

Volume 22 No. 2, April 2006

ISSN 0963-2638

Selection & Development Review

PUBLISHED BY THE BRITISH PSYCHOLOGICAL SOCIETY

SDR

CONTENTS

Cognitive Ability Tests and unfairness against minority ethnic groups	3
Stephen A. Woods	
The psychology of 360 degree feedback	9
Peter Goodge & Jane Coomber	
The case for coaching	12
John Toscano	
Spurilitarianism	14
Glen Fox	
Making waves	17
Rob MacIver and the Development Team	
Letter to the Editor	24
Francis Catermole	

Selection & Development Review Editorial Team

Dr John Boddy

16 Tarrws Close, Wenvoe, Cardiff CF5 6BT.
Tel: 029 2059 9233. Fax: 029 2059 7399.
E-mail: JBoddy2112@aol.com

Stuart Duff, Stephan Lucks & Ceri Roderick

Pearn Kandola Occupational Psychologists,
76 Banbury Road, Oxford OX2 6JT.
Tel: 01865 516202. Fax: 01865 510182.
E-mail: stephan.lucks@oxford.pearnkandola.com

Philippa Hain

98 Plymouth Road, Penarth CF64 5DL.
Tel: 07816 919857. E-mail: philippa.hain@ntlworld.com

Consulting Editors: Dr S. Blinkhorn; Professor V. Dulewicz; Professor N. Anderson.

Published six times a year by the British Psychological Society,
St Andrews House, 48 Princess Road East, Leicester LE1 7DR at £37
(US \$50 overseas) a year. Tel: 0116 254 9568. Fax: 0116 247 0787.
E-mail: mail@bps.org.uk. ISSN 0963-2638

Aims, objectives and information for contributors

SDR aims to communicate new thinking and recent advances in the theory and practice of assessment, selection, and development. It encourages critical reviews of current issues and constructive debate on them in readers' letters.

SDR is strongly oriented to the practice of selection, assessment and development, and is particularly keen to publish articles in which rigorous research is presented in a way likely to inform and influence the work of practitioners. It also seeks articles from practitioners drawing on their experience to indicate how practice can be improved.

SDR is not intended to be an academic journal. Articles are reviewed by the editorial team for their relevance, rigour and intelligibility, but not all papers are referred to independent referees. The aim is to get new, practitioner-relevant data and ideas into print as quickly as possible. *SDR* is also open to book reviews in its area.

The Editorial Team aim to give a platform for a range of views that are not necessarily their own or those of the BPS. Articles (2000 words maximum) should be sent as an e-mail attachment, saved as a text or MS Word file, containing author contact details. References should follow the BPS Style Guide (available from publications page of www.bps.org.uk).

Editorial

COGNITIVE ABILITY TESTS (CATs) are widely used in assessment for selection and Steve Woods' article will cause many of us discomfort by highlighting the difficulty of avoiding discrimination against different ethnic groups. Particularly important is the issue that CATs may have differential predictive power for job performance between different groups. Steve proposes appropriate statistical procedures for assessing fairness that many selectors may find it difficult to get their heads round and which most organisations are unlikely to be willing and able to devote the resources to. However, by not doing it, they are potentially exposed to litigation by rejected ethnic minority candidates. As Steve says, there is some work here for the test publishers to carry out.

There is a growing appreciation of the value of 360 degree feedback to help managers gain a balanced view of their performance and specific data on their strengths and weaknesses. Peter Goodge and Jane Coomber underline how the effectiveness of this sort of feedback is dependent on the quality of the process and provide us with some useful guidelines for assuring that quality.

Coaching has enjoyed a great rise in popularity in recent years, particularly evident in the very large number of people who have subscribed to the new BPS Special Group for Coaching Psychology. It is, therefore, appropriate for John Toscano to lay out the case for coaching being a particularly effective means of personal development.

In Issue 21, No. 6, in his article 'Is it Worth It?', Mark Parkinson attempted to bring utility analysis back to centre stage in promoting the benefits of getting occupational psychologists to design selection processes. In this issue Glen Fox takes Mark to task for missing out some important objections to the way utility analysis is sometimes used. She highlights the dangers of psychologists undermining their own position by selling their services on spurious utility coefficients assuming a starting point of random selection and failing to take into account the incremental validities in the multiple selection processes usually deployed. The utility of selection processes is a key issue in occupational psychology, establishing the value it can add to business and we would hope to see a continuation of this debate.

