Using Clean Language to explore the subjectivity of coachees' experience and outcomes

Authors

Dr Susie Linder-Pelz

Independent researcher and director of Good Decisions, Sydney

susie@gooddecisions.com.au

au.linkedin.com/pub/dr-susie-linder-pelz/8/9a0/5b4

James Lawley

Independent researcher and partner in The Developing Company, London

james@cleanlanguage.co.uk

www.cleanlanguage.co.uk

Correspondence:

Dr Susie Linder-Pelz

Email: susie@gooddecisions.com.au

Abstract

Objectives

This paper aims to contribute methodologically and substantively to understanding how coachees experience and evaluate coaching. First, we explore the use of 'Clean Language' as a phenomenological approach to coaching research, including the eliciting and analysing of data into findings and insights for coaches and coach trainers (Tosey *et al*. 2014, p. 630). Second, we explore the nature of events, effects, evaluations and outcomes reported by coachees after a single coaching session.

Design

Three coaches accredited in the same coaching methodology each delivered a single session to two randomly allocated coachees. The coachees were subsequently interviewed twice using Clean Language, in person two days after the coaching and by telephone two weeks later.

Methodology

The transcribed follow-up interviews were analysed by an expert in Clean Language (the second author), using a form of thematic analysis within a realist/essentialist paradigm (Braun & Clarke, 2006, p. 85).

Findings

The interviews elicited detailed information on many aspects of coaching without the interviewer introducing any topics. Coachees' events, effects and evaluations happened during the coaching session, between that session and the first interview, and during the two weeks between the first and second interviews. Coachees emphasised coaches' style of

repeating back, pacing, setting goals and questioning, maintaining the focus of the session, confronting and challenging, as well as their responsiveness (or lack of it). Increased self-awareness was mentioned by all coachees. Outcomes occurring after the session were maintained two weeks later, at which time new outcomes were also reported.

Conclusions

Clean Language Interviewing supplements and extends existing methods of phenomenological interviewing and data coding. The study yielded nuanced findings on the coach behaviours that led coachees to give favourable versus unfavourable evaluations, with implications for coaching psychologists with regard in particular to coaches' ability to calibrate and respond to coachees' ongoing evaluation of the coaching, the pace of the session and how the timing of coachees' feedback affects the findings.

Keywords: Clean Language Interviewing**,** phenomenological research, coachees' experience, coaching outcomes, coach training, coaching evaluation

Introduction

At the heart of interviewing is an interest in understanding peoples' experience and the meaning they make of it (Knox & Burkard, 2009, p. 2) and the value of qualitative interviewing in coaching psychology research is widely recognised (de Haan & Nieß, 2012; Greif, 2007). While researchers may believe that their interviews are free of prejudice and presupposition, the wording of a question can inadvertently influence an interviewee's recall and response (Loftus, 1975). For example, in a study of coaching outcomes two questions were asked about the perceived benefits of the programme and 'this may have inadvertently precluded participants from identifying unhelpful aspects of the coaching programme.' (Grant, 2013, p. 19).

While O'Broin & Palmer (2010) used semi-structured Repertory Grid interviews based on Personal Construct theory, their wording is drawn from existing evidence-informed literature (p.125). Many of their probe questions contain strong researcher metaphors (e.g. 'handle a rupture') and presuppositions (e.g. 'how important is the coach-client relationship to outcome?', p. 143). This last statement presupposes a causal link between the relationship and the outcome; it also presupposes that the link is important. The interviewee is therefore more likely to answer within this specified frame – whether or not they held this belief beforehand.

Passmore (2010) used a semi-structured interview method centered around six predefined themes (p.51). Unfortunately he does not give examples of the precise wording of the additional prompts and questions asked to stimulate interviewees to respond, so we cannot know whether the interviewer introduced questions whose syntax inadvertently led or constrained the interviewee's answers. This could cast doubt on the authenticity of the data.

'The coaching psychologist…needs to be aware of the effects of different questions and the most appropriate timing for each.' (Hieker & Huffington, 2006, p. 48). 'Even the subtlest instantiation of a metaphor (via a single word) can have a powerful influence [and furthermore] the influence of the metaphorical framing effect is covert: people do not recognize metaphors as influential' (Thibodeau & Boroditsky, 2011, p. 1). 'People' in this context includes both interviewer and interviewee.

Clean Language was developed from the clinical work of David Grove (Grove & Panzer, 1989) and extended into other areas such as management development, organisational change and education (Lawley & Tompkins, 2000). By paying careful attention to their language, researchers can minimise undesired influence and unintended bias during all stages of research—design, data gathering, analysis and reporting (Van Helsdingen & Lawley, 2012). In particular, Clean Language can refine interviewing by minimising the introduction of researchers' metaphors and constructs (Tosey *et al.*, 2014).

The Clean Language interviewing (CLI) method fits within a phenomenological methodology (Owen, 1989; Tosey, 2011; Worth, 2012) which addresses widely recognised difficulties in the process of exploring and explicating a person's self (Greif, 2007, p. 223). CLI is also grounded in NLP-based modelling (Tosey & Mathison, 2010) behind which are constructivist and systemic assumptions (Linder-Pelz, 2010, pp. 81-82). The intent of Clean Language is akin to the practice of 'bracketing' in other phenomenological research; it is an attempt to suspend prior knowledge or belief about the phenomenon under study (Tosey *et al.*, 2014, p. 633; Vansickel-Peterson, 2010, p. 56).

