach behavior emerged as the best predictor. The authors further found that coach mood (and also coachee mood) correlated significantly with their own dominant-friendly behavior and with the working alliance as ultimately scored by the coachee, but only on the good mood and calm subscales, not on that of awakeness, although the latter was negatively correlated with dominant-friendly coachee behavior.

3. Gessnitzer and Kauffeld (2015) next coded the video recordings of the first, third, and fifth (last) coaching sessions of 31 coaching dyads (i.e., 93 sessions in total). They showed that the working-alliance aspects “agreement on goals and tasks” and “affective bond” coded by observers scarcely correlated with the same working-alliance aspects measured by coach and coachee using a standard questionnaire. Only the observed working-alliance behaviors with respect to “agreement on goals and tasks” initiated by the coachee turned out to correlate with coaching outcomes (as measured by coachees in terms of the progression toward their own goals).

4. Ianiro et al. (2015) looked for the first time at sequences in the coding, that is, at the order of behaviors between coach and coachee. All five coaching sessions of 31 psychology students, who were again coaching students from other faculties, were this time recorded and coded per behavioral segment. Frequencies of behavior sequences were measured and by comparing with random fluctuations Ianiro’s team could demonstrate that friendly coaching behavior significantly evoked friendly coachee behavior and vice versa. Moreover, the degree of this “mutual affiliation behavior” predicted a higher score on working alliance. In addition, dominant-friendly coaching behavior was found to be the only behavior that significantly evoked dominant coachee behavior and vice versa (whereas other dominant coach behavior, both neutral and hostile, in fact evoked dependent coachee behavior). Dominant coachee behavior related in turn to positive outcomes from the third session onward (outcomes in terms of self-scored effectiveness on goals). Moreover, the complementary hypothesis also appears to be true: Dependent coachee behavior was significantly inversely proportional (negatively correlated) to the same outcomes.

5. Will et al. (2016) analyzed 19 video recordings of one-off sessions between 19 coach–coachee pairs in the same university setting. They measured not only self-perceptions of coaches’ (cognitive) empathy but also their coachees’ perceptions and compared these with observed empathic interventions (consisting of paraphrasing and naming coachee feelings). The authors found no correlation between coaches’ and coachees’ assessments of the
coach’s empathy. They went on to demonstrate that in these coaching sessions only coach paraphrasing was correlated with the coachee’s perception of coach empathy. However, sequential analysis showed that both (observer-rated) aspects of empathy (paraphrasing and naming coachee feelings) in the behavioral analysis of the video images induced a positive response in coachees with a frequency that could be classed as significant.

6. Gessnitzer et al. (2016) have looked at the development of self-efficacy over time, by analyzing video recordings of the first, third, and fifth session of 31 coaching dyads. Their analysis of video images showed that self-efficient coachee statements (i.e., self-beliefs about success) increased with the amount of coaching and predicted goal-attainment as well. They were even able to show that in many cases open questions and offering support by coaches led to more self-efficient coachee statements (as did offering solutions, a very directive intervention, but only in the very first session) and that the support of coaches after a self-efficient statement could significantly lead to more self-efficient coachee statements immediately after. In this way, Gessnitzer et al. (2016) managed to clarify and deepen many results from quantitative coaching results about the high predictive value of coachee self-efficacy.

In summary, these six articles showed convincingly that a coach’s positive-friendly mood prior to the session can lead to measurable (dominant-friendly) coach behaviors, which in turn can evoke similar coachee behavior, leading to demonstrably better results in later sessions. They also showed that self-belief by coachees and (observer-rated) empathic behavior from coaches can make a positive difference to outcomes. At the same time, the coach’s own assessment of both his or her own empathy and the quality of the coaching relationship showed no correlation with the empathy and working alliance that can be measured by observers on the basis of video recordings (and only the latter correlated with outcome). These studies found in particular that the coachee-initiated “agreement on tasks and goals” (as measured by observers) showed a significant correlation with outcomes.

