

Steps towards the Benchmarking of Coaches' Skills

Susie Linder-Pelz, Edgecliff NSW 2027, Australia

Contact Email: susie@gooddecisions.com.au

Abstract

This paper describes a study of the development and implementation of a benchmarking methodology within a coaching community. It asks: how robust and valid was the process used to identify the key coaching skills and their specific behavioural indicators? It finds that the initial identification of core coach skills was driven by a literature review and by modelling based on Neuro-linguistic Programming (NLP). The ongoing collaborative testing and refinement of the benchmarking meets established criteria of valid action research. This approach to assessing coach competencies is based on making fine-grained behavioural distinctions, uses direct observation and is used both for training new coaches and maintaining practising coaches' skills. Further research is needed with regard to inter-rater reliability, outcomes and coach supervision.

Keywords: Benchmarking, competencies, coach training, NLP, Meta-Coaching

Introduction

Benchmarking is about continuous improvement and learning about the processes that enable unusual performance excellence (Weeks, 2000). Benchmarking competencies for training and maintaining practice standards is increasingly recognised in mature professions such as nursing (Chaboyer et al., 2004) and human resource development (Rodriguez et al., 2010; Olian et al., 1998).

Competencies result from integrative learning experiences in which knowledge, skills and abilities interact; they are a blend of skills, abilities and knowledge needed to perform a specific task (Foss et al., 2004; Cowan et al., 2007). There is a view that “one size fits all” is not appropriate when it comes to method of competency assessment (Butler et al., 2011, p.298); competencies can be assessed through self-report measures, oral case presentations, behaviour-based interviews as well as live and video demonstrations of applying new knowledge and skills. There are also standardised, psychometrically sound, objective measures of behaviour, using various observer-rating inventories (Perosa & Perosa, 2010).

Several commentators have noted that professional training has a natural connection with competency-based education because of the clear relationship between learners' performance and work force expectations (Foss et al., 2004, p. 369, citing Voorhees, 2001). Competency-based learning allows for flexibility, reduction of duplicity, and building on previous knowledge (Foss et al., 2004).

However, some critics of competency-based education are concerned that it is reductionist and teaches to the task rather than developing critical thinking (Foss et al., 2004, p. 369). There is also a view that it is best to abandon attempts to assess the quality of less tangible qualities—such as professionalism or interpersonal interaction—by standards-based approaches to the assessment of competence (Gallagher, 2012). Others advocate the empowerment of trainees in adult and

The current issue and full text archive of this journal is available at
<http://business.brookes.ac.uk/ijebcm>

professional education (Freire, 1970; Mezirow, 1978) and there is widespread support for various forms of participative assessment in adult and professional education (Reynolds & Trehan, 2000). For some who hold this view, peer assessment is a type of ‘formative assessment’ where students are learning from each other; this is seen as an easier and more accessible way of learning than through someone in a position of authority. (Burns & Gillon, 2011).

The drive to professionalise coaching and to benchmark competencies parallels developments in other fields and Grant point out that: “the identification of the key competencies of effective coaches is important for coach education and training [and] there have been a number of attempts to place these behaviours, skills and attributes within a competency model” (Grant et al., 2010, pp. 28-29). There is much discussion about the skills and competencies valued by coachees (Grant, 2011), the need for trainers to understand which coach behaviours have impact and how they interact (Passmore & Gibbes, 2007), as well as ethical guidelines and benchmarks for best practice, including the use of professional supervision (Linley, 2006; Grant, 2012). Among those who have identified key coaching skills are Griffith and Campbell (2008), Passmore (2008) and Grant & Cavanagh (2007): the latter also measured coaching skills as “a means of benchmarking” those skills (Grant & Cavanagh, 2007, p.752).

In 2003, Michael Hall and Michelle Duval started to identify core coaching skills and develop behavioural benchmarks for coaches (Hall & Duval, 2004). Their approach was based on NLP. Most coaches and psychologists know—or have heard of—NLP as a set of change techniques. However, the original purpose of NLP was to explore the structure of experience, using language and other calibrating skills to gather intra-personal information. This is called *modelling* and it enables coaches to track, map and make explicit the patterns and sequences of internal representations of information coming in through a person’s senses and memory and by which s/he makes meaning, creates his/her experience and takes action (or not). The key modelling question is ‘how do you/they do it?’ There are many ways to answer that, in terms of the levels of detail that the NLP modeller tracks: skills, physiology, behaviour, beliefs, concepts, identity and symbols. The key to modelling is the meta-model of precision questioning (Grinder and McMaster, 1993), which enables the modeller to recognise details from a ‘meta’ position (Hall, 2011).

One of the applications of NLP modelling is as a type of benchmarking (Dilts, 1998) and “both benchmarking and modelling, as investigative processes, inquire into the inner structured processes that explain a given experience” (Hall 2011b, p.54). So NLP modelling may be considered a naturalistic and inductive inquiry method for identifying competencies (Smith, 2009).

Hall’s idea of benchmarking ‘intangibles’, such as coaches’ skills, was also based on Maslow’s (1968) view that “we ought to keep seeking objective correlations or indicators of subjective states” (Hall 2011, p.14). Hall wanted to drill down into sensory-specific and behavioural descriptions of coaching skills. He recalls the beginning in 2002:

... Meta-Coaching arose as I modeled four expert coaches, two Executive Coaches [...], one Personal Coach [...] and one Group Coach [...]. I modeled from them the core competencies of a professional coach which you know as the Seven Core Competencies of Meta-Coaching. What I found in interviewing and watching and modeling from them was the innate processes that allow deep connection, intimate exploration, confrontative mirroring, and experiential feeling to occur. [...]

Since 2002 I have modeled 27 other experts in the field of coaching and so the models of Meta-Coaching continue to grow, evolve, and develop. And while these experts mostly come from a NLP background, some are from many other Schools of Coaching. This is the powerful facet of modeling excellence, identifying the best and the best of the best as processes and then developing processes for transferring those skills. (Hall, 2011e, p.102).