Peter Saville has played a leading role in the development of psychometric assessments for occupational settings in this country, but, for various reasons, has been quiet for some time. Now we know why. He has been gestating a new, highly innovative, suite of assessment tools in the new organisation that he has set up. The *SDR* Editorial Team has to tread a fine line in keeping the academic and practitioner community informed of new developments in psychometrics, whilst not providing a blatant marketing vehicle for test publishers. We felt that the Saville Consulting Wave tools represent important developments and that it was, therefore, right to publish an invited article by Rab MacIver and the Development Team, giving key technical information and particularly outlining innovative features and their rationale, such as the integration of normative and ipsative measures, making it 'validation centric' and indexing the match between talent and motivation. We believe that the instruments involve a number of new concepts that are worth debating and we welcome discussion of them.

We hope that you enjoy the variety and the relevance of this April issue of *SDR* and, as usual, will be delighted if it stimulates debate in the pages of future issues.

John Boddy

Cognitive Ability Tests and unfairness against minority ethnic groups: Two practical ways to check for unfairness in selection

HOW CAN WE ASSESS whether Cognitive Ability Tests (CATs) are unfairly discriminatory against minority ethnic groups in the UK? Most test publishers and test users in the UK accept that we don't have a satisfactory answer to this question. Despite a high volume of evidence demonstrating the potential for adverse impact against minority ethnic groups from the use of CATs in selection, test users do not currently have an agreed method for evaluating the fairness of CATs as selection methods.

The absence of rigorous evaluations of test fairness in some of the most widely used CATs in the UK is easily illustrated. Take a browse through the test manuals you use, and try to find some data on ethnic group differences in test score means or validities. I bet that only a small proportion of test manuals report these data, and that none of them give any clear guidelines on what practitioners should do if the test is unfair against a particular group.

As a result, most test users apply CATs in selection activities assuming that they are fair because there is no evidence to the contrary. Compare this with test reliability and validity. We never assume a test to be reliable or valid without sufficient empirical evidence. It is only with regard to fairness that we apply this faulty reasoning.

This article presents two practical checks that test users can perform to identify whether the tests they use in selection activities are unfair against minority ethnic groups. The article has two sections. In the first, the literature on ethnic-differences in test performance, and its implications for unfairness in testing are examined. The second describes and discusses the procedures

Stephen A. Woods
University of Nottingham

proposed by Woods (2006) for checking unfairness in CATs.

CATs, validity, and score differences across ethnic groups

In selection research and practice, CAT scores have consistently been found to be strong predictors of job performance criteria across a range of jobs in the US and Europe (Schmidt & Hunter, 1998; Salgado *et al.*, 2003). CATs are widely used, and are a relatively low cost selection tool. Despite this, the use of CATs in selection is somewhat controversial because the score distributions of white and minority ethnic groups are systematically different (e.g. De Shon *et al.*, 1998).

The generally accepted figure for differences across white and non-white groups on assessments of general intelligence (*g*) is one standard deviation in favour of white groups (1.0 SD; Ryan, 2001; Hunter & Hunter, 1984). This figure is generally based on the direct comparison of the scores of black and white candidates (Roth *et al.*, 2001), and so may not represent differences across other ethnic groups. Moreover, this figure does not reflect some of the complexities of score differences within certain types of job. In their meta-analysis, Roth *et al.* (2001) report differences of between 0.99 SD and 1.12 SD for assessments of *g* within specific job groups (e.g. education, military, industrial), but lower differ-

ences when results are grouped according to job complexity. Within specific job groups, differences ranged from 0.63 SD for high complexity jobs to 0.86 SD for low complexity jobs. Hough, Oswald and Ployhart (2001) examined differences on specific abilities rather than general ability and reported comparisons between white groups and Hispanics, Blacks and Asians. Their data showed that specific abilities demonstrated lower differences than general ability and that among the ethnic groups, East Asians scored higher than the white group, with differences in favour of the white group tending to be smaller for Hispanics than for the Black group.

Whilst it makes sense to treat the accepted value of 1.0 SD cautiously, the finding of some substantive differences in test scores across ethnic groups is pervasive. What are the effects of these differences in selection practice? Let's take the lowest reported difference as an example (0.63 SD) and assume it applies to a test we are using for selection. If we set a conservative cut-off score of the mean of the white group distribution, then 50 per cent of the white group will meet the cut-off, compared with 26 per cent of the non-white group. The implications of the research findings are that in this selection we could expect half of the white group to pass the criterion, compared with a little over a quarter of the non-white group.