This research program also demonstrated a wider point: Process research on the basis of a large volume of data from qualitative data collection can make powerful, verifiable predictions about all coaching assignments, just as the more traditional quantitative research. In principle, this applies to all qualitative research (including case studies and interviews): With sufficient measurements based on a representative sample, one can turn any qualitative exploration into a sustainable and generalizable quantitative research result. However, to make statistically significant statements, the numbers of these measurements can run into the tens of thousands, as in each of the “Braunschweig Gruppe” articles cited above.

Over the period between 2008 and 2015, the Ashridge Critical-Moments Study Group has published process research based on descriptions of critical moments as experienced during coaching. They compared the descriptions given by inexperienced coaches, experienced coaches, coachees, sponsors, and coachees/coaches directly after their sessions. This work can be briefly summarized as follows:

1. de Haan (2008a) analyzed 80 descriptions of critical moments encountered by relatively inexperienced internal and external coaches in their work. Around three quarters of the participants in this study were organizational consultants who recently completed a full year of studies in management consulting, with the other participants being beginning executive coaches. Content analysis showed that all of the critical moments could be interpreted as expressions of doubt on the part of the coach.

2. de Haan (2008b) and Day, de Haan, Sills, Bertie, and Blass (2008) analyzed 78 and 49 descriptions, respectively, of critical moments of coaches with at least 8 years of coaching experience. Content analysis showed that all of these critical moments could be characterized as anxieties in the coach. In other words, these experienced coaches struggled with recurring anxieties and to a lesser extent doubts in their practice with coachees. Day et al.

---

1For a more extensive overview of the seven research articles by the this study group, please see the book *Critical Moments in Executive Coaching* (De Haan, 2019).
(2008) made use of in-depth telephone interviews to investigate the temporal process around critical moments. They showed that these coaches consistently had come across a minor or major “rupture” in the relationship (e.g., incomprehension, anger, recontracting, referral, nonattendance, or termination). The presence of ongoing, shared reflection determined in nearly all cases whether the rupture in the relationship was overcome. If there was no shared reflection, then the rupture usually led to an agreed ending or referral.

3. de Haan, Bertie, Day, and Sills (2010a) analyzed 59 descriptions of critical moments of coachees. Content analysis showed that coachees considered new forms of awareness and new insights to be the most important experiences in their coaching process. This finding differed from de Haan’s earlier findings (2008a and 2008b) with regard to executive coaches. At the same time, this supported findings by, for example, Hall et al. (1999) and Blackman (2006), who also argued that coachees want (actionable) information, insight, or challenge from their coaching conversations.

4. de Haan, Bertie, Day, and Sills (2010b; with a field study) and de Haan and Nieß (2012; with a case study) analyzed 86 and 32 descriptions of critical moments, respectively, of coachees and coaches that were recorded immediately after agreed coaching sessions; that is, in this case the response did not come from coaching sessions selected by the respondent but from the particular session on the day of writing. In this research, all previous data sets were also recoded by four raters using a simple coding structure with 12 codes (see Table 2). The coding confirmed that, in the case of inexperienced coaches, Code 12 for doubt in the coach was given most frequently; in the case of experienced coaches, the code given most frequently was Code 10 for anxiety in the coach, whereas for coachees the codes most often used were those for new insights and reflections (Codes 1 and 2). In this research, both coaches and coachees cited new reflection and insight as being of primary importance in their direct coaching experiences.

The authors hypothesized that the six data sets so far largely fall into two different classes of descriptions. In the data sets of de Haan (2008a and 2008b) and Day et al. (2008), coaches cited critical moments from their coaching work in the previous year, thus mainly having contributed