The Meta-Coaching Foundation (MCF) and the first level of coach training, the Associate Certified Meta-Coach training (ACMC), are described in Appendix A.

Benchmarking, as a skill, enables us to operationalize our terms, ‘denominalize’ verbs-turned- into-nouns [eg ‘rapport’] and generate a set of see-hear-feel variables that we can use to determine and measure our skill level and give feedback against (Hall & Duval, 2004). Benchmarking can also be described as the ability to:

- identify expert or best practice of a skill/quality
- demonstrate the best expression of that skill/quality
- model the component variables of the skill/quality
- create a scale from the lowest to highest (best) expression of that skill/quality
- distinguish very specific see-hear-feel indicators at each point of skill/quality development on the scale
- use that scale to give and receive sensory-based feedback for learning and development
- communicate this process with precision (Hall, 2011, p15 & p.272).

Knowing a little about how Hall developed his coaching models, I wondered if the development of the coaching benchmarks and the benchmarking methodology can be described and evaluated through the lens—frames, practices and criteria of validity—of ‘action research’. The aim was to evaluate Hall’s benchmarking process with some objectivity and compare it with other approaches in the field of coaching. The research question was: *how robust and trustworthy was the process Hall used in identifying the core coaching skills and deriving their specific behavioural indicators?* Although I am a trained Meta-Coach, I stepped back from coaching practice in 2005 to focus on research and was not actively involved in the MCF community.

Methodology

In order to systematically explore, describe and evaluate of the process of developing the benchmarking methodology, I analysed documents and conducted semi-structured interviews with a purposive sample of members the Meta-Coaching community. The study had an explicit theoretical basis, namely, an essentialist/realist approach which looks at motivations, experience and meaning in a straightforward way, because it assumes that “language reflects and enables us to articulate meaning and experience” (Kelly, 2004, p. 474). A particular strength of this methodology is that “through providing detailed description of what happens when interventions are carried out [...] we begin to look more closely at how things work (or do not work)” (Kelly, 2004, p. 470).

I re-read Hall’s writings from 2003-2011, observed a coach training and attended some MCF chapter meetings and then prepared a set of open-ended questions about the development, testing and use of coaching benchmarks in training and certifying Meta-Coaches (Appendix 2). Given the availability of Hall’s writings and his constant travel commitments, he responded in writing to a few specific questions (Hall, personal communication, December 15, 2011). His early collaborator, Michelle Duval, was also invited to check historical details (personal communication, April 26, 2012).

A purposive sample (Gobo, 2004) of nine Meta-Coaches was selected. The rationale for selecting three of these interviewees was their interest and engagement with research and evidence-based coaching in the Meta-Coaching community. Five other interviewees were long-term members of the MCF Leadership Team who had not only been involved in the development and testing of these NLP-based coaching benchmarks by being trainers for several years but also had used the benchmarking methodology in additional ways in their own practices. Another interviewee had been active early in the benchmarking project but had later chosen to disengage from the Meta-Coaching community. Interviewees were from seven countries; all had attended at least two ACMC trainings, six had been

certified Meta-Coaches for five or more years. Two had not been on Assist Teams, meaning they had not benchmarked others. The interviews (averaging 54 minutes) were recorded and transcribed.

The written materials and interview transcriptions were analysed deductively, using Ladkin's (2004) criteria of action research. This involved several iterations of reading Hall's written materials and linking evidence from this written evidence with interviewees' transcribed comments. When the analysis was complete, interviewees received the draft report and a transcript of their own interview for member checking and giving feedback. They were invited to comment on accuracy, themes, conclusions and anything else that came to mind as they read the report. Findings of these interviews would be indicative of Meta-Coach trainers' and trainees' experiences and not generalisable to all trainers and trainees.

In the next section the findings are presented as a chronology of the steps taken in the development of the benchmarks.

Findings

Hall first modelled Duval's live and recorded coaching sessions and identified over 120 questions she used without her making any comments or statements (Hall and Duval, 2003/2004, pp. 41-43). An interviewee (G) recalled:

They modelled out some of the competencies that Michelle Duval believed were absolutely essential to have as a coach. And then they started using NLP, and the communications model and the Meta-Model for specific language to start to get sensory-based information. From there they realised there are sub-skills to the benchmarks, the behaviours of the skill.

Hall's approach to benchmarking intangible coaching skills was very explicitly based on his own Meta-States Model, which explains how abstract knowledge is created out of empirical, see-hear-feel knowledge. He modelled not only behaviours but also meaning-making, since his earlier work had demonstrated how it is possible to model not only behaviours but also self-reflexivity, beliefs and abstract knowledge (Hall 1995/2000; Linder-Pelz & Hall 2008; Hall, 2011). This modelling of both meaning and performance follows Gallwey's (1974) distinction between the 'inner game' and 'outer game' of coaching as well as Maslow's view that self-actualisation involves both *doing* and *being* (Hall 2011, p.160). Thus Hall added the 'vertical modelling' of meta-level frames and states to the original linear or horizontal modelling of NLP (Linder-Pelz, 2010). Modelling may be considered an inductive inquiry method useful in identifying the competencies that define the exemplary role (Smith, 2009).

Hall and Duval (2003/2004) reviewed the literature on benchmarking in businesses as well as on the core competencies required for a professional coach by the International Coach Federation (ICF) at that time. That led them to identify five basic skills: listening, rapport, questioning, inducing states and giving feedback. Hall's modelling led him to add two more core competencies: meta-questioning and receiving feedback (email to author 8/8/11; Hall et al., 2004/2011).