In this example, the effect of applying the test to a diverse group of job candidates means that being in the white group automatically puts you at an advantage. The 'fairness' of this advantage is typically evaluated using the four-fifths rule of thumb. The four-fifths rule states that a selection process (or test in this case) causes adverse impact when the proportion of the lower scoring group of candidates meeting a selection criterion (or test cut-off) is less than 80 per cent of the proportion of higher-scoring candidates meeting the same criterion. In our example, 26 per cent (non-white group) divided by 50 per cent (white group) is equal to 0.52; we would conclude that the test causes adverse impact against the non-white group (because 0.52 is less than 0.80).

Test users and developers justify these differences in test scores by pointing out that tests are good predictors of job performance. The accepted definition of test fairness in US law is that a fair test should predict performance in the same way for all sub-groups of the population. Imagine a situation where candidates from

different ethnic backgrounds tend to achieve different test scores, but the test itself predicts their performance in the job equally well. To resolve the problem of adverse impact, we could adjust the norm-referencing of scores so that equivalent proportions of candidates from the two groups are selected. However, this strategy clearly favours selecting for diversity over productivity, a decision which needs to be made by individual businesses based on their values and needs. The important point is that the fairness of the test rests on demonstrating that it predicts job performance in the same way for the different ethnic groups. This kind of evaluation is rarely conducted in practice.

Procedures to check for unfairness in CATs for selection

The themes from the literature on test fairness in selection lead to two applied questions. First, are the proportions of people meeting the test cut-off equal across different groups? Second, does the test predict performance equally across different groups? Woods (2006) proposed two checks for answering the first question and partially answering the second. Respectively, these consisted of testing differences in mean scores across groups and testing 'regression slope' differences in the prediction of job performance from the test. These were designed to be accessible and practical to implement.

Testing differences in score distributions

This first check for test users to make is based on the differences in score means of different ethnic groups. The process allows two different ethnic groups to be compared and may be applied to as many combinations of groups as are available. First, the groups should be combined and the standard deviation of the combined sample calculated. Next the groups are divided and within-group means calculated. The higher within-group mean is then subtracted from the lower within-group mean. The resultant difference value is then divided by the standard deviation of the combined sample to give a difference between the two within-group means in standard deviation terms.

The calculated difference should then be compared with the values reported in Table 1. These values indicate the maximum difference in group means for the test to avoid causing adverse impact (according to the four-fifths rule of thumb) given a particular test cut-off. These

calculations are possible when abilities are assumed to be normally distributed, and reflect a statistical technique that is discussed elsewhere (Sackett & Wilk, 1994).

To illustrate, imagine that a test was applied in a selection activity with the test cut-off set at one standard deviation above the mean ($z=1$). Using the procedure described above, the difference between two different group means on the test is found to be 0.24 SD. It could be concluded that because the observed difference exceeds the maximum allowable difference indicated in Table 1 (0.11), the test is likely to cause adverse impact against the lower scoring group. The values in the Table also demonstrate that if the cut-off were lowered to $z=0$ (or the mean of the combined sample distribution), adverse impact would be avoided (because the observed difference is lower than the maximum difference of 0.25 SD).

Table 1: Maximum score difference values (standard deviations) between grouped distributions given test cut-off x , to avoid adverse impact.

Test cut-off (x ; Z-Score)	Max difference to avoid adverse impact
1.50	0.11
1.25	0.13
1.00	0.14
0.75	0.16
0.50	0.18
0.25	0.21
0.00	0.25
-0.25	0.30
-0.50	0.37

Testing differences in prediction across groups

As outlined earlier, the widely accepted definition of test fairness in selection is that the test should predict job performance in the same way for all sub-groups of the population. From an analytical point of view, this means that if we test the relationship between CATs and performance in different ethnic groups, the within-group regression lines should be the same (Cleary, 1968).

‘Differential prediction’ of job performance from test scores in two different ethnic groups can be confirmed by examining three points of comparison in the predictive relationship: (1) the regression slope; (2) the regression inter-

cept; and (3) within-group errors of prediction. The method proposed here relates to the first.