A comparison with the established phenomenological method Interpretive Phenomenological Analysis (IPA) is instructive. Both CLI and IPA aim to explore and understand meaning-making. With IPA the researcher is active in interpreting the participant's experiences; by contrast CLI aims to facilitate interviewees by restricting interviewer

interpretation, impositions or effects especially because 'individuals may have difficulties reporting what they are thinking or/and they may not want to self-disclose' (Gyllensten & Palmer, 2007, p. 170). While CLI restricts interviewer influence it does not eliminate it; the interviewer is still selecting aspects of the interviewee's description to focus on, the Clean Language questions to ask and how to ask them.

Clean Language also offers an alternative to the semi-structured interview method which explicitly centres on themes (Braun & Clarke, 2006, p. 94; Passmore 2010, p. 51). By being more attuned to the individual interviewee rather than following the same path for all respondents, Clean Language can enhance rapport and mitigate against the possibility that findings may represent participants' authentic responses (Tosey et al., 2014, p.641). The adherence to a strict protocol prevents the interviewer from introducing into the conversation content or leading questions such as, 'Have you ever experienced something that felt like a "critical moment" … an exciting, tense or significant moment?' (de Haan *et al*., 2010, p. 8). This ensures that the descriptions obtained are sourced exclusively in the interviewee's personal vocabulary and experience. However, Clean Language is not appropriate where the interviewer is aiming to co-create meaning (Tosey *et al*., 2014, p. 641).

One application of Clean Language interviewing is to facilitate interviewees to describe 'what' and 'how' they evaluate an experience (Lawley & Tompkins, 2011). It can address what Hall (2013) calls 'evaluative vagueness' where researchers often only get the 'bottom line' of a person's evaluations and do not know the criteria by which the person made those judgments. Hence CLI could be used by coaching psychologists who want to better understand how coachees evaluate coaching (de Haan *et al*., 2011).

Research questions

The two research questions addressed in this paper arise from the potential of CLI to explore in detail *how* coachees evaluate.

1. Does Clean Language interviewing of coachees about their experience and evaluation of coaching yield insights for coaches and trainers of coaching?

2. What events, effects and outcomes, during and after a single coaching session, do coachees use to evaluate coaching?

Methodology

This paper reports on part of a mixed method study which triangulated the views of coaches, coachees and a coach trainer/expert with regard to single coaching 'breakthrough' sessions (Lawley & Linder-Pelz, 2014). Six new/naive volunteer coachees were randomly assigned to one of three practising coaches certified in the same coaching methodology. Convenience and purposive sampling was used via a request from the lead researcher to colleagues and acquaintances. Each coachee participated in a single coaching session and two subsequent Clean Language interviews, all of which were recorded. The participants were unknown to the coaches and the interviewer.

The coaching method adopted was Meta-Coaching, a goal-oriented coaching methodology based explicitly on cognitive-behavioral psychology (Linder-Pelz & Hall, 2008). We chose this method because Meta-Coaches are trained and assessed in specific, benchmarked competencies as a requirement for certification (Hall, 2011; Linder-Pelz, 2014). There is no unanimity as to whether coaching method plays a part in effectiveness (Grant et al., 2010); neither is there unanimity regarding what specific behaviours constitute competency in coach-coachee relationship skills such as listening (Passmore & Fillery-Travis, 2011). Hence it is reasonable to expect that studying a single coaching method would

reduce the variability compared to using coaches with different approaches and requiring different experts to assess their competencies.

Expecting within-coach variability as well as between-coach variability—and given resource constraints—we recruited three coaches and six coachees in order to explore whether or not coaching skills related to the coachees' experiences and evaluations.

We developed protocols for selecting and briefing coaches and coachees as well as for data collection and for the analysis of the Clean Language interviews. These are available at www.cleanlanguage.co.uk/articles/articles/350/.

The coaches

The three coaches had demonstrated coaching competency, having been certified by the Meta-Coaching benchmarking system (Linder-Pelz, 2014). The core Meta-Coaching skills are listening, supporting, questioning, meta-questioning, inducing states and giving and receiving feedback (Hall, 2011). All coaches were women aged in their 30s and 40s, running their own practices with paying clients and active in Meta-Coach training and mentoring programs. Their task was to conduct a single 1.5-hour Meta-Coaching session with two randomly allocated coachees.

Coachees

Given that there would be variability in ages, genders and prior experience of change work, we recruited six coachees. They comprised a convenient and purposively-selected group of volunteers who met our criteria of: (1) having no prior experience of Meta-Coaching; (2) having something meaningful they wanted to change in their life; (3) not currently seeing a coach, psychologist or psychotherapist; and (4) never having been diagnosed with a major psychological disturbance. All were aged from mid-30s to early 60s, five were women, three had had previous coaching or counselling and two had some prior experience of NLP (Neuro-

Linguistic Programming) on which Meta-Coaching is based. The topics they chose to work with included health, building a business, confidence at work, self-worth, a relationship concern and managing money. All gave informed consent to the recording and use of their interviews on the condition of anonymity. Although small, the sample size was sufficient for the purpose of investigating the value of the CLI method and exploring the resultant findings (Tosey *et al.*, 2014, p. 634).

