Making the most of Psychometric Profiles – Effective Integration into the Coaching Process

Dr Almuth McDowall and Dr Rainer Kurz


Abstract

This practical paper is based on a skills session as delivered at the first International Coaching Psychology Conference held in 2006. It commences by discussing the use of psychometrics in general by emphasising the four psychometric principles as hallmarks of a good test; and outlining some of the advantages and potential limitations of psychometrics. In this paper a contemporary instrument, the Saville Consulting Wave® is introduced and its application is discussed in relation to coaching, with particular reference to a performance coaching context at work. It is concluded that no psychometric test is a panacea for each and every situation, but that skilful use greatly enhances the coaching process.

Keywords:
Psychometrics, reliability, validity, freedom from bias, standardisation, objectivity, test user, test taker, Wave TM, Five Factor Model of Personality, Great Eight Competencies, competency at work.

Introduction

Psychometrics are not for everyone, and certainly not for every coach. To illustrate, coaches who see their practice rooted in Humanistic or Rogerian approaches may favour an on-going dialogue over the use of assessments during the coaching relationship. Anecdotal evidence for instance suggests that a number of coaches practicing career coaching favour other techniques, such as interviews, value card sorts or questioning techniques derived from counselling psychology over the use of tests or questionnaires. However, we believe that psychometrics can make an effective contribution to any coaching relationship, if, like a good seasoning, they are used sparingly and with care. This article is based on the skills based session that was delivered at the First International Coaching Psychology Conference in December 2006, and thus necessarily takes a practical rather than theoretical focus.
Our discussion focuses on the use of coaching and psychometrics at work, rather than other contexts.

First, we debate the value of psychometrics in coaching in a general sense, by discussing what psychometrics are, how to choose psychometrics, how to use psychometrics and when to use them in a coaching process. This will also entail a critical perspective on their potential limitations, with particular reference to the end user. Next, we will take our readers through a case study, using an actual profile, offering different approaches for interpretation and future use, leading to a final conclusion and recommendations for best coaching practice.

**What defines a psychometric test?**

Psychometric measures or instruments divide into ability tests where answers are scored as 'right' or 'wrong' and self-report questionnaires where there are no 'right' or 'wrong' answers. For convenience we will use the term test as shorthand in line with current practice.

It is important that coaches understand what psychometrics are, as even versed practitioners may find it difficult how to tell whether a measure is ‘fit for purpose’. Whilst this following section may at first glance seem simplistic, the fundamental understanding of psychometric principles is core to their use. Defining characteristics of a psychometric measure are four psychometric principles: reliability, validity, freedom from bias and standardisation (Rust, 2004).

Reliability refers to whether a measure is consistent, across time, across people and different applications (Rust & Golombok, 1999). The most commonly reported form of reliability is internal consistency, measured through Cronbach’s Coefficient Alpha, which tells us to what extent all items measure the same underlying construct (Cronbach, 1951). Reliability can however be more powerfully established through research procedures such as test re-test or alternate form comparisons. If the association between scores people obtain on different occasions at time one and time two or across two versions is sufficient, satisfactory reliability has been established.

Validity refers to whether a measure actually measures what it says it does. This principle is perhaps the most complex, as there are different forms of validity.
Content validity assesses whether an instrument measures everything that it should be measuring. For instance, a work-based competency measure would in all likelihood need to measure different behavioural constructs, and not just one. Construct validity tells us whether the underlying psychological construct of a measure holds up. This can be established through procedures where scores from a new measure of a construct are correlated with an existing test of that construct. Criterion-related validity tells us whether test scores are associated with objective criteria, for instance whether results of a test correlate with productivity or performance scores.

Reliability is a pre-requisite for validity (Rust & Golombok, 1999) as any measure has to be reliable to be valid – if we cannot observe reliable results, we cannot be sure that these are not only due to chance alone. Saville and Nyfield (1975) summarised their interplay as: ‘Reliability is about getting the test right, validity is about getting the right test.’