### Table 2

**Critical-Moments Coding Scheme as Found in de Haan et al. (2010b)**

<table>
<thead>
<tr>
<th>Code number</th>
<th>Short description of the code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A moment of learning: a moment in which new insight was created for coach and—particularly—coachee.</td>
</tr>
<tr>
<td>2</td>
<td>A moment of learning: a moment of working through, reflecting, gaining new perspectives, or making sense of existing material.</td>
</tr>
<tr>
<td>3</td>
<td>A change in the relationship in the moment (positive).</td>
</tr>
<tr>
<td>4</td>
<td>A change in the relationship in the moment (negative).</td>
</tr>
<tr>
<td>5</td>
<td>Significant action in the moment (coach-led): applying oneself to a unique scripted process such as drawing, visualization, role-play, GROW, . . .</td>
</tr>
<tr>
<td>6</td>
<td>Significant action in the moment (coachee-led): organizing future sessions, negotiating the session, taking away action points, making notes, . . .</td>
</tr>
<tr>
<td>7</td>
<td>Significant emotional experience in the moment: joy (coachee); heightened positive emotion.</td>
</tr>
<tr>
<td>8</td>
<td>Significant emotional experience in the moment: joy (coach); heightened positive emotion.</td>
</tr>
<tr>
<td>9</td>
<td>Significant emotional experience in the moment: anxiety (coachee); heightened positive emotion.</td>
</tr>
<tr>
<td>10</td>
<td>Significant emotional experience in the moment: anxiety (coach); heightened positive emotion.</td>
</tr>
<tr>
<td>11</td>
<td>Significant emotional experience in the moment: doubt (coachee); fundamental not-knowing, often a starting point for reflection.</td>
</tr>
<tr>
<td>12</td>
<td>Significant emotional experience in the moment: doubt (coach); fundamental not-knowing, often a starting point for reflection.</td>
</tr>
</tbody>
</table>

*Note.* The coding scheme is based on four broader categories: moments of learning (Codes 1 and 2), moments of relational change in the coaching relationship (Codes 3 and 4), moments of significant action (Codes 5 and 6), and moments of significant emotional experience (Codes 7 to 12).
examples of “exceptional circumstances,” such as ruptures in the coaching relationships. In de Haan et al. (2010a), coachees looked at a single coaching relationship, and in de Haan et al. (2010b) and de Haan and Nieß (2012), coachees and coaches looked at a single coaching session; on average, therefore, they probably had in mind more “everyday” coaching sessions rather than the “coaching in extraordinary circumstances” of the other three data sets. This hypothesis could explain why the coach data over a longer period in the first three articles (de Haan, 2008a, 2008b; Day et al., 2008) and the coach and coachee data over a much shorter period (de Haan et al., 2010a, 2010b; de Haan & Nieß, 2012) consistently failed to correlate when they were coded using the 12 codes.

De Haan et al. (2010b) and de Haan & Nieß (2012) also found that in around half of cases, coachees and coaches cited the same moments when asked independently what they regarded as “critical” in their session.

5. de Haan and Nieß (2015) contained 147 descriptions of critical moments from 49 coaching relationships as recorded by coachee, coach, and sponsor, plus 30 other critical-moment descriptions from sponsors. The results of this research show that there is much more alignment between coachees and coaches on the critical moments in their coaching relationship than there is between both those parties and the sponsors. In the main, coachees and coaches still described as “critical” those moments in which they gained new perspectives and insights. Sponsors, for their part, appear to attach more importance to new initiatives and behavioral changes initiated by the coaching coachee (Code 6 in Table 2). There were again indications that coachees and coaches quote the same moments to an extent that goes beyond pure chance.

I have found one more field study and two process studies on a smaller scale focusing on critical moments in coaching: Lightfoot (2019) recorded one session of six coaching dyads and interviewed the coach and coachee about one critical moment each. Following de Haan et al. (2010b), he found full overlap between coach and coachee accounts and was able to follow various recurring steps of these “insight events.” Fatien-Diochon and Nizet (2015) analyzed 27 “moments of ethical challenge” from the practice of an equal number of French accredited coaches. They concluded that codes of conduct have very limited use in such challenges: Ethical codes are experienced by coaches to be either irrelevant, insufficient, or even an obstacle when it comes to resolving the challenges. Turner and McCarthy (2015) collected and analyzed descriptions of “coachable moments” of line managers, that is, examples of informal, unplanned opportunities to coach their direct reports.