Procedures

The coaching sessions all took place on the same day in an office setting that was not any of the coaches' work premises. The first interviews using Clean Language took place in person two days after the coaching session and lasted, after preliminaries, for between 37 and 51 minutes. The second interview, by telephone, was two weeks later and took between 10 and 22 minutes. All interviews were audio-recorded and transcribed. Coachees were informed that their coach would not be privy to anything said in the interviews.

Clean Language Interviews

Using Clean Language as an interview methodology meant that the interviewer aimed to not introduce *any* topic or content into the conversation, ensuring that the descriptions obtained were sourced exclusively in the interviewee's personal vocabulary and experience.

Analysing the interviews

The method of analysis (see Appendix) followed the stages of thematic analysis: familiarisation with the data, generating initial codes, searching for themes, then reviewing, defining and naming themes (Braun & Clarke, 2006, p. 87). However, a new method was employed in Stage 2, 'generating initial codes', where the analyst distinguished between

items that were 'events', 'effects', 'outcomes' and 'evaluations' *from each coachee's perspective*. An 'event' is a description of what happens during the coaching session, an 'effect' is the impact on the coachee during the session, an 'outcome' is an effect that happens *after* the session and an 'evaluation' involves the coachee making a favourable or unfavourable assessment of the experience. We prefer the terms 'favourable' (expressing approval) and 'unfavourable' to 'positive' and 'negative' because the former retain the sense of being the *interviewee's* preferences rather than an external observer's assumptions. Our findings show that favourable evaluations are not necessarily tied to beneficial outcomes and unfavourable evaluations do not necessarily result in adverse outcomes.

Examples from two coachees on the same topic of 'coach neutrality' illustrate the data coding:

> I liked the way that she stayed very neutral [favourable EVALUATION of EVENT]. It was soothing, calming [favourable EVALUATION of EFFECT]. If I did drop a bomb [or] say something that was quite personal [EVENT], I couldn't see it anywhere on her face or in her speech. She wasn't moved in any way [EVENT]. I appreciated that [and] didn't feel judged [favourable EVALUATION & EFFECT].

> I admire that … she could maintain that neutrality [favourable EVALUATION of EVENT] … But I will say on the other hand … I did feel like there was a lack of rapport in that sense [unfavourable EVALUATION of EFFECT].

The coachee's evaluation—and whether it is favourable or unfavourable—can be explicit or implicit. In the first example above, 'I liked' and 'I appreciated' clearly indicate favourable evaluations, while 'soothing and calming' and 'didn't feel judged' *imply* a favourable

evaluation because it is a our assumption that the coachee considers these qualities to be desirable.

The detailed analysis was undertaken by the second author and can be considered dependable (Mays & Pope, 1995, p. 112) or trustworthy (Sousa 2014, p.213; Worth, 2012, p. 70) because of adherence to protocols, adequacy of data, audit trail and the repeated reflexive discussions between the analyst and the first author. The Appendix shows the systematic steps taken by the analyst, which could be replicated by someone else trained in the method.

Findings

The identified events, effects, outcomes and evaluations of the coaching were separated into five time frames, depending on when the interviewees reported they occurred: at the beginning of the session, throughout the session, towards the end of the session, in the first two days after the session (i.e. before the first interview) and in the two weeks between the first and second interviews.

The findings are presented in order of time frame and are illustrated with quotes selected from all six coachees. These findings need to be considered in the context of the interviewees' overall assessment of their coaching; we show elsewhere that all coachees evaluated their session highly and that their numerical ratings and verbal reports shed further light on how they evaluated the effects (Lawley & Linder-Pelz, 2014).

BEGINNINGS OF SESSIONS

Four of the six coachees referred in unfavourable terms to coach behaviour occurring in the first 30 minutes of the coaching session. Words were repeated back too often, the overall pace

was too fast or too slow and the manner of goal setting too direct. However, only one of these coachees said that the disliked behaviour affected the final evaluation.

Repeating back words and managing the pace of the session

These four coachees found the manner and frequency with which their words were repeated back at the beginning of the session 'irritating', 'just constant', 'distracting' or 'a bit repetitive'. However, the coachees reported that after 20-30 minutes their coaches did more than simply reflect what was said; the coaches were 'listening for what was going on underneath' and 'had the thread of where I was going', which received favourable evaluations.

Three of these four coachees also did not appreciate the speed of the beginning of the session. It either 'went at a very rapid pace' or was 'a little bit slow'. One coachee was thinking: 'Come on, let's move on and talk about something'. Again, the opinion of these coachees changed as the coaching session progressed.

Goal setting

Three coachees had difficulties with their coach's approach to goal setting. Being asked, 'What do you think we can achieve?' was for one coachee 'very difficult because I don't like to commit myself on a path.' Similarly, another 'didn't want to pre-empt the endpoint; I wanted to go on the journey and find out where the endpoint was.' When a third coachee was asked, 'If there was one thing that could change your life, your world, for the next year, what would that be?' the coachee thought, 'Wow, that's a big question; it set some expectations that I felt like I wasn't sure I lived up to.' However, overall, these three coachees all evaluated their sessions highly, sometimes achieving more than they had expected.