Freedom from bias means that a test should produce consistent results for everyone. If a test is biased, it means that groups of people, such as men and women, are getting different scores. One of the main sources in our multi-cultural society is item bias, where speakers of English as a second language interpret colloquial items, such as ‘beating around the bush’, or idiomatic items in a different way to native speakers (e.g. Rust, Daouk & McDowall, 2005). At the same time, it is possible that there are real and genuine differences between different groups of people, women for instance score higher on the inter-personal aspects as measured by the Trait Emotional Intelligence Questionnaire [TEIQ] (Petrides & Furnham, 2004).

Standardisation means that the test is administered and used under standardised conditions, and interpreted in a standardised way. Tests are typically either norm-referenced, where an individual’s score is compared against general (e.g. General Population) or specific norms (e.g. Senior Executives), or criterion-referenced against a common benchmark that should be achieved on a test. An example for a norm-referenced ability test in the context of work would be the Watson Glaser Critical Thinking Appraisal UK where the manual would refer both to the general working population as a potential benchmark comparison group, but also specific norm groups such as MBA students or senior managers. Standardisation also refers to standardised procedures when administering, scoring and interpreting a test, as the same procedure should be applied in each and every situation and for every test.
taker to ensure that human error and situational influences are minimised. The test manual would provide test takers with guidance here; some tests such as the Rust Advanced Numerical Reasoning Appraisal [RANRA] for instance can be administered either timed or un-timed.

Thus, a good instrument should adhere to all four principles and have corresponding information in the test manual. In the UK, test reviews are available to members of the British Psychological Society via the Psychological Testing Centre while in the US the Buros Foundation publishes test reviews (see web-links at the end of this article).

**Potential challenges for using psychometrics in coaching**

There are some limitations around the use of psychometrics that we note upfront. There are some tests, such as the famous Rorschach Ink Blot test (e.g. Vernon, 1933), that are used with subjective interpretation. Most standardised instruments are purportedly objective, however, although even their use entails an element of subjectivity. Particularly personality profiles rely on the skilled interpretation of the test user to ensure objectivity. Another limitation is their choice of test as practitioners may prefer to keep using tried and tested measures. The research on personality measures is an example. Research over the last decades has consistently shown that the ‘Big Five’ model of personality factors provides an excellent summary of personal characteristics that is very robust (e.g. Barrick & Mount, 1991). There are many tests that build on this trait model, such as the NEO NFI, the HPI or the Orpheus. Not all practitioners have taken this information on board however, and stick in every situation with measures based on older models that they are used to, such as older ‘type’ measures. This can result in ‘evangelical test users’ – practitioners who solely rely on one or few tests. Whilst we do understand, and even advocate, that repeat use of a measure will make coaches better at providing in-depth and rounded assessments, this also brings the danger of over-interpretation. Test users may interpret individual scores too literally, or read too much into overall profiles, without corroborating their validity with the coachee.

Current training models, such as the British Psychological Society (BPS) Certificates in Occupational Testing in the UK, set minimum qualification requirements for test use. Level A training covers ability tests and is universally accepted by all reputable test publishers. Level B (Intermediate) training allows practitioners to use one
personality instrument in the workplace only with short conversion training required for other questionnaires to safeguard accurate standards of interpretation. Further conversion training that safeguards correct interpretation of a specific instrument is usually required but may be expensive and thus somewhat limit practitioners appetite to move beyond a limited ‘toolkit’ that they are familiar with.

Despite these possible limitations, skilful use of psychometrics can add value to any coaching process. However, it is important to consider some of the common misconceptions about psychometrics which have long been noted (Rust, 2004). There is a commonly held belief that psychometrics treat everyone in a robotic simplistic way, and fail to draw out rich individual differences for instance with reference to stable underlying preferences (personality). This can be argued to the contrary. Whilst the psychometric profile itself should be robust, objective and standardised, its value lies in the discussion with the coachee, where care is taken to verify this evidence by drawing out relevant examples that illustrate typical behaviours. For some coachees, the profile may be accurate as it stands. Others may have found ways of compensating for natural preferences, for instance overcoming a tendency to overlook detail by utilising appropriate checking mechanisms.

The second misconception addressed here is that tests are impersonal. Again, we highlight the fact that their value lies in the individual discussion of the profiles, which should always be a two-way process and give the coachee ample room to share their experiences. It is up to the coach to use this evidence with her or his professional judgement and, triangulate with the profile itself to formulate a comprehensive psychometric assessment.