In April 2003 Hall and Duval started asking, "What are the behavioural equivalents of these core competencies?" (Hall 2011, pp. 169-181). Drawing from the literature on benchmarking in organisations, Hall and Duval posed specific questions such as:

- How can we establish a valid and practical way to determine if a person is actually skilled as a Meta-Coach? (Hall 2011, p.5 & p.24)
- How can we model best competencies, identify the component parts, the critical factors that facilitate change? (Hall 2011, p. 5)

- What processes can we use to set benchmarks on the skills of coaching, so that we can identify best practices in coaching and create training that replicates that expertise?" (Hall & Duval 2004, pp. 131-153).

Table 1 shows these seven core skills. The first sentence alongside each of the skills refers to the states that an effective coach will be operating from in order to perform that skill, while the second sentence refers to the intended effect on the coachee. Appendix 2 gives the example of the benchmarks for listening. Benchmarks for the other six core skills and 23 more advanced coaching skills for developmental coaching—whether in organisational, executive or life coaching contexts—appear in Hall, 2011.

1)	<i>Support</i>	Warm and caring — while supporting coachee.	So coachee feels cared for and believed in by the coach.
2)	<i>Listening</i>	Compassionate attention to the person — while listening to coachee.	So coachee feels deeply understood and 'heard'
3)	<i>Questioning</i>	Respectfully curious — while questioning coachee.	So coachee feels the questions are respectful, in service of actualizing his or her highest and best.
4)	<i>Meta-Questioning</i>	Exposing empowerment — while meta-questioning the higher frames of meaning.	So coachee feels led on an inward journey of discovery, expanded self-awareness and a safe vulnerability to the coach.
5)	<i>Receiving Feedback</i>	Respectful reflective openness — while receiving feedback from coachee during the session.	So coachee feels safe to disagree, to correct the coach and able to express his or her self in a completely open way.
6)	<i>Giving Feedback</i>	Fierce succinct compassion — while giving feedback to coachee in a session.	So coachee feels seen, exposed and fascinated by the mirroring of the coach.
7)	<i>Inducing States</i>	Expressively persuasive — while inducing a coachee into states.	So coachee feels connected, open, respected, understood, safe and curious, engaged, responsible, appreciated, motivated, solution focused.

Table 1: Seven Core Coaching Skills (Hall 2011a, Reproduced with permission)

From 2003, Hall and colleagues refined the behaviours indicating the different levels of competency for each skill; they did this at every ACMC training and in MCF chapter meetings where trainers and trainees do live coaching sessions followed by debriefing of the coachee and getting feedback from other coaches and trainees who were present. In doing so, they tested Hall's hypothesis that the more often one observes the behaviors of competency in a coach, the more likely a coachee will evaluate the coaching to be successful in facilitating their goals for coaching (Hall 2010; Hall 2011; Hall 2011c and Hall, personal communication, January 12, 2012). Interviewee D explained:

Michelle and Michael developed the benchmarking model... and the question was, how do we get more specific, because if you can't measure skills then how do we improve? So as they got

more and more specific in fully mapping out what the skills were, a couple of things happened. One is that in terms of the scale, it is quite obvious what 'not listening' is; it's interrupting. But what is less obvious is what's level five? Level five was what Michael was seeing from modelling [expert coaches] and he'd describe that. And then the question was: we have the two polarities but what's half way? What are the other steps and stages [in developing the skill]? Michael (and Michelle initially) was training Meta-Coaching on a consistent basis, so they got a chance to review and refine this continuum.... So there was a constant revision and refining process based on his experience and the feedback from the people on the courses.

Hall has stated that this is the “measurement question in coaching” (Hall 2011, pp. 5-6). From 2004, Hall had referred to Maslow (1968), who had sought “objective correlations or indicators of subjective states” (Hall & Duval, 2004, p.132).

Describing how Assist Teams systematically benchmarked trainee coaches in all trainings and gave them feedback, interviewee G said:

Typically, there's one Assist Team leader to every three participants on the training. And so their job is to monitor their team, benchmark, take the score of each day, report that back. That gets reported back each morning to the Leadership Team, monitoring how people are going, what's working, what do they need help with. And then the final morning, the Leadership Team confers on who's nowhere near, who's close, how close are they and what do they need today.

Interviewee H reported on how participants reflect on developments and communicate them to the community of researchers:

Often after some of the trainings, the trainers have added to the model... I know a few years ago we had an extra day with Michael and we just spent it on the benchmarking and Michael was refining the project. While he was going through it, the whole group, probably 15 of us, were going through the benchmarking model with him.

Over the years 2006-2011, Hall conferred frequently with four particular Meta-Coach trainers and the various groups of Team Leaders in ACMC trainings (Hall, personal communication, December 15, 2011). In every training he spent one or two days with the teams, watching them do sessions, talking them through their findings and refining the benchmarking. For example, he found that in different language communities (Spanish, French, Chinese), different words for the benchmarks carry different connotations.

According to interviewee D, “*there is constant revision and refining based on Michael's experience and feedback*”, and E described how Hall “*focuses mainly on the way the Assist Team is addressing problems, addressing difficulties, benchmarking the people; he asks for examples, he does demos, he asks for role plays. He's always more than correcting ... I notice a constant reflection in order to reorganize; that is one of the sources of inspiration Michael uses for writing.*”

Hall constantly asked questions—for himself and for the listener to reflect on and respond to. In his books he also often sets tasks for the reflection, learning and action of the Meta-Coaching community (e.g. Hall & Duval, 2004, pp.157-162). ‘Reflections’ are posted regularly on the e-group.