THROUGHOUT THE SESSION

There were six commonalities – occurring at varying times during the session – that coachees indicated contributed to their evaluations: their increase in self-awareness, the coaches' style, responsiveness, questioning, maintenance of focus and confronting or challenging. Five coachees mentioned all these factors while the sixth coachee mentioned five of the factors.

Self-awareness

Increased self-awareness was mentioned by all coachees and was described variously as 'greater awareness', 'realisation', 'insight', 'aha moment', 'self-discovery' and 'looking at myself in the mirror'. While the *kind* of experience appears similar, *what* coachees found revealing varied widely: 'patterns', the relevance of 'events in my life', 'why I've done things and why I haven't', 'getting a payoff from the problem behaviour', 'things that are not working for me' and 'having the experience of imagining myself without the hang-ups'.

Some recognised these moments as turning points; for others they contributed more generally. One coachee used words associated with self-awareness 31 times in the first interview while another realised in the two weeks after the session that even increased self-awareness 'can have a downside too; I was not aware of the limitations that those other people placed on my behaviour until I developed this heightened awareness.'

Coach style

The interviewees recognised the relationship-building qualities their coaches exhibited. There were numerous comments highlighting coach qualities. One from each interviewee is presented here:

Very engaging, natural, warm, sincere;

Saw me as a whole person;

Made me feel extremely at ease;

Comfortable with whatever was said;

Neutral throughout the whole session;

Sympathetic, not judgemental.

The effects were:

You could be yourself;

A sense of being listened to for what was going on underneath;

Not embarrassed;

Like talking to someone at a coffee shop;

Not feeling judged;

Trust and could bare your soul.

Even though one coachee '… really genuinely liked that style of coaching; better than the other stuff I got which was rubbish crap coaching', there was repeated concern that '… because [the coach] was being so neutral, I did feel like there was a lack of rapport. Maybe I did hold myself back … I don't know whether that would have changed the outcome.'

Coach responsiveness

All coachees commented on how the coach responded to key things that did (or did not) happen; four remarked when it worked well and all six remarked when it didn't. For example:

It was really pertinent to me that the coach picked up that that was really important to me and chose to go with that.

> [The coach] highlighted certain words that I wasn't paying that much attention to. She was paying a lot more attention to the way that I framed things, which basically just led me on that path to the end.

In addition, all coachees commented on their coach's *lack* of responsiveness at one time or another and found it unhelpful.

> When I'm feeling vulnerable that I'm putting myself out there … I need other people to validate that that's ok. [It] ended up feeling a bit condescending.

> She was asking me questions that I would generally want to think about before I gave an answer and I didn't feel I had the opportunity to give it as much thought.

> Sometimes I got stuck in ruts … I would have liked [it] slightly more structured in certain parts.

Questioning

All six coachees mentioned the style of questioning. Most appreciated the quality of the questions and the skill in asking those questions:

> Her questions forced me to look at myself;

> The simplicity [of] the wording, the questions, everything was quite basic; the tone was really soothing, calming, which made me respond the same way.

However, one coachee said that at times the 'questioning was too open in style' and another found it 'a bit repetitive'.

Maintaining focus

Five coachees valued the coaches' ability to maintain the purpose or focus of the session and bring their attention back to what was important. 'I had lost the link and [the coach] brought it back'. When another coachee had 'run out of steam', the coach 'somehow managed to get it back on to track'. Another talked of how in other coaching, 'they have flitted right across my life', while this coach 'kept me really glued on the one path … I found that really awesome'.

Confronting and challenging

Coachees were clearly sensitive to the degree and style of challenging since it was directly addressed by five of them, sometimes in strong terms. The very first words of two coachees in their interviews were, 'It was quite confronting' and 'Very confronting; she challenged me on just about everything I said.' They all, however, appreciated the value of the challenge. Sometimes this happened at the moment of challenge: '[It] actually stopped me there in my tracks and I had to rethink whether my behaviour actually was giving me a payoff and that's the point in the session where there was a change'. Sometimes appreciation of the challenge happened after the session. On reflection two weeks later, an interviewee said: 'I realised perhaps the questions were framed deliberately to turn the whole process back on myself, I guess akin to looking at myself in the mirror and staring at it.'

By contrast, one coachee would have liked *more* challenge. 'I stayed in story for quite a long time. And the story was a bit safe.' The coach needed 'to butt in' and 'get me out of story'.

ENDING OF SESSIONS

All six coachees had something to say about how their coach ended the session and, in particular, the setting of tasks or 'homework'. While one interviewee commented on what

worked ('I left the session having developed a strategy to be able to deal with that conflict, which was extremely beneficial') most focussed on what didn't work and what they would have preferred instead. To give two examples:

> [The coach] wanted to really nail me down. But in my heart it was still like 'I may or may not do that. I'll promise you I'll do it today but in my head I don't know that I can'.

> [The task] didn't deal with the reality, the practicality [of life]. If we [had] talked a bit more about what obstacles might get in my way for actually doing it, I would have been more clear at the time about what else I was going to do.

This matches previous findings that coachees had mixed responses to takeaway tasks. Passmore (2010, p.55) suggests that where the task was reflective, coachees expressed value in the task; more action-orientation tasks were less valued. While Passmore links the variability to management seniority, our findings suggest there may be other, more general, factors involved.