Third, there is a widely held belief that people cheat on personality questionnaires to present themselves in the best possible light. A well designed questionnaire however should have build in checks that pick up any exaggerations or inconsistencies that flag to the coach whether this is a problem. Tests in the past used specific items to form a Social Desirability scale to flag possible attempts to fake, an example being the EPQ (Eysenck Personality Questionnaire). Modern questionnaires often profile consistency and acquiescence using sophisticated computer scoring algorithms; the Orpheus for instance has four different audit scales which assess potential positive or negative distortion, as well as checking for inconsistent or contradictory responses. More fundamentally though, there should be no need for the coachee to try and bias
his or her responses in a trusting coaching relationship. Whilst candidates trying to distort their responses may be a problem in a selection context, the context of coaching should be such that trust and transparency are established from the beginning, making such attempts to manipulate unnecessary and unlikely.

Last but not least, there is also a misconception that psychometrics put people into ‘boxes’. This is likely to be due to the popularity of ‘type’ measures that presume that people pertain to a number of psychological types that remain stable over life-time. Contemporary instruments that are designed for use in the workplace acknowledge however that people’s personal preferences or styles can change and thus be targeted through activities such as coaching. Examples for measures that combine the measurement of stable characteristics with the prediction of work-based competencies are the Occupational Personality Questionnaire (Saville, Holdsworth, Nyfield, Cramp & Mabey, 1984) and Saville Consulting Wave® (Kurz, Saville & MacIver, in press) both developed under the leadership of Professor Peter Saville. The scales in these tools are work relevant rather than designed to measure general or clinical constructs. They operate at the level of detail that reflects the true complexity of people and jobs rather than at the parsimonious yet highly abstract level that academics prefer. The tools distinguish between the measurement of psychological traits on one hand and the reporting of derived Competency Potential scores that translate this specialist terminology into the everyday language of competencies backed by large scale criterion validation evidence on the other hand.

The advantage of such competency-oriented measures is that they can be used at various stages in the coaching process. Good coaching lets a coachee become more aware of what they can do with their life and prepares them to take more responsibility for it. This is achieved through sessions structured around questioning. The test results can for example support the use of each stage of the GROW (Goal, Reality, Options, Will) approach originally put forward in the 1980s (Alexander & Renshaw, 2005) (that sequences or orders those questions in sessions.

First, they can be used as a baseline measure that helps coach and coachee understand the general goals aspired to, the current reality, the option of building on specific strengths or tackling development needs, and what the coachee is likely to commit to. It is likely to be useful to bring in other evidence, too, such as a ‘value elicitation’ task or a ‘lifeline’ exercise that generates evidence over and above the psychometric profile. Second, as these measures allow for behaviour change, they
can also be used as an evaluation tool towards the end of a long-term coaching process, or in a stand alone follow up session to measure where progress has been made.

The next section illustrates the use of psychometrics in coaching using actual, albeit anonymised, profile excerpts generated using the Saville Consulting Wave® Professional Styles questionnaire. This section will commence with an introduction to the underlying model, then move to explain the profile; resulting in suggestions for interpreting and using the information present.

Using Saville Consulting Wave® Styles Questionnaires in Coaching

The Saville Consulting Wave® model (see MacIver, Saville, Kurz, Mitchener, Mariscal, Parry, Becker, Saville, O’Connor, Patterson & Oxley, 2006) is hierarchical, providing four levels of detail and utilising a century of personality research as well as technological advances to create an integrated suite of tests. Figure 1 shows the four behaviour clusters at the apex that provide a broad overview of the key characteristics that underpin work performance. Each cluster is comprised of three sections that are particularly suitable for finely grained assessment. Each section consists of three dimensions that measure behaviour at the level of detail expected by experienced psychometric test users. Each dimension breaks down into three facets that jointly define the dimension. These facets provide breadth of measurement while maintaining clarity of meaning.

The model incorporates as shown in Table 1 the Great Eight competencies (a model of generic workplace effectiveness) as well as widely accepted psychological constructs such as the Big Five personality factors, motivational need factors and intelligence.