Michael is very open to learning and many times he is just reflecting aloud and asking me to be a mirror, to be a bouncing board. He believes a lot on working as a team. In those moments, he thinks aloud, I ask questions and we both have a lot of things that come up. They are not well formed, just ideas around the way to benchmark... It's like building things through reflection and through memories of the experience.” (E)

The term 'team' is widely used in the MCF community: 'assist team', 'leadership team', 'team of benchmarkers'. With several colleagues, Hall decided on a particular benchmarking scale (Hall, 2011, p.5) and others contributed to testing and refining the measurement tool and to scaling distinctions (Hall 2009, pp. 255-276; Hall, 2011, p. 43). A few, including interviewees in this study, were in effect partial co-investigators in the sense of giving feedback, asking questions and helping generate next steps in training to those benchmarks. One interviewee recalled:

I was instrumental in developing the scaling approach, in asking "how can we measure this?" I brought the triad work with marking sheets, as in counselling, which involved marking [a candidate] on a scale and giving written examples. I was also involved in breaking them down to behaviours and looking at these behaviours as a sequence, based on my experience with the key skills process in UK... Feedback from assist teams was well received by Michael; there was to-ing and fro-ing and he had the final say... In the early days of the scaling approach, we were developing paperwork through feedback to Michael at the daily debriefs (Interviewee I).

Interviewee G talked about the benchmarking project in the context of the MCF community:

We are a community that's developing a community practice. So we've got the networks, we've got the chapters, we've got several worldwide e-groups that have voice, putting out information and connecting with the students during and post the training. Then we've got an international leadership team member in each country that has the responsibility to maintain contact and support for anybody that would like it. And then below that, there's a regional director for each city that we've delivered Meta-Coaching in and their job is to be in contact with those people that have got through and those people that need additional support. So it's not necessarily the most robust system in the world, but it is a community... There's a lot of support, a lot of connection.

The next section considers whether the process of developing Meta-Coaching benchmarks can be considered valid 'action research' and compares Hall's approach to identifying and assessing coaching competencies to other approaches. Issues of reliability and fairness in peer assessment of competencies are discussed and further research is suggested.

Discussion

The validity of the action research

To evaluate the development of benchmarking led by Hall, I considered how well it meets the 12 criteria of action research which derive from Ladkin's (2004) review of the theoretical and research literature and accounts of researchers at University of Bath Centre for Action Research in Professional Practice.

1. Intention that there will be practical consequences

Hall's intention from 2004 on was to create a system as objective as possible for recognising the level of competency of those learning coaching skills and to set up a process for continual improvement of skills and expertise (Hall, 2011, p.192).

2. Defining the problem clearly

Hall clearly stated the problem as the "measurement question in coaching." (Hall, 2011, pp. 5-6).

3. Specifying the plan of action, including 1st, 2nd & 3rd person inquiry

Hall recorded his model of working which included reading and reflecting on his learnings (1st-person inquiry). Evidence of 3rd person inquiry is how the entire MCF membership and e-group, trainers and Assist Teams on APMC trainings took part in the sense-making process by being subjects of benchmarking and giving feedback. There was face-to-face small group work; exercises;

leadership team discussions, one-to-one conversations and e-group exchanges. Hall also had Meta-Coach trainees and team leaders learn and comment on designing and developing a benchmarking project (Hall, 2011, pp.182-190).

4. Testing hypotheses by applying action to the problem

Trainers and Assist Teams were involved in the testing and refining of the skills measurement tool and scaling distinctions (2nd person inquiry). As interviewee E said, “*I can recall many times at Starbucks at the airport very early in the morning just reflecting on the benchmarking and the new insights that pop up because of our observations and experiences.*”

5. Evaluating by monitoring effectiveness of action taken

Interviewees often used words such as ‘structure’, ‘standardisation’, ‘tighter and tighter standards’ to describe the benchmarking and giving of feedback. Interviewee D commented: “*The pool of top coaches is growing and the feedback is getting better... There are more skilful people giving feedback and being benchmarked. And that’s why the distinctions are getting better.*”

6. Participants reflect on, and explain developments and then communicate results to community of researchers

All interviewees were clear that, while the development of the benchmarks was mostly Hall’s work, the refining and training to those benchmarks has involved the practitioner community. At trainings there were frequent formal and informal meetings between the Assist Team, the Leadership Team and Hall.

7. Action research is underpinned by specific beliefs and assumptions which Ladkin (2004) specifies and which are italicised in the next six paragraphs.

The primary purpose of action research is to *produce practical knowledge* that is useful to people in the everyday conduct of their lives. Historically, the core concern of NLP practice has been to do whatever gets results for coachees (Linder-Pelz, 2010, p.103) and the benchmarking project reflects this purpose.

All interviewees attested that the process of developing the benchmarks had been very *participative*, though two made statements suggesting they felt it wasn’t truly *democratic*.

The constructivist assumptions that underpin NLP imply that *all observations are biased* and that in any situation there is a plurality of meanings and interpretations (Linder-Pelz, 2010, p.81). ‘Perceptual positioning’ is the way NLP practitioners work with *multiple ways of understanding* (Linder-Pelz 2010, pp. 29-30); in this benchmarking project, this is achieved by the use of triads where there is always a benchmarker, a coachee and an observer who notes and gives sensory-based feedback to the benchmarker.

Taking into account the *emotional, social and spiritual dimensions* of those with whom the researchers interact also accords with NLP. However, *political influences* are not the focus in NLP and were never considered in this project.

Unpicking and unearthing the frames was achieved by the use of Meta-Model language. This means, for example, challenging a coachee’s assumption about cause and effect, inviting the coachee to be more specific and precise by turning static and vague nouns into verbs (‘denominalising’) and translating a principle into something more concrete or precise. Having an awareness of coachees’ and benchmarkers’ perceptual filters and thinking styles also helps unpick hidden frames of mind (Hall, 2011).

Apart from being benchmarked, candidates in ACMC trainings learn how to benchmark skills (Hall 2011); this is evidence of *learning and evolving*. It is also how coach trainers and trainees come to *see themselves as part of the research frame/situation*.

8. *Using accounts and recordings of practice*

Beginning in 2003, Hall modelled live and recorded coaching sessions. Since then, there have been many additional examples in his writings of *using accounts and recordings of practice as raw data* as the way the coaching skills and benchmarks were identified.