Testing for regression slope differences across groups is straightforward, and is described in detail by Cohen and Cohen (1984). The procedure involves conducting hierarchical regression analyses on data sourced from two different groups of the population, with job performance set as the criterion. In step one of the regression, CAT score and ethnic group are entered as predictors, followed by their product in step 2. A significant increase in explained variance (R^2) in step 2 means that the within-group regression lines are unequal, confirming differential prediction.

From a practitioner point of view, there are two problems with this procedure. First, the test of increase in R^2 is a significance test and is, therefore, dependent on sample size. In small samples, even large differences cannot be reliably detected, and in large samples, the significance test becomes highly sensitive. Second, the criterion for deciding whether differential prediction applies is very (overly) precise. Prediction is either equal or unequal, and there is no procedural flexibility to allow the practitioner to determine whether small differences in the regression slope can be tolerated in the selection process. In summary, the procedure doesn’t specify whether differences in prediction actually constitute unfairness in selection.

Woods (2006) proposed that tests of differential prediction should be based on a broad applied principle. CATs are imperfect predictors of job performance, and so applying a CAT for selection inevitably leads to the selection of true positives (people who meet the test cut-off, and perform at the right level in the job) and false positives (people who meet the test cut-off, but not the job performance requirement). The essence of fairness in prediction is that even though some people are incorrectly selected, CATs should identify the true positives equally well for different ethnic groups. By applying the four-fifths rule to this problem, we can specify this more clearly:

Differences in regression coefficients are unfair when they mean that the proportion of true positives among those selected from one population sub-group is less than four-fifths of those from another

By adopting this principle, it is possible to work through the regression equations to determine when differences in regression slopes may be

considered unfair.¹ The results of the calculations are listed in Table 2. The values in the first column represent the higher within-group validity. Values in the second column represent corresponding ‘maximum changes in R²’ from hierarchical regression analyses for the test to avoid unfairness. Practitioners can use this table to evaluate differences in validity across groups.

The broad result of the calculations is a trade-off between test validity and acceptable validity differences across groups. As validity coefficients increase, so larger differences in within-group validities can be tolerated.

Table 2: Maximum changes in R² (ΔR^2) in hierarchical regression analyses to assume fair differences in regression slopes (based on higher correlation R).

R	ΔR^2
0.300	0.038
0.310	0.040
0.320	0.042
0.330	0.044
0.340	0.045
0.350	0.047
0.360	0.049
0.370	0.051
0.380	0.053
0.390	0.055
0.400	0.058
0.410	0.060
0.420	0.062
0.430	0.064
0.440	0.066
0.450	0.068
0.460	0.071
0.470	0.073
0.480	0.075
0.490	0.077
0.500	0.080

Sampling. The values in Table 2 have implications for deciding on an acceptable sample size for testing this component of differential prediction. The smallest value in the second column is 0.038. This means that the smallest value that needs to be reliably detected in the hierarchical analyses is 0.038. Examining the formula for the F-test of significance in regression analyses tells us that the minimum sample required to reliably

detect this effect size is 104 participants. These participants should be arranged into dichotomous groups with roughly 50 per cent of the sample contributed from each group.

How to implement the check for differences in regression slopes. To carry out this check, test users should compile a dataset that comprises test and performance data from two ethnic groups. The hierarchical regression procedure described earlier should then be conducted as normal. If the change in R² in Step 2 is significant, then within-group correlations between test score and performance should be computed and the higher of the two taken to represent the values in column one of Table 2. The ΔR^2 value can then be compared with the second column in Table 2, and where the ΔR^2 value exceeds the value in Table 2, the differences in the regression slopes may be considered unfair against the lower validity group.

To illustrate, suppose an evaluative study used the hierarchical regression analyses and found a significant ΔR^2 value of 0.061, with a higher within-group validity of 0.40. This test would be viewed as unfair according to this method, because the computed ΔR^2 exceeds the maximum value in Table 2 of 0.058.

Summary of the proposed checks

The recommendation from this article is that in order to investigate whether CATs are unfair for selection purposes, test users should perform two checks.

1. Compare the CAT score distribution means from different ethnic groups to assess adverse impact, using maximum score differences shown in Table 1.
2. Conduct a test of differential prediction by comparing the regression slopes of the two groups statistically. Hierarchical regression analysis of data from at least 104 participants (with roughly equal numbers from each group the two ethnic groups) should be computed. A significant interactive effect can be taken as showing that the test has differential prediction for the two groups with values reported in Table 2 determining whether the differential prediction may be considered unfair.

¹ The details of these calculations can be obtained from the author.