Note to editor: insert Table 1 about here

Barrick and Mount (1991) outlined the research base for the broad Big Five trait factors that have frequently been found and replicated in personality research Openness to Experience, Conscientiousness, Extraversion, Agreeableness and Neuroticism (the initials of which form the handy OCEAN mnemonic). They traced back the origins of the Big Five to the work of Norman (1963) and acknowledged that
over the years different names had been used for what is now understood to be essentially the same construct set. Neuroticism is increasingly referred to in the wake of Positive Psychology as Emotional Stability or Confidence while Agreeableness and Openness to Experience are sometimes measured through their opposite pole e.g. Independence and Conventionality respectively.

The emergence of the Big Five model as the higher-order categorisation of self-report personality factors eventually led to the development of the ‘Great Eight’ competencies model by Kurz and Bartram (2002), whose model added Need for Power, Need for Achievement and Intelligence oriented competency constructs. They defined competencies in relation to their significance for performance at work as ‘sets of behaviours that are instrumental in the achievement of desired results or outcomes’. Ability or personality traits in contrast ‘exist’ and can be measured in isolation from a work context.

Saville Consulting Wave® integrates the two leading assessment models into the ‘Fab Four’ higher-order factors and the ‘Terrific Twelve’ sections as shown in Figure 1 that provide a more detailed differentiation to reflect the true complexity of people and jobs.

Note to editor: insert Figure 1 about here

Solving Problems is important in all jobs. Individuals have to first critically evaluate problems, then investigate the issues and generate innovation.

Influencing People is also usually important in terms of building relationships, communicating information and leading people.

Adapting Approaches is related to constructs of Emotional Intelligence that underpin resilience at the work place, flexibility in the face of change and support of others.

Delivering Results finally is about detailed implementation of assigned work, structuring of work tasks and motivational drive to bring things to conclusion.
Whilst a range of tools are available based on the Saville Consulting Wave® model, the remainder of this section will concentrate on the Professional and Focus Styles measures as they are most pertinent to coaching situations.

**Using the Psychometric Profile**

Three features of Saville Consulting Wave® Styles questionnaires are briefly explained below with reference to Figure 2.

*Note to editor – insert Figure 2 about here*

The report always profiles the dimension scores on a 1-10 standardised ‘Sten’ scale. Each dimension of the questionnaire is comprised of three facets for which a Sten value is given after the verbaliser text. If the facet Sten results differ by three or more Stens then this is shown through ‘Facet Range’ hatching on the report. Such ranges pick up very subtle aspects of the individual that represent their uniqueness; and provide valuable information to the coach that can feed into an entire series of coaching sessions based on the individual’s profile.

The Styles questionnaires present blocks of six items that have to be rated on a 9 point ‘agree-disagree’ rating scale (normative rating). If any items receive the same rating they presented once more but this time in a ranking task screen where individuals have to state which item is ‘most’ and ‘least’ true for them (ipsative ranking). This dynamic dual response format is unique to Saville Consulting Wave®. It enables validity checks (was the test taker honest?) but also homes in on the areas where the individual is most likely to experience conflicts under pressure. ‘Normative-Ipsative Splits’ are displayed in the profile if there is a substantial difference between normative rating and ipsative ranking results. Normative ratings are likely to reflect everyday behaviour while ipsative rankings are likely to reflect behaviour under pressure when time and resources are limited. The ipsative scores will pull down the results of those who have been very generous on themselves in the normative rating, and boost the scores of those who have been overly self-critical in their ratings. Again, this provides valuable information to the coach, particular when being asked to coach in an organisational context where impression management is in fact part of everyone’s job.
Each block consists either of Motive items that are personality oriented or Talent items that are competency oriented. Each facet in the model has one item of each type. A ‘Motive-Talent Split' is shown if there is a substantial difference highlighting areas where individual preferences and actual behaviours are misaligned. Where Talent is lower than Motive, individuals effectively have identified themselves a development area. Where Motive is lower than Talent, individuals may, at worst, be at risk of burn-out as their behaviours are not supported by underlying motivation. This information can also be utilised to good effect in coaching sessions, for instance by conducting a gap analysis (“where do you think you are at the moment, and where would you like to be?”).