9. *Rigorous reflection and experimentation with new understandings or behaviours*

This is seen in the constant revision and refining based on systematic feedback. Other examples of rigour are the 'discovery process' Hall used and the 22 revisions of the benchmarking feedback form between 2003 and 2011.

10. *Undertaking cycles of action and reflection leading to knowledge and creation of theory*

Hall is explicit about his process of reading, observing, reflecting and writing. The constant engagement in reflecting on, reading and trialling his models around the world combines 1st person, 2nd person and 3rd person inquiry. Outcomes have been the explication of a facilitation model (Hall, 2010, Ch.12) as well as the model for benchmarking of intangible skills, qualities and needs (Hall, 2011: Ch. 3, 14).

11. *Collaborating so that those involved in the research are co-subjects and co-researchers whose thinking and decision-making contribute to generating ideas, designing and managing the project and drawing conclusions*

For more than seven years collaboration has been the key to all the modelling, listing of behaviours for core competencies, getting information and feedback about coaches' performance, and testing the benchmarking/distinctions (Hall, 2011, pp.5-6). Hall has documented the substantial contributions of community members (Hall, 2009, pp. 255-276; Hall & Duval, 2003/2004; Hall & Duval, 2004; Hall, 2011, p.43).

12. *Awareness of emergent processes*

Hall alludes frequently to 'emergent process' (e.g. Hall, 2011: Ch. 2, 4, 6). He describes his personal experience of this work in terms of "*messiness, fog, dealing with the unexpected and setbacks*" and of not knowing in 2003 where/what would emerge over the next seven years (Hall 2011, pp. 2-12). "*The challenge for benchmarking the intangibles of humans is the complexity of our mind-body-emotion system.... It is out of that complexity that emergent properties arise...*" (Hall & Duval. 2004, p.135, emphasis in original). Interviewees also alluded to emergence: "*We are finding more pieces to the puzzle [and] nailing down the detail in practice and supervision sessions*" (Interviewee D).

13. *Finding presentational form and going public with transparency*

For Hall, writing is central to action research; it includes book publication, weekly e-group posts and annual revisions to the coaching manual. "*He's all the time observing then reflecting then writing and when he writes, he knows. I have heard him say 'When I write I know how much I know.'*" (Interviewee E).

According to Ladkin (2004, p. 481), the *validity or trustworthiness* of action research is evidenced by:

- demonstrating emergence and enduring consequences
- dealing with pragmatic issues
- demonstrating quality of relational practice such as democracy and collaboration
- dealing with questions of significance
- taking into account different ways of knowing.

According to these criteria, it is reasonable to conclude that the development of NLP-based benchmarks of coaching competency has been robust and valid. It contributes to development of competency-based assessment in coaching which, as Voorhees (2002) points out, requires the development of three distinct, interactive components: a description of the competency, a means of measuring or assessing that competency and a standard by which someone is judged to be competent.

Other approaches to identifying and assessing coach competencies

For Grant & Cavanagh (2007), the validity of core goal-focused coaching skills was first evidenced in theoretical and empirical literature and followed by a study to validate a self-report measure of these five core competencies, looking at face validity, discriminant validity, convergent validity and test-retest reliability. Griffiths & Campbell's (2008) approach to validating the 11 core ICF competencies was through a qualitative grounded theory study of the reported/remembered and self-reported experiences of ICF coaches and their clients.

Hall also began with competencies mentioned in the literature and 'denominalised' them by modelling expert coaches. With the resulting list of sensory and behavioural indicators, rather than evaluative statements and recollections, this understanding of coach competencies can be considered more fine-grained. The benchmarking is done during observations of actual coaching, using a detailed checklist and rating protocol that is intended to reduce observer bias. This differs from self-reporting (Grant & Cavanagh, 2007) or self-assessment of competencies (Amin & Amin, 2003). Although the ICF has developed a detailed checklist for assessing competencies, some appear to rely on interpretations or evaluations more than on what the rater observed or heard directly (ICF 2009; ICF 2009a).

Reliability and fairness

While numerous studies show that peer assessment is beneficial for learning (Welsh, 2007; Khabiri et al., 2011; Rush et al., 2012), there are widespread concerns around observer bias and fairness in peer assessments (Welsh, 2007; Reynolds & Trehan, 2000). Despite the checklist of detailed, behavioural indicators used in Meta-Coach benchmarking, there is still much variability in the benchmarking which can be explained in terms of individual styles (Vikerman, 2009; Linder-Pelz, forthcoming). One way to improve inter-observer reliability is through providing even more training of peer assessors and benchmarkers (Dochy et al., 1999). In this regard it is worth noting that "variations in judgements are one reason rating scales may have a lower reliability than checklists [...]" (Bashook, 2005, p. 577). Further independent study is needed regarding the inter-observer reliability the benchmarking method used by the MCF.

Further research

The assessment challenge, according to Bashook (2005) is to "develop and use valid and reliable assessment methods...Validity refers to the accumulated evidence about how well an assessment of competencies measures what it is intended to measure" (pp. 564-566). Even with the conclusion that Hall's benchmarking scales and method do differentiate levels of coach competency and can be considered trustworthy and useful for coach training and selection, further research needs to address the predictive validity of the six-point ordinal rating scales of competency currently being used in the MCF. Predictive studies are needed in order to establish whether or not there is support for Hall's hypothesis that the more often one sees the behaviours of competencies in a coach, the more likely a client will evaluate the coaching to be successful in facilitating the goals s/he has for coaching.

Research is also needed to investigate how effective the benchmarking is in assisting coachees to attain their outcomes over the longer term. While the effect on coachees of greater levels of Meta-Coach competency has been observed in coach trainings and chapter meetings, this association has not yet been systematically researched with 'real clients'. Answering the call for researchers to embrace a range of investigative paradigms (Grant et al., 2010, p.33) means designing longitudinal studies where

coaches' benchmarked skills are compared to or correlated with clients' outcomes. However, longitudinal *qualitative* studies can also shed light on how coaches experience and evaluate coaches' different skill levels.