Discussion and implications

What are the implications of these procedures for the use of CATs in selection? Firstly, a comparison of the values in Table 1 with the ethnic group differences reported in the research literature shows that almost all CATs are likely to result in adverse impact against some minority ethnic groups. By contrast, values in Table 2 show that in selection, reasonably substantial differences in regression slopes can be tolerated in assessments of fairness. Smaller differences become particularly important when test validity is low. Unfair discrimination is therefore likely to affect some, but not all CATs.

Secondly, the practice of using CATs in selection needs to be examined carefully. Recent statistics indicate that 53 per cent of employers in the UK are using CATs as part of their recruitment and selection activities (CIPD, 2004). There is, therefore, considerable potential for applicant complaints about unfair discrimination, particularly given that in such cases, it would be up to the employer to present data to support their case rather than the complainant. The availability of these data, therefore, needs to be addressed. Test users should request data from test suppliers about ethnic group differences in test scores and validity, or conduct their own validity studies. When data are unavailable, test users should consider avoiding the test in selection activities.

Limitations of this approach

One limitation of the proposed checks is that testing differences in regression slopes is not a complete test of differential prediction. By disconfirming unfairness in these analyses we would not conclude that the test is fair. That conclusion would require additional comparisons of regression intercepts and errors in prediction across the groups. The logical next step to this work is, therefore, the design of standard checks for assessing these aspects of differential prediction with reference to CATs and selection.

Final comments

This article has described two practical checks for assessing whether the CATs are unfair against minority ethnic groups in selection. The checks are not definitive, although are arguably the most accessible alternative to date. If others disagree, then their critique may lead to a better alternative. Either way, test users and test-takers are the beneficiaries.

References

- Chartered Institute of Personnel and Development (2004). *Recruitment, retention and turnover survey 2004*. London: CIPD.
- Cleary, T.A. (1968). Test bias: Prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement*, 5, 115–124.
- Cohen, J. & Cohen, P. (1984). *Applied multiple regression/correlation analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- DeShon, R.P., Smith, M.R., Chan, D. & Schmitt, N. (1998). Can racial differences in cognitive test performance be reduced by presenting problems in a social context? *Journal of Applied Psychology*, 83, 438–451.
- Hough, L.M., Oswald, F.L. & Ployhart, R.E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment*, 9(1/2), 152–194.
- Hunter J.E. & Hunter, R.F. (1984). Validity and utility of alternate predictors of job performance. *Psychological Bulletin*, 96, 72–98.
- Roth, P.L., Bevier, C.A., Bobko, P., Switzer III, F. & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, 54, 297–330.
- Ryan, A.M. (2001). Explaining the Black–White test score gap: The role of test perceptions. *Human Performance*, 14, 48–75.
- Sackett, P.R. & Wilk, S.L. (1994). Within-group norming and other forms of adjustment in pre-employment testing. *American Psychologist*, 49, 929–954.
- Salgado, J.F., Anderson, N., Moscoso, S., Bertua, C., De Fruyt, F. & Rolland, J.P. (2003). A meta-analytic study of general mental ability validity for different occupations in the European Community. *Journal of Applied Psychology*, 88, 1068–1081.
- Schmidt, F.L. & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: 85 years of research findings. *Psychological Bulletin*, 124, 262–274.
- Woods, S.A. (2006) Cognitive ability tests and adverse impact against minority ethnic groups: A proposal for standard evaluative procedures. *Proceedings of the Division of Occupational Psychology Conference, 2006*.

Correspondence

Dr. Stephen A. Woods

Institute of Work, Health & Organisations
(I-WHO)

The University of Nottingham,
8 William Lee Buildings,
Nottingham Science and Technology Park,
University Boulevard,
Nottingham NG7 2RQ.
Tel: + 44 (0) 115 846 6483
E-mail: Stephen.woods@nottingham.ac.uk

Stephen Woods is a Lecturer in Occupational Psychology at the Institute of Work, Health and Organisations, at the University of Nottingham. Stephen's work in the Occupational Psychology Research Group focuses on Personality Assessment and Psychometrics.

Research to investigate the procedures described in this article requires a collaborative sponsor, who will contribute £12K to the project over three years.

Further information can be obtained from Stephen.woods@nottingham.ac.uk.

This article was prepared based on the Focused Paper: 'Cognitive ability tests and adverse impact against minority ethnic groups: A proposal for standard evaluative procedures' presented by Stephen at the Division of Occupational Psychology Conference, January 2006.