In addition, Smith and Brummel (2013) recorded interviews with 30 executives about their coaching and had these coded in terms of three proposed active ingredients of executive coaching: the coachees’ involvement with coaching, their perceptions of improvability, and their personal development plans. This coding was compared with the coachees’ self-reported competency development and the examples coachees gave of their development. The statistical analysis of the codes showed significant relationships with the reported competency change, for all three proposed ingredients.

All in all, this broad range of process research provides an understanding of how coachees, coaches, and their sponsors each experience coaching contracts and sessions differently—although a surprisingly large correspondence can also be demonstrated between these different perspectives on coaching.

Relevance of Process Research to RQ1, RQ2, and RQ3

With process research in coaching we have recently achieved a scale of investigation with which we can demonstrate generalizable results, that is, beginning to address RQ3. There is a variety of strikingly different process-research studies that report that coaches and coachees notice or look for different things in coaching conversations. Broadly, coaches emphasize the importance of listening; attending to the conversations themselves; and noting emotion, doubt, and anxiety. Coachees, in comparison, place more emphasis on practical new help, insights, challenges, and perspectives (Bickerich et al., 2018; de Haan et al., 2010a; Hall et al., 1999; Myers, 2014). de Haan and Nieß (2015) also traced other demonstrable differences in perspective between coachee, coach, and
sponsoring. Likewise, the Braunschweig studies demonstrated several significant predictors of coaching outcome, which will be treated more fully in the Discussion section.

**Discussion**

If we turn back to the original research questions, it is possible to see what qualitative research is already demonstrating and how qualitative findings can be synthesized together to form a deeper understanding of the coaching intervention and provide a starting point for doing more generalizable research.

With respect to RQ1, about potentially important variables in coaching that qualitative research has unearthed, the following seem to be relevant particularly as they are all currently underresearched variables in quantitative studies.

Important success criteria of coaching seem to be related to the coachee: the development of trust in, acceptance of, and commitment to coaching, and respect for the coaching contract (see, e.g., Alvey & Barclay, 2007; Audet & Couteret, 2012; Smith & Brummel, 2013). An important success criterion for both coaches and coachees seems to be the ability for both to achieve a deep level of psychological reflection and understanding (Machin, 2010; Styhre, 2008).

Many important barriers have also been explored, particularly from the coachee perspective: for example, lack of agreement of goals with coaches (Carter et al., 2017) or a lack of trust, even in some cases cynicism (Gray et al., 2015) and sexism of coachees (Gray & Goregaokar, 2010). Hidden agendas and politics are perhaps more important than we think (Ben-Hador, 2016; Gray & Goregaokar, 2010; Peil, 2008; Skinner, 2012), as well as lack of transfer to the workplace (Gray et al., 2011; Levenson, 2009). Lack of support for coaching from third parties does not seem as much of an issue as these barriers (Carter et al., 2017).

On the part of the coach, a crucial factor seems to be the ability to help with the potential success criteria above. In particular, it seems important for coaches to help actively to gain new (psychological) insight (Bickerich et al., 2018; de Haan et al., 2010a; Hall et al., 1999; Machin, 2010; Myers, 2014), as this seems to be a key ingredient for the coachee.

In all these apparently important ingredients of coaching, the eye of the beholder matters a great deal: Coachees, coaches, and observers look very differently at their common endeavor (see, e.g., de Haan & Nieß, 2015; Myers & Bachkirova, 2018; Will et al., 2016). When studying important aspects such as the coaching relationship, trust, and the coach’s empathy, it is important therefore to specify the perspective. Indications are that measurements by different parties do not correlate, and it may be that aspects that cannot be well reported by coach or coachee are nevertheless important; for example, in the area of perceived empathy it seems that only the observers’ views are correlated with goal attainment (Will et al., 2016). Similarly, it is very important to understand how coaches themselves experience the prevailing ethical codes of practice and their value for coaching (Fatien-Diochon & Nizet, 2015).