Grant et al (2012) recommend adding to the growing body of randomised between-subject studies and they also suggest that metacognitive techniques in coaching will become an increasing focus of research (Grant et al., 2012, p. 40). Given that meta-questioning is a key and distinctive competency in Meta-Coaching, it would be useful to carry out a randomised study of two coach groups: those using precise or Meta-Model questioning only versus those using precise questioning plus meta-questioning.

The Meta-Coaching benchmarking system is used for training coaches, for maintaining and enhancing practising coaches' competence, as well as for credentialing. The requirement to maintain competency levels through formal, ongoing assessment can be regarded as a form of supervision and more research is needed in order to clarify whether ongoing benchmarking does contribute to maintaining and increasing skill levels.

Specific conclusions

Systematic action research has led to the development and testing of competency-based assessment of coach trainees and practising coaches based on fine-grained behavioural distinctions enabled by training in NLP. As such, it contributes to the discussion on assessment in the coaching profession.

The method of validating the competencies—reviewing published literature and using NLP-based modelling—is distinguished from other methods of validating core coach competencies. The ongoing assessment of competencies involves trained benchmarkers who work with both checklists and rating scales while directly observing the coach-coachee interaction.

The benchmarking in the Meta-Coaching community is a demonstration of the researcher-practitioner model for coach training and practice (Grant and Cavanagh 2004; Kemp 2011) and offers a model of using benchmarking as a form of supervision for practising coaches as well as of trainees (Grant, 2012).

Further research is needed to test inter-observer/rater reliability and address issues of fairness. Outcome studies are needed, using various investigative methods including randomised studies of different coaching competencies and competency levels. Studies of the predictive validity of the Meta-Coaching benchmarking rating scales are a priority.

Acknowledgements

I wish to thank the nine interviewees as well as Michelle Duval, Michael Hall and the anonymous reviewers of this paper.

References

- Amin, M.R. & Amin, N.A. (2003). Benchmarking learning outcomes of undergraduate business education. *Benchmarking: An International Journal*, 10(6), 538-558.
- Bashook, P.G. (2005). Best practices for assessing competence and performance of the behavioral health workforce. *Administration and Policy in Mental Health*, 32(5/6), 563-592.
- Burns, L. & Gillon, E. (2011). Developing a teaching agenda for coaching psychology in undergraduate programmes. *The Coaching Psychologist*, 7(2), 90-97.

- Butler M.P., Cassidy I., Quillinan B., Fahy, A., Bradshaw, C., Tuohy, D. & Tierney, C. (2011). Competency assessment methods - tools and processes: a survey of nurse preceptors in Ireland. *Nurse Education in Practice*, 11(5), 298-303.
- Chaboyer, W., Willman, A., Johnson, P. & Stockhausen, L. (2004). Embedding evidence-based practice in a nursing curriculum: A benchmarking project. *Nurse Education in Practice*, 4(3), 216-223.
- Cowan, D., Norman, I. and Coopamah V.P. (2007). Competence in nursing practice: a controversial concept - a focused review of literature. *Nurse Education Today*, 25(5), 355-362.
- Dilts, R. (1998). *Modeling with NLP*, Capitola, CA: Meta Publications.
- Dochy, F., Segers, M. & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24(3), 331-350.
- Foss, G.F., Janken, J.K., Langford, D.R. & Patton, M.M. (2004). Using professional specialty competencies to guide course development. *Journal of Nursing Education*, 43(8), 368-375.
- Freire, P. (1970). *Pedagogy of the Oppressed*. New York: Continuum Publishing
- Gallagher, P., Smith, T. & Ousey, K. (2012). Problems with competence assessment as it applies to student nurses. *Nurse Education in Practice*, 12(6), 301-303.
- Gallwey, T. W. (1974). *The Inner Game of Tennis*, New York: Random House
- Gobo, G. (2004). Sampling, representativeness and generalizability. In C. Seale, G. Gobo, J. F. Gubrium & D. Silverman (Eds.), *Qualitative Research Practice* (pp.405-426). London, UK: Sage Publications.
- Grant, A. (2011). Developing an agenda for teaching coaching psychology. *International Coaching Psychology Review*, 6(1), 84-99.
- Grant, A. (2012). Australian coaches' views on coaching supervision: A study with implications for Australian coach education, training and practice. *International Journal of Evidence Based Coaching and Mentoring*, 10(2), 17-33.
- Grant, A. M. & Cavanagh, M. (2004) Toward a profession of coaching: Sixty-five years of progress and challenges for the future. *International Journal of Evidence Based Coaching and Mentoring*, 2(1), 1-16.
- Grant, A.M. & Cavanagh, M.J. (2007). The goal-focused coaching skills questionnaire: Preliminary findings. *Social Behaviour and Personality*, 35(6), 751-760.
- Grant, A.M., Passmore, J., Cavanagh, M.J. & Parker, H. (2010). The state of play in coaching today: A review of the field. *International Review of Industrial and Organisational Psychology*, 25, 125-167.
- Griffiths, K E. & Campbell, M A. (2008) Regulating the regulators: Paving the way for international, evidence-based coaching standards. *International Journal of Evidence Based Coaching and Mentoring*, 6(1), 19-31.
- Grinder, J. & McMaster, M. (1993). *Precision: A New Approach to Communication: How to get the information you need to get result*. Scotts Valley, CA: Grinder DeLozier & Associates.
- Hall, L.M. (1995/2002). *Meta-states: Managing the higher levels of your mind's reflexivity*. Clifton CO: Neuro-Semantics Publications.
- Hall, L.M. (2009). Measuring the synergy of self-actualization. In L. Hall, *Achieving peak performance* (pp 255-276). Clifton CO, Neuro-Semantic Publications.
- Hall, L. M. (2010). The facilitation model. In L.M.Hall, *Coaching conversations* (3rd ed, Chapter 12). Clifton CO: Neuro-Semantics Publications.
- Hall, L. M. (2011). *Benchmarking intangibles: The art of measuring quality*, Clifton CO: Neuro-Semantics Publications.