OUTCOMES AFTER THE COACHING

An outcome in coaching is an effect that happens after the session. Every coaching session in this study had beneficial outcomes and, at minimum, these were maintained for the next two weeks. By that time, five of the six coachees could also report *new* beneficial outcomes. Of the 38 outcomes described as occurring *after* the coaching session, 36 were reported as beneficial and two adverse. A selection is given in Table 1.

Beneficial outcome	Illustrative quote
In sleep and dreams	The first night [after the session] I slept like a log … which is something I don't do. … I remembered the dream vividly. [It] was about a catharsis, a cleansing, an emptying – that to me was something significant.
Reduced 'negative' feelings	[The problem] doesn't send shivers through my body any more.
Increased 'positive' feelings and thoughts	Something has shifted that has allowed me to be more confident.
Taken action	Now that I've actually done some action, that helps. I'm happy.
Increased motivation	Before it was a chore, and now it was like 'Ok, I'll wake up early and I'll do this'.
Developing new strategy	I am trying to develop strategies to deal with those changes, and that's an ongoing process.
Change in attitude	I was just much more open-minded towards what the options were.
New desired outcome	Because I have that clarity from that coaching session, other things have come up and now I want more clarity.
New behaviour	I am treating [my husband] differently.
Increased awareness	I use that awareness [from the coaching session] to help me to act rather than react.
Adverse outcome	**Illustrative quote**
Concern change would not last	By the time I walked out of [the coaching session] I had no doubt in my mind that that's locked in, [but] it's floating at the moment.
Increase in 'negative' feelings	The downside is that I have become far more aware of a lot of other things which are causing frustration.

Table 1: Some beneficial and adverse outcomes reported by interviewees as having happened *after* the coaching session.

Changes in outcomes

At the first interview, two days after the coaching, four coachees reported experiencing

beneficial outcomes, one reported both a beneficial and an adverse outcome, and one reported

that the beneficial effect of the session had been maintained with no further outcomes

noticed.

At the second follow-up interview two weeks later, three interviewees reported experiencing further beneficial outcomes and two others said that the beneficial outcomes reported two weeks earlier were maintained (although one of these had also experienced a new adverse outcome). The sixth interviewee said that there had been no change since the first interview. Table 2 summarises for each interviewee the changes in outcomes experienced after the coaching session.

Coachee	Two days after coaching	Two weeks after coaching
1	Beneficial effect of session maintained with no further outcomes	Beneficial effect of session maintained with no further outcomes
2	Beneficial outcome + Adverse outcome	More beneficial outcomes
3	Beneficial outcomes	Beneficial outcomes maintained + Adverse outcome
4	Beneficial outcomes	Beneficial outcome maintained
5	Beneficial outcomes	More beneficial outcomes
6	Beneficial outcomes	More beneficial outcomes

Table 2: Summary of changes in outcomes reported by each interviewee after the coaching.

Table 2 shows five different profiles of outcome changes among the six coachees. This diversity highlights the individuality of the responses coachees can experience in the two weeks following a coaching session.

Discussion

Interviewer as researcher

This study included processes that might have affected the outcomes (de Haan & Duckworth, 2013, p. 12). Given the second author was both interviewer and analyst, he minimized 'over involvement' due to his dual roles (Allmark *et al*., 2009, p.7) and the possible influence on coachees' responses of being interviewed (Brannick & Coghlan, 2007; Shamai, 2003) by explaining to interviewees the purpose of the interviews and sticking closely to the Clean Language interviewing protocol. At the end of the second interview, the interviewer asked each coachee to comment on the interviews themselves. Three mentioned the extra value they received over and above the coaching: 'It gave different insight into what was going on', 'It added another layer of reinforcement about what had happened in the coaching session' and 'That is probably where I really did crystallise that idea'. Clearly we could not avoid some 'interview effect', namely, that the interview inevitably prompted some new sense-making.

Issues of rigour

To check the 'cleanness' of the interviews, we invited nine experienced Clean Language practitioners and researchers, working in teams, to give each of the interviewer's questions and statements a 'cleanness rating' (classically clean, contextually clean, mildly leading or strongly leading). The tabulated results were used to arrive at an overall assessment for each interview. The reviewers found that, on average, the interviewer contributed 50 questions or statements in each interview, 40 of which were classically or contextually clean, eight mildly leading and two strongly leading. While this was not quite as 'clean' as other Clean Language-based research (Tosey *et*. *al*., 2014), the reviewers concluded that the interviews substantially adhered to the CLI protocol and were therefore fit for the purpose of this research.

The research approach and analysis were congruent with the research questions (Passmore & Fillery-Travis, 2011, p. 80). While the aim of the study was to investigate CLI as a means to explore coachees' experiences of coaching rather than to draw generalisable conclusions, our findings may have some validity or transferability for contexts of 'proximal similarity' (Polit & Beck, 2010, p. 1453), such as other goal-oriented coaching practices to which coachees bring health, career, business and relationship concerns. As a study that analyses the data according to conceptual themes, it provides some 'evidence-for-practice' despite being limited by a lack of diversity in the sample (Daly *et al.*, 2007, p. 43).