With respect to RQ2, which asked about applications of executive coaching in different contexts, coaching has now been undertaken in many different organizations and cultures. Case studies show how coaching can support a wider organizational agenda, such as radical change (e.g., Fahy, 2007; Nanduri, 2018; Schnell, 2005). Several studies looked at coaching in British SMEs and argue that such contracts will be different because of the more personalized nature of work relationships and the dominant nature of senior management (Gray & Goregaokar, 2010, Gray et al., 2011; Peil, 2008).

The extant literature also shows that we can learn through case studies and field research undertaken by specialized coaches, for example, about how to handle aspects such as psychometric instruments, internal coaching, leadership dynamics in SMEs, coaching for entrepreneurs or expats, and coach-coachee matching. Case studies illustrate how the approach to executive coaching has changed profoundly in the last 25 years (compare, e.g., Winum, 1995, or Diedrich, 1996, on the one hand, with Audet & Couteret, 2012, or Nanduri, 2018, on the other). There is some support for specific approaches, such as those inspired by cognitive therapy and positive psychology. Coaching in a second language is entirely possible and has fewer disadvantages than expected for the coaching...
experience and its outcomes (Cox, 2012), and the same is true for coaching across racial divides (Motsoaledi & Cilliers, 2012).

With respect to RQ3, about rigorous quantitative results stemming from purely qualitative research, there are quite a few of those coming into view in recent years.

First a caveat. Quantitative research need not be the ideal “end phase” or “final objective” of qualitative research. Smaller collections of measurements can also yield surprising, innovative, or even theory-confirming insights with a significant influence on researchers and professionals. Take, for example, the pioneering work of Sigmund Freud into helping conversations. The only research he undertook to test some of his many bold statements about the human mind and about therapy consisted of a small number of highly subjective (i.e., described and analyzed by Freud himself) case studies, the most famous of which are those of Anna O. and Dora (see Breuer & Freud, 1895/1960; Freud, 1905/1960). These are only limited measurements (“N = 1” for each case study, so by no means statistically generalizable). However, these two case studies have had a profound impact on the field: Anna O. on the introduction of the psychoanalytical method, and Dora on the understanding of transference (and as an example of the use of dream analysis in therapy). Many generations of psychoanalysts, therapists, and even coaches have employed these ideas and anchored their work in them, so their influence has far exceeded any (generalizable) significance they might have. This is still the case with original and influential qualitative studies.

The long-running program in Braunschweig shows how self-efficacy, empathy, and the working alliance can all be expressed in observable behavior, which can subsequently be rated reliably by observers. These studies have also shown examples of behavior that might be related to the effectiveness of coaching. Specifically, they found that (observer-rated) “dominant-friendly” behavior by both coachee and coach (Ianiro & Kauffeld, 2014) and “self-efficient” communications from the coachee (Gessnitzer et al., 2016) were significantly related to outcome. Moreover, the (again, observer-rated) support given by the coach to self-efficient statements from the coachee was significantly followed by more of those statements. Finally, a strong working alliance, particularly in terms of “agreement on tasks and goals initiated by the coachee,” led to better goal attainment (supporting many quantitative studies and also what coachees report about serious barriers when there is no agreement on goals, in Carter et al., 2017).

The 10-year-long program in Ashridge also informs us in reliable ways about micro-outcomes from coaching sessions. This research has shown that coaches make sense of their critical moments in coaching by often referring to their own doubts and anxieties (de Haan, 2008a and 2008b), whereas coachees experience their key moments more as new learning and realization (de Haan et al., 2010a, confirmed by several other qualitative studies cited above), and sponsors mainly experience new initiatives and behavioral changes (de Haan & Nieß, 2015). Even though coachees and coaches therefore look back on their critical moments rather differently, they do seem to select the same moments in a surprising number of cases and describe these in the same terms after joint sessions, namely as new learning and realizations (de Haan et al., 2010b).