- Hall, L. M. (2011a). *Coaching Mastery: Module III of Meta-Coaching for APMC Certification*. Clifton CO: Neuro-Semantics Publications.
- Hall, L. M. (2011b). Replicating Excellence in Benchmarking. In *Acuity* (2), 54-59.
- Hall, L. M. (2011c). Meta-Coach Reflections #15, April. Retrieved 24 Jan 2014 from <http://www.neurosemantics.com/wp-content/uploads/2013/02/2011-Meta-Coach-Reflections-A.pdf>
- Hall, L.M. (2011d). Meta-Coach Reflections #30, July. Retrieved 24 Jan 2014 from <http://www.neurosemantics.com/wp-content/uploads/2013/02/2011-Meta-Coach-Reflections-A.pdf>
- Hall, L. M. (2011e). Meta-Coach Reflections #44, Nov. Retrieved 24 Jan 2014 from <http://www.neurosemantics.com/wp-content/uploads/2013/02/2011-Meta-Coach-Reflections-A.pdf>
- Hall, L.M. & Duval, M. (2003/2004). *Coaching conversations*, Clifton CO: Neuro-Semantic Publications.
- Hall, L M & Duval, M (2004). *Meta-Coaching, Volume 1: Coaching change*, Clifton CO: Neuro-Semantics Publications
- Hall, L. M, Duval M. & Expert Coaches (2004/2011). *The Meta-Coaching certification trainings*, Clifton CO: Neuro-Semantics Publications.
- ICF (2009). ICF-Competencies-Levels-ACC-PCC-MCC-revised 7-29-09.pdf. Retrieved 16 February 2013 from <http://www.coachfederation.org/icfcredentials/core-competencies/>
- ICF (2009a). ACC-Assessment-Rating-Form.doc. Retrieved 25 April 2013 from <http://www.coachfederation.org/search/searchresults.cfm?QuickSearch=scoring+table/>
- Kelly, M. J. (2004). Qualitative evaluation research. In C. Seale, G. Gobo, J. F. Gubrium & D. Silverman (Eds.), *Qualitative research practice* (pp. 463-477). London, UK: Sage Publications.
- Kemp, T. (2011). Developing an agenda for teaching coaching psychology – invited responses. *International Coaching Psychology Review*, 6(1), 119-121.
- Khabiri, M., Sabbaghan, S. & Sabbaghan, S. (2011). The relationship between peer assessment and the cognition hypothesis. *English Language Teaching*, 4(1), 214-223.
- Ladkin, D. (2004). Action research. In C. Seale, G. Gobo, J. F. Gubrium & D. Silverman (Eds.), *Qualitative Research Practice* (pp. 478-490). London, UK: Sage Publications.
- Linder-Pelz, S. (2013). Benchmarking coaches' skills: Experiences of benchmarkers and of trainees being benchmarked. *Current Research in NLP Vol 3*, UK: ANLP International CIC.
- Linder-Pelz, S. (2010). *NLP coaching: An evidence based approach*, London, UK: Kogan Page Ltd.
- 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.
- Linley, P A. (2006). Coaching research: who? what? where? when? why??. *International Journal of Evidence Based Coaching and Mentoring*, 4(2), 1-7.
- Maslow, A. (1966). *The psychology of science: A reconnaissance*. New York: Harper & Row.
- Maslow, A. (1968) *Toward a psychology of being*. New York: Van Nostrand.
- Mezirow, J. (1978). Education for perspective transformation: Women's re-entry programs in community colleges. New York: Columbia University Center for Adult Education.
- Olian J.D., Durham, C.C., Kristof, A.L., Brown, K.G., Pierce, R.M. & Kunder, L. (1998). Designing management training and development for competitive advantage: Lessons from the best. *Human Resource Planning*, 21(1), 20-31,
- Passmore, J. (2008). *The character of workplace coaching: The implications for coaching training and practice*. Stratford, UK: University of East London.
- Passmore, J. & Gibbs, C. (2007). The state of executive coaching research: What does the current literature tell us and what's next for coaching research? *International Coaching Psychology Review*, 2(2), 116-128.

- Perosa, L.M. & Perosa, S.L. (2010). Assessing competencies in couples and family therapy counselling: A call to the profession. *Journal of Marital and Family Therapy*, 36(2), 126-143.
- Reynolds, M. & Trehan, K. (2000). Assessment: A critical perspective. *Studies in Higher Education*, 25(3), 267-278.
- Rodriguez, D., Patel, R., Bright, A., Gregory, D., & Gowing, M.K. (2002). Developing competency models to promote integrated human resource practices. *Human Resource Management*, 41(3), 309-324.
- Rush, S., Firth, T., Burke, L. & Marks-Maran, D. (2012). Implementation and evaluation of peer assessment of clinical skills for first year student nurses. *Nurse Education in Practice*, 12(4), 219-226.
- Smith, T. F (2009). Going beyond competencies: An exploratory study in defining exemplary workplace learning and performance practitioners, *Performance Improvement Quarterly*, 22 (1), 27-51.
- Vickerman, P. (2009) Student perspectives on formative peer assessment: An attempt to deepen learning? *Assessment and Evaluation in Higher Education*, 34 (2), 221-230.
- Voorhees, A.B. (2001). Creating and implementing competency based leaning models. In R.A. Voorhees (Ed.), *Measuring what matters: Competency-based learning models in higher education* (pp. 83-95). San Francisco: Jossey-Bass.
- Voorhees, R. (2002). Defining and assessing learning: Exploring competency-based initiatives. Report of the national postsecondary education cooperative working group on competency-based initiatives in postsecondary education. Retrieved 26 January 2014 from <http://nces.ed.gov/pubs2002/2002159.pdf>
- Weeks, P. (2000). Benchmarking in higher education: An Australian case study. *Innovations in Education and Training International*, 37(1), 59-67.
- Welsh, M. M. (2007). Engaging with peer assessment in post-registration nurse education. *Nurse Education in Practice*, 7 (2), 75-81.