As explained earlier, the analysis and interpretation of the interviews can be considered trustworthy because of the design, protocols, adequacy of data, audit trail and repeated reflexive review of the analysis. That said, trustworthiness will be enhanced (or not) when similar studies are undertaken by other Clean Language interviewers and analysts. While longitudinal studies of three months or more may in principle be preferable, it is worth considering that the extra lapsed time may make it difficult for coachees to separate the effects of a single session from other things that happened in the meantime.

Addressing the research questions

1.   The first research question asked whether Clean Language interviewing of coachees about their experience and evaluation of coaching yields insights for coaches and trainers of coaching.

This paper reports many insights and demonstrates CLI is useful for understanding coaching through the subjective lens of the coachee (de Haan *et al.*, 2011, p. 25). This suggests that existing evaluation methods can be supplemented and extended through the use of CLI and the coding of first-person accounts into events, effects, outcomes and evaluations.

2.   The second research question asked what events, effects and outcomes – during and after a single coaching session – coachees used to evaluate coaching,

We found that coachees evaluated coaching in terms of *relational* events and effects. Many other researchers have shown that the coaching relationship is a key factor in determining how coachees perceive the outcome of coaching (de Haan & Duckworth, 2013). Relationship factors include trust, bond, engagement, collaboration and working alliance (O'Broin & Palmer, 2010, p. 139); the present study supports and fleshes out the association between working alliance and coachees' outcomes (O'Broin & Palmer, 2010, p. 137). The reports by coachees in this study confirm the importance of such relationship factors. Coachees said they were *unfavourably* affected by their coaches' behaviour when:

> Repeating coachee words back was overdone;
> The pace of the coachee was misjudged;
> Goal setting was too direct;
> Responsiveness to changes in the coachee's internal state was lacking; and
> Tasking was inappropriate.

In addition, our findings indicate that coachees' evaluation of the coaching was favourably influenced when coaches were able to:

> Display relationship qualities;
> Facilitate coachee's self-awareness and insights;
> Ask high-quality questions;
> Maintain the focus of the session;
> Respond to what was, and was not, happening; and
> Challenge and confront appropriately.

These findings confirm previous research relating to the value of increased self-awareness (Greif, 2007; Froese *et al*., 2011; Shamai, 2003), being challenged and the provision of reflective space (Seamons, 2006). It also supports the evidence that both challenge *and* support are needed for effective coaching (Passmore & Fillery-Travis, 2011, p. 81).

However, we did not find a single aspect of the coaching which every interviewee evaluated the same way. There were always exceptions –*too much* or *too little* self-awareness, *too much* or *too little* challenge, *too much* or *too little* goal setting. This suggests that while these aspects are important, so is how they applied at the time to the individual coachee. Coaches need to balance being engaging *and* being neutral, putting the coachee at ease *and* challenging them, not judging *and* confronting, having a chat at a coffee shop *and* focusing on outcomes, and so on.

In all cases, the benefits accrued during the coaching sessions were maintained or added to in the following two weeks. However, the pattern of these post-session outcomes showed no consistency (Table 2). It seems each coachee follows their own path and it would be difficult to predict, on the basis of what happened during the session, what additional changes will occur after coaching.

Implications for coach practice and training

The importance of recognising phenomena related to the beginning, middle and end stages of a coaching journey (de Haan & Nieß, 2012, p. 12) also seems to apply to a single coaching conversation. Coach trainings may need to address coaches' ability to calibrate and respond to coachees' ongoing evaluation of the coaching, the pace of the session and how the timing of coachees' feedback affects the findings.

Much of the coachees' feedback seemed to be about *calibration* and *responsiveness*. O'Broin & Palmer (2009) have documented the theoretical basis for bond, task and collaboration in the coaching alliance and the cognitive-behavioural dynamics thereof. Using skills of calibration (Linder-Pelz, 2010, p. 22; Tompkins & Lawley, 2011) coaches can, for example, get clues that the coachee wants less repeating back, wants the pace to change or has had an insight. Responsiveness is the ability to then act effectively. Together, calibration and responsiveness form a complex competency that involves noticing shifts in coachee behaviour and language, gauging his or her internal state and evaluations, creating questions related to those calibrations and tracking the direction of the session – all without allowing one's own internal commentary, interpretation or mind-reading to override the coachee's experience.

More time and attention could be allocated in coach trainings to the ending stage of coaching sessions, in particular to how to formulate takeaway tasks appreciated by the coachee. Time needs to be given to collaboratively designing these tasks.

Passmore (2010, p.49) asks what coaching psychologists need to learn to be effective. The present study suggests that they need to learn how coachees evaluate coaching and to calibrate coachees' in-the-moment responses so that their interventions are more informed by what their coachees regard as important. Coaches need to be aware that they can be considerably out of step with their coachees in terms of the pace at the beginning of the session, their degree of confrontation and their suggestions for takeaway tasks.

What happens *after* a session is a vital aspect of the value of the coaching. Given that most of the coachees in this study said they benefitted from the review of their evaluations during the follow-up interviews, coach training could support coaches to learn how to seek this kind of information more directly; this can be done during the first coaching session when coach and coachee do not know each other well, and in subsequent sessions.

Further research

Clean Language interviewing techniques could provide researchers with more detailed accounts of coachees' experiences of difficult-to-define aspects of coaching such as 'balancing challenge and support', 'stimulating problem-solving', 'effective communication', 'staying focused' (Passmore 2010) and the association between 'working alliance' and coachee outcomes (O'Broin & Palmer, 2010, p. 137).