Using a competency potential profile

Coaching that is focused on improving performance is one of the most common applications in organisational settings. Thus, it makes intuitive sense to incorporate a psychometric measure focused on competence into the coaching process. The Wave® contains such a competency profile that maps people's preferences concisely against behaviours that are effective and valued in the workplace. Please refer to Figure 3 for an example profile for a ‘Peter Purple' showing scores on 12 Behaviour Sections with scores on the three Behaviour Dimensions that sit under each summary heading. Scores are also reported in Stens based on the validation evidence where hundreds of managers were rated on the effectiveness of their behaviours. The Behaviour Competency model is structurally parallel to the Professional Styles model with sophisticated equations that optimise criterion-related validity linking the personality and competency taxonomies (see Kurz, Saville & MacIver, in press).

Note to editor: insert Figure 3 about here

Whilst such profiles are immensely useful, as their language and structure is easily mapped against typical organisational competencies, they can also be overwhelming for the coachee, who is presented with a wealth of information, as each of the 12 competencies has a separate score. Plus, as stated earlier, there is always the necessity to corroborate the information with the coachee.
One way of eliciting this information that is useful is to use ‘talent traffic lights’. As the competency profiles use colour coding, where high scores are marked in green, medium ones in amber, and low scores in red; it can be conducive to get coaches to think about their ‘red, green and amber lights’ before disclosing the profile; which we set out below.

**Turning Competency Profiles into ‘Talent Traffic Lights’**

**Step 1**: First, describe the competency areas that are particularly important in this context (‘Requirement Profiling’). This could be based on consultation with the client, the coachee, or both. Accurate profiling of the requirements contributes to a coaching session by highlighting the areas that are key to the role and avoiding undue emphasis on less important areas that would detract from the coaching process. This needs to be done as preparation for the relevant coaching session, and can feed into the introduction and setting the scene in the actual session itself. In terms of discussing the profile itself, we have found it helpful to take a very open approach, where we get coachees to think about the competencies in their own words, before disclosing the actual profile. In the context of Peter Purple’s profile outlined in Figure 1, required competencies were ‘thinking outside the box’ (e.g. Creating Innovation, Evaluating Problems, Presenting Information) and also ‘moving away from the silo mentality’ (Providing Leadership, Communicating with People, Providing Support).

**Step 2**: Having introduced the purpose of the assessment (“e.g. today, I would particularly like to use this assessment to explore how you innovate and work with others at work”) the profile is discussed in more detail. It can be useful to start off with ‘green lights’, in order to get coachees into a positive and receptive frame of mind, and enable them to later utilise their areas of strength to work on other aspects. Ask the coachee to describe their ‘green lights’ as they see them at the moment – which are their current strengths in general? Ensure to elicit concrete examples. Then, it may be necessary to corroborate this through further questioning. To illustrate, Peter Purple is high on Creating Innovation, but less so on Presenting Information. Appropriate questions might be “When, and under what circumstances are you at your best when generating new ideas and developing new strategies?” or “How can you ensure that you communicate these ideas to other people?”
Step 3: Now ask the coachee to describe their ‘amber lights’ – these are competency areas where they are doing ok, but could do better; or perhaps do not have the time or opportunity to do what they want to do. These may not be as salient to the coachee as ‘green’ or ‘red’ areas, and thus require quite specific prompting. Referring to the profile in Figure 2, a potential amber area for exploration is Adjusting to Change, where there is a difference between how the individual embraces change (very readily) and how they invite feedback from others (not so effectively). An effective question might be “how do/ how can you ensure that others think about change at work in the same way that you do?”

Step 4: Ask the coachee to describe their ‘red lights’ – any areas that definitely need more work? The profile in Figure 2 would indicate someone who is less competent at people skills, than at managerial transaction, for instance the score on Providing Support is extremely low. Open questions should draw out relevant examples (e.g. “How do you support others at work?”, “Can you describe a time when you did this effectively?”). As there might be a discrepancy between the score and how people see themselves, it may be necessary to disclose the scores at this point; and structure the conversation around the difference between the score and examples given. Using Peter Purple’s example, it turned out in the discussion that this individual could actually support others when required, but would give priority to operational requirements over people skills in his day to day tasks.