Susie Linder-Pelz has worked in business, academia and professional development. Susie has a PhD in Sociomedical Sciences from Columbia University and was a behavioural science researcher before developing a successful NLP-based career coaching practice. She has authored 22 articles in peer-reviewed journals, 19 reports of commissioned research and five books.

Appendix 1: The Meta-Coach Foundation and Associate Certified Meta-Coaching training

The Meta-Coach Foundation (MCF) includes a Leadership Team, Regional Leaders, certified Meta-Coach trainers and ACMC Assist Teams (comprising past ACMC participants) which are created specifically for each training. Typically there is one Assist Team member for every three trainees; triads are formed where an Assist Team member observes/benchmarks a trainee coach as s/he 'coaches' another trainee.

The pre-requisite for Associate Certified Meta-Coaching training (ACMC) is a minimum three-day course in NLP essentials and a three-day training on the Meta-States model. At the end of an ACMC program, candidates can receive either provisional or full certification. There is a regional MCF chapter structure with 52 chapters at different stages of development. At local chapter meetings coaches' skills are benchmarked in order to gain certification or increase skills.

What they get in the chapter meeting is much closer attention; we give them sensory-based feedback: 'Here's what we're seeing, here's what we're hearing, here's the benchmark, here's what it looks and sounds like, and can you see the gap?' (Member of MCF leadership team interviewed for this study)

The MCF community includes all certified Meta-Coaches. Membership depends on being in good standing (payment of fee) and maintaining standards through the benchmarking process. In December 2011, 1,421 people in 37 countries had been licensed as Meta-Coaches.

Appendix 2: Scaling the levels of a coach's listening skill in behavioural and sensory-specific terms (Hall 2011:204). Reproduced with permission.

- 0 No or little eye contact with coachee (25% or less of time), no tracking of words or actions being said, talks over, tells, teaches, makes evaluations and interrupts.
- 1 Makes some eye contact (less than 60% of the time), paraphrases coachee's words, only partially keeps general track of the precise words, misses points, asks 'where are we?', takes notes on things other than coachee's statements, his/her eye movements indicate internal processing while coachee is speaking.
- 2 Maintains eye contact 75% of the time, repeats back specific words and some paraphrasing that matches coachee's words, speaks 60% or more of the time, gives little or no time for coachee to speak. Speaks immediately when coachee stops, rushes to make comments or ask questions.
- 3 Actively explores what the coachee speaks about to understand fully, encourages coachee to speak by using head nods, encourages with sounds such as 'hmmm', 'ahhh', 'yes, go ahead', 'say more'. Uses pauses and coachee speaks at least 50% of the time.
- 3.5 Looks at coachee, makes eye contact, speaks 20% of the time or less.
Turns body to coachee to be facing and physically present to the coachee.
Acknowledges the talking by maintaining eye contact.
Uses soft 'sparkling eyes' (eyes opening wider when responding to coachee).
Nods the head in acknowledging what coachee says.
Uses sounds, nods and words that signal coachee to talk.
Tracks words and gestures over time.
Asks about what coachee has not said.
Asks questions that invite coachee to ask self more questions or share thinking and feeling.
Asks questions that probe for details about coachee's view or the meaning of coachee's words.
Asks questions for self-listening ('Did you hear what you just said?') to enable coachee to know what's 'in the back of the mind'.
Gives space and time for person to be with those thoughts and feelings.
Is silent as the coachee speaks 60% to 80% of the time.
Asks awareness questions, e.g. 'How aware are you that you have said lots of things about X, but nothing about Y?'.
Uses extended pauses of more than five seconds.
Feeds back coachee's words and gestures as a mirror for coachee.

Appendix 3. The interview questions

“The aim of this study is to throw light on the development and use of the NLP Neuro-Semantic benchmarking methodology. So I would like to ask you a few specific questions, first about the benchmarking of seven core + 13 advanced coaching skills in ACMC (Meta-Coach certification trainings).

1. Can you tell me about your participation in ACMC? ACMC Assist Team? Neuro-Semantic Trainers training? Leadership Team? Describe your role in benchmarking coach/trainees' skills.
2. What's happening at MCF chapter level? How many? How often? Agenda?
3. Michael's hypothesis is that ‘the more often we see the behaviours of competencies in a coach, the more likely a coachee will evaluate the coaching to be successful in facilitating the goals of coaching (e.g., clarity, decision, planning, resourcing, changing, exposing incongruencies).’ How was this tested? And do you have—or know of—any evidence/ (see, hear, experience, read) that lack of practice reduces behavioural competency?
4. Re distinguishing behaviours indicating below competency/ competency/ enriching the competency and making it fuller, more robust. How was this done? (Prompt questions: By getting feedback from coaches & coachees’ in M-C trainings, demos?) Role of Assist Teams? Team Leaders? When, where, with whom? How exactly?
5. Hall’s conclusion to date is that ‘Today the standards are rigorous enough and most of the Assist Teams are trained to the level that only 25% to 40% of people can reach the standards during the Coaching Mastery course’. How was that percentage obtained? Who decides whether the participants reach the standard? What record-keeping by Assist Team and/or Michael at end of ACMC are you aware of?
6. Hall says ‘Most will reach competency level at some time later, through one of the MCF chapter meetings.’ To what extent is this happening? Who is monitoring post-ACMC competency levels? How is it documented and fed back to Hall?
7. In your experience and observation of ACMC programs and MCF chapter meetings, what are the key challenges in training to benchmarks? And key learnings?”