A research methodology similar to the one described in this paper could be employed to establish coachee evaluations of second and subsequent coaching sessions. It could also be used in a longitudinal study to look at how coachees perceive the outcomes of coaching after a series of coaching sessions has ended.

In addition, the use of video coupled with coachee accounts of a session could shed light on the clues that a session is or is not working well for the coachee. In a similar way, *coaches'* experience of the coaching process could be investigated.

While further study is needed, the potential of CLI as a method for evaluating coachees' experience has been demonstrated. Ongoing qualitative studies like this will enable the coaching profession to learn from exceptions to the coaching experience as well as from averages and generalisation.

Bibliography

Allmark, P., Boote, J., Chambers, E., Clarke, A., McDonnell, A., Thompson, A. & Tod, A. M. (2009). Ethical issues in the use of in-depth interviews: Literature review and discussion. *Research Ethics Review, 5*, 48-54.

Brannick, T. & Coghlan, D. (2007). In defense of being "native": The case for insider academic research. *Organizational Research Methods, 10*(1), 59-74.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Daly, J., Willis, K., Small, R., Green, J., Welch, N., Kealy, M. & Hughes, E. (2007). A hierarchy of evidence for assessing qualitative health research. *Journal of Clinical Epidemiology, 60*(1), 43-49.

de Haan, E., Bertie, C., Day, A. & Sills, C. (2010). Critical moments of clients of coaching: Towards a 'client model' of executive coaching. *Academy of Management Learning & Education, 9.4*, 1-31.

de Haan, E., Culpin, V. & Curd, J. (2011). Executive coaching in practice: What determines helpfulness for clients of coaching? *Personnel Review, 40*(1), 24-44.

de Haan, E. & Duckworth, A. (2013). Signalling a new trend in executive coaching outcome research. *International Coaching Psychology Review, 8*(1), 6-19.

de Haan, E. & Nieß, C. (2012). Critical moments in a coaching case study: Illustration of a process research model. *Consulting Psychology Journal: Practice and Research, 64* (3), 198-224.

Froese, T., Gould, C. & Barrett, A. (2011). Re-viewing from within: A commentary on first-and second-person methods in the science of consciousness. *Constructivist Foundations*, *6*(2), 254-269.

Grant, A. M. (2013). The efficacy of executive coaching in times of organisational change. *Journal of Change Management*, (ahead-of-print), 1-23.

Greif, S. (2007). Advances in research on coaching outcomes. *International Coaching Psychology Review, 2*(3), 222-249.

Grove, D. J. & Panzer, B. I. (1989). *Resolving traumatic memories: Metaphors and symbols in psychotherapy*. New York: Irvington Publishers Inc.

Gyllensten, K. & Palmer, S. (2007). The coaching relationship: An interpretive phenomenological analysis. *International Coaching Psychology Review, 2*(2), 168-177.

Hall, L. M. (2011). *Coaching mastery: Meta-Coaching module III, 2011 edition, certification training for ACMC credentials*. Grand Junction, CO: Neuro-Semantic Publications.

Hall, L. M. (2013, Jan 9). Evaluative vagueness in the coaching context. Message posted to E-Group for Certified Meta-Coaches.

Hieker, C. & Huffington, C. (2006). Reflexive questions in a coaching psychology context. *International Coaching Psychology Review*, *1*(2), 46-55.

Knox, S. & Burkard, A. W. (2009). Qualitative research interviews. *Psychotherapy Research, 19*(4-5), 566-575.

Lawley, J. & Linder-Pelz, S. (2014). 'Evidence-based' on what? Exploring coach, coachee and expert evaluations of coaching. Manuscript submitted for publication.

Lawley, J. & Tompkins, P. (2000). *Metaphors in mind: Transformation through Symbolic Modelling*. London: The Developing Company Press.

Lawley, J. & Tompkins, P. (2011). Clean Evaluative Interviewing. www.cleanlanguage.co.uk/articles/articles/317/ retrieved 16 November 2014.

Linder-Pelz, S. (2010). *NLP coaching: An evidence-based approach for coaches, leaders and individuals*. London: Kogan Page Limited.

Linder-Pelz, S. (2014). Steps towards the benchmarking of coaches' skills. *International Journal of Evidence Based Coaching and Mentoring, 12*(1), 47-62.

Linder-Pelz, S. & Hall, L. M. (2008). Meta-Coaching: A methodology grounded in psychological theory. *International Journal of Evidence Based Coaching and Mentoring*, *6*(1), 43-56.

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7*, 560–572.

Mays, N. & Pope, C. (1995). Rigour and qualitative research. *British Medical Journal, 311*, 109-112.

O'Broin, A. & Palmer, S. (2009). Co-creating an optimal coaching alliance: A cognitive behavioural coaching perspective. *International Coaching Psychology Review*, *4*(2), 184-194.

O'Broin, A. & Palmer, S. (2010). Exploring key aspects in the formation of coaching relationships: Initial indicators from the perspective of the coachee and the coach. *Coaching: An International Journal of Theory, Research and Practice, 3*(2), 124-143.