Step 5: Last, triangulate the evidence from this discussion with the actual profile, and probe further particularly where there were gaps or contradictions. Then work with the coachee on concrete action planning – how can they use their ‘green lights’ (current strengths) to work on their amber and red areas? Relating this back to the Peter Purple, he acknowledged that Providing Support to the entire team of workers was something he found difficult; consequently his specialist skills were better used in a more strategic and operational role. However, it also transpired that the potential for innovation could be facilitated at an individual and interpersonal level, as this person proved adept using his innovation potential to support particular individuals at work who were experiencing very unusual challenges, such as a subordinate who had faced very taxing personal circumstances.

Steps 1 to 5 could feed directly into an action plan for future coaching sessions, perhaps ideally with some support from the line manager to ensure that behaviour
change takes place; even if details of the coaching sessions remain confidential. For Peter Purple, further information from the Professional Styles report could prove valuable as the response checks would indicate someone who tended to be ruthlessly honest, and thus rather harsh on himself; which would need to be born in mind by the coach, particularly if other individuals in the same organisation were also to receive coaching.

**Conclusion**

The approach outlined above is of course only one potential application of psychometrics. The Saville Consulting Wave® suite, and particularly the competency profile, is particularly suited to performance coaching in a work context for managerial levels. For other coaching purposes, for instance providing career coaching to young individuals who are about to enter or entering the world of work, a different instrument, or indeed an approach without psychometrics, might be more suitable. We do not believe that there is any such thing as a ‘best test’, but rather that test users should use the right test for any given purpose. This article, nevertheless, demonstrated how psychometric profiles can make a valuable contribution to the coaching process. Modern psychometric measures such as Saville Consulting Wave® provide not just psychometric profiles but output in competency oriented terminology that is easy to understand for coach and coachee alike. The complexity of standardised Sten scores can be reduced by the ‘Talent Traffic Lights’ approach outlined to pinpoint development needs as well as areas of strengths that can be leveraged to maximise performance and well being at work.

Dr Almuth McDowall, University of Surrey

Psychology Department
Faculty of Arts and Human Sciences
Guildford GU2 7HX, United Kingdom
a.mcdowall@surrey.ac.uk

Dr Rainer Kurz, Saville Consulting UK Ltd

Harley House
94 Hare Lane
Claygate
Surrey
KT10 0RB, United Kingdom
rainer.kurz@savilleconsulting.com
References


**Useful Web-links**

Psychological Testing Centre, web-site hosted by British Psychological Society:
www.psychtesting.org.uk

Buros Library of Mental Measurements:
http://www.unl.edu/buros/bimm/html/library.html
Figure 1: The Saville Consulting Wave® Behaviour Model
Table 1: Mapping of Great Eight Competency Factors and Psychological Constructs against the Saville Consulting Wave® Behaviour Clusters

<table>
<thead>
<tr>
<th>Wave Behaviour Clusters</th>
<th>Great Eight Competencies</th>
<th>Psychological Constructs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solving Problems</td>
<td>Analysing &amp; Interpreting</td>
<td>Intelligence</td>
</tr>
<tr>
<td></td>
<td>Creating &amp; Conceptualising</td>
<td>Openness to Experience</td>
</tr>
<tr>
<td>Influencing People</td>
<td>Interacting &amp; Presenting</td>
<td>Extraversion</td>
</tr>
<tr>
<td></td>
<td>Leading &amp; Deciding</td>
<td>Need for Power</td>
</tr>
<tr>
<td>Adapting Approaches</td>
<td>Supporting &amp; Co-operating</td>
<td>Agreeableness</td>
</tr>
<tr>
<td></td>
<td>Adapting &amp; Coping</td>
<td>Emotional Stability</td>
</tr>
<tr>
<td>Delivering Results</td>
<td>Organising &amp; Executing</td>
<td>Conscientiousness</td>
</tr>
<tr>
<td></td>
<td>Enterprising &amp; Performing</td>
<td>Need for Achievement</td>
</tr>
</tbody>
</table>

Figure 2: Saville Consulting Wave® Psychometric Profile Excerpt of the Professional Styles Expert Report
Figure 3: Peter Purple Competency Potential Profile