These research programs offer an appreciation of coaching in terms of micro-outcomes inside sessions, either in terms of behaviors (Braunschweig group) or in terms of critical moments as experienced by all parties (Ashridge group). Together, these endeavors are beginning to open up the coaching sessions themselves to quantitative research and verifiable findings. The findings themselves are amenable to coach training and supervision, as they show how parties experience the sessions differently and how particular coach behaviors are more likely than others to achieve good results in terms of the coachees’ goal attainment. Here are some practical suggestions for coach development and supervision:

- An increased use of observed sessions, with observers spotting important aspects for effectiveness such as “agreement on tasks and goals,” “dominant-friendly coaching behaviors,” “received empathy,” “new realizations for coachees,” and “self-efficient statements of coachees.”
- Training and supervision on contracting for coaching results, including (a) coachee-initiated agreement on contracts, (b) ways of enhancing coachees’ trust and their commitment to the
contract, and (c) an overall sensitivity to the psychological contract where lingering barriers
to commitment can be found.

- Learning how to look at coaching sessions from multiple perspectives and to appreciate
differences that can be found in these perspectives, such as the perspective of the coachee
(mainly on insight and new perspectives), the sponsors, and colleagues (mainly on new
initiatives and actions), as can be measured by observers.

- Learning to work across diversity to make coaching contracts more ethical, particularly
when powerful others in the organization might interfere with the contract, for example, in
ways that are more relational and less dependent on detached and objective codes of
conduct.

**Limitations**

First, because of my deliberate choice to survey the whole purely qualitative-research literature in
executive and workplace coaching, this systematic review has wrestled with a large heterogeneity
of individual studies. These studies lie across the widest possible range of qualitative research and,
moreover, are situated in a new domain of research (which only emerged in the late 1990s). This has
meant that most studies reviewed stand on their own and were not anchored in a structured,
methodical research base. This heterogeneity has meant that it was harder to synthesize and draw
general conclusions true for all studies; only mutually confirming trends could be described. Also,
some relevant qualitative outcomes may have been missed altogether, because of excluding the
mixed-methods studies.

Second, an important limitation has been the differing quality of the individual studies, which
has been somewhat counteracted by synthesizing findings and drawing mostly on the more rigorous
studies and the larger samples, on which all conclusions in this Discussion section have been based.

Third, this systematic review has had to deal with the limited generalizability of all qualitative
research. Generalizability has much improved through finding many areas of overlap between the
individual studies; however there are still very few specific recommendations coming out of this
research base. We should eventually expect improved generalizability coming from more rigorous
research done on a larger scale; hopefully that will emerge over the next decades.

Fourth, being very inclusive about the primary studies incorporated entailed the risk of reducing
the range and depth of interpretations of the coaching phenomenon. Nevertheless, this systematic
review encompassing a wide spectrum may constitute a marker for future qualitative coaching
studies in the domain of workplace and executive coaching. Although most reviews have to rely on
a representative sample of published studies to generate a conceptual consolidation, this metasyn-
thesis offers also an empirical consolidation based on a wide and exhaustive search strategy.

**Future Research**

This synthesis of existing research has shown the great value of a more rigorous research design that
is replicated over several data sets (such as, e.g., the longstanding Braunschweig and Ashridge
partnerships mentioned above) or executed in a systematic and statistically informed way (such as
can be argued about Audet & Couteret, 2012; Gray et al., 2011; or Myers & Bachkirova, 2018).
These are designs that are easily replicated and are offering a great promise for the future. I hope
that researchers thinking of conducting a new piece of qualitative research in coaching would
consider the possibility of adapting their design to some of those more rigorous ones: Their new
study will stand well on itself and will also become a valuable extension of the thorough work
previously undertaken by others. Finally, we need more careful action-research studies that, even if
not easily generalizable, are providing valuable access to the inner world and rich experience of
coaching professionals—and also of their clients.

**References**

References marked with an asterisk denote all studies included in the systematic
review.


THE EMERGENCE OF QUALITATIVE RESEARCH IN COACHING


Received January 1, 2019
Latest revision received June 30, 2019
Accepted July 18, 2019