Owen, I. R. (1989). Beyond Carl Rogers: the work of David Grove *Journal of Interprofessional Care, 4*, 186–196.

Passmore, J. (2010). A grounded theory study of the coachee experience: The implications for training and practice in coaching psychology. *International Coaching Psychology Review, 5*(1), 48-62.

Passmore, J. & Fillery-Travis, A. (2011). A critical review of executive coaching research: A decade of progress and what's to come. *Coaching: An International Journal of Theory, Research and Practice, 4*(2), 70-88.

Polit, D. F. & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies, 47*, 1451–1458.

Seamons, B. L. (2006). *The most effective factors in executive coaching engagements according to the coach, the client, and the client's boss*. (PhD), Saybrook Graduate School and Research Center, California.   (UMI Dissertations Publishing 3206219).

Shamai, M. (2003). Therapeutic effects of qualitative research: Reconstructing the experience of treatment as a by-product of qualitative evaluation. *The Social Service Review, 77*(3), 455-467.

Sousa, D. (2014). Validation in Qualitative Research: General Aspects and Specificities of the Descriptive Phenomenological Method. *Qualitative Research in Psychology,* 11:211–227.

Thibodeau, P. H. & Boroditsky, L. (2011). Metaphors We Think With: The Role of Metaphor in Reasoning. *PLoS ONE, 6*(2), 16782.doi: 10.1371/Journal.

Tompkins, P. & Lawley, J. (2011). Calibrating whether what you are doing is working – or not. www.cleanlanguage.co.uk/articles/articles/308/, retrieved 16 November 2014.

Tosey, P. (2011). 'Symbolic Modelling' as an innovative phenomenological method in HRD research: the work-life balance project. Paper presented at the 12th International Conference on HRD Practice and Research across Europe, University of Gloustershire, 25-27 May.

Tosey, P. Lawley, J., Meese, R. (2014). Eliciting metaphor through Clean Language: An innovation in qualitative research. *British Journal of Management, 25*(3), 629-646.

Tosey, P. & Mathison, J. (2010). Exploring inner landscapes through psychophenomenology: The contribution of neuro-linguistic programming to innovations in researching first person experience. *Qualitative Research in Organizations and Management, 5*(1), 63-82.

van Helsdingen, A. & Lawley, J. (2012). Modelling Shared Reality: avoiding unintended influence in qualitative research. *Kwalon: Journal of the Netherlands Association for Qualitative Research, 1*(3), www.cleanlanguage.co.uk/articles/articles/328/, retrieved 16 November 2014.

Vansickel-Peterson, D. L. (2010). *Coaching efficacy with academic leaders: A phenomenological investigation*. Open Access Theses and Dissertations from the College of Education and Human Sciences. Paper 87. University of Nebraska, Lincoln. Retrieved from http://digitalcommons.unl.edu/cehsdiss/87, retrieved 16 November 2014.

Worth, S-A. (2012). *An exploration of coaching women towards authenticity in the workplace: A heuristic study with women in academia*. (DCM), Oxford Brookes University. Retrieved from www.brookes.ac.uk/go/radar. Available from Oxford Brookes University Research Archive and Digital Asset Repository database.

Appendix: Methodology of present study compared with Braun & Clarke's phases of Thematic Analysis (2006)

Braun & Clarke's Thematic Analysis 2006, p 87	Analysis of Clean Language Interviews
1. **Familiarise yourself with data.**	For each interview: Read and reread to get general sense of *each* interviewee's perspective. Noted how individuals expressed themselves and organised their experience, especially through metaphor.
2. **Generate initial codes.** Coding interesting features in a systematic fashion across the entire data set, collating data relevant to each code.	For each interview: Modelled the sentences of each interviewee *separately*, highlighting events, effects, outcomes and 'favourable' or 'unfavourable' evaluation statements (Tompkins and Lawley 2010, Lawley & Tompkins, 2011).

3. **Search for themes.** Collating codes into potential ("candidate") themes, gathering all data relevant to each potential theme.	For each interview:
	a. Co-located 'data extracts' into clusters *for each interviewee* by:
	  - repetition (same words)
	  - restatement (same idea said in a different way)
	  - inherent logic (e.g. cause-effect)
	  - sequence (when what happened)
	b. Added remainder of examples to 'unallocated cluster'.
	c. Allocated 'candidate theme' names for clusters

4a. **Review Themes - Level 1.** Checking if the themes work in relation to the coded extracts	Across all data extracts: a. Co-located examples of similar themes (retain source of each quote). b. Reviewed examples in 'unallocated clusters' for new clusters or allocated to an existing cluster. c. Retained remainder of examples in global 'unallocated cluster'.
4b. **Review Themes – Level 2**	Across themes: a. Reviewed themes for 'internal homogeneity' – removed examples that no longer 'fit' and reallocated. b. Reviewed themes for 'external heterogeneity' – combined or split clusters that covered the same or different themes. c. Reviewed *interviews* to check original context and look for missed or misallocated data items.
5. **Define and name themes:** Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme.	Across themes: a. Organised implicit structure inherent (e.g. the five time-frames of coachee evaluations) b. Defined 'prevalence' criteria for an 'acceptable theme' (e.g. example from at least 3 of the 6 interviewees). c. Settled on names for selected themes.
6. **Produce Report.**	Wrote up results with reference to research questions.