CALL FOR SUBMISSIONS

Division of Occupational Psychology
Annual Conference 2007

10 – 12 January

Bristol Marriott Hotel City Centre

The Conference Committee welcomes submissions, from both academics and practitioners, which address the broad range of topics within occupational and organisational psychology. Empirical studies, theoretical papers and practitioner case studies are all encouraged.

Further details including the submission guidelines can be obtained from the BPS website, www.bps.org.uk under the Conferences and Events section.

All submissions will now be online. The closing date for all submission is Monday 7 August 2006.

For further details about the conference please also see the BPS website where information will be continually updated.

All enquiries should be addressed to the BPS Conference Office. Tel: 0116 252 9555.
E-mail: dopannual@bps.org.uk

The psychology of 360 degree feedback

EVERY DAY THOUSANDS of 360 questionnaires and reports are generated and fed back to slightly apprehensive managers, sales people, specialists, etc. But, no one has a clear idea about how 360 works. What psychological mechanisms are involved? How do we explain 360's proven ability to change actions and performance?

In this article we attempt to provide answers to those and other questions. We offer a helpful theory that should explain (and perhaps challenge) your use of 360. You may be able to develop a better 360 feedback process.

All theory begins with an assumption, and ours is that people are constantly checking their perceptions to see if they are true or false. That's the assumption made by Personal Construct Psychology (see Fransella, 2005, for a great introduction) which offers a hugely practical approach to human action that is intuitively reasonable. You and I base all our decisions and actions on how we see and understand things. Perceptions are all we have to guide us; we depend totally on them. It makes sense to test and checkout perceptions whenever we can. Perceptions, hunches and assumptions that are disconfirmed by our experiences probably need changing. But, as we will see later, that's not always easy.

Managers, leaders, salespeople, etc., are no different. Perceptions about markets, technologies, problems, opportunities and people shape their decisions. Good managers use a multitude of sources to check their perceptions – reports, presentations, discussions, publications, data, etc. Indeed, a characteristic of all good managers seems to be concerted, ongoing attempts to challenge their views, to think differently, to try new ways of seeing things.

But, for nearly all managers there is one set of perceptions that is based on very little data. They are perceptions which are very difficult to check, and hence easy to get wrong. Surprisingly, they are a crucial set of perceptions; ones which underpin strong personal performance. They are the individual's perceptions of him/herself.

Peter Goodge & Jane Coomber
the360.co.uk

Ask any manager for information on almost any aspect of the business and he/she can produce reports, statistics, examples, etc. Yet, ask about his/her skills and approach, and he/she will probably have very little information, and almost certainly nothing that can be relied upon.

Why do we know so little about ourselves? There seems to be a root cause in our business culture; we don't honestly tell one another how we are doing. In other areas of life we do honestly tell; in fact we welcome it. Alice, Peter's horse riding instructor, is great because she both challenges and supports his perceptions with relentless precision. In business we hardly ever do that; even feedback in the safe environment of the training course is soft, diplomatic and imprecise by Alice's standards.

Good 360 feedback provides clear, concise information about the individual. That's a rare thing in business, and it enables individuals to check and clarify their perceptions of their skills. Often 360 disconfirms self-perceptions; the feedback report isn't entirely what was expected. By challenging our perceptions 360 leads individuals to rethink themselves and rethink their performance, and it seems to work in three ways.

- a. *Negation* – feedback results are the opposite of expected strengths or weaknesses. For example, an individual may believe that all his/her team rate their teamwork highly, yet the 360 results show some people rate teamwork poorly. Or, the individual may discover his/her critical (low) self-rating of their planning is not shared by others who rate his/her planning favourably.
- b. *Significance* – results confirm a self-perception but show it is much more (or less) important than the individual thought. For instance, an individual may discover everyone rates his/her time-management poorly, not just

himself. Or, it is not just one aspect of leadership that is strong; it is several aspects.

- c. *Explanation* – results highlight unexpected relationships. For example, where great attention to detail produces a controlling management style, or relatively weak business understanding produces poor business presentations.

A 360 skilled coach or workshop facilitator will be looking for all three types of disconfirmation, and making full use of them. It is Explanation that seems to have the biggest psychological impact, perhaps because it challenges the person’s perceptions of how the world works. Inexperienced 360 coaches tend to expect only Negation, and they miss some of the most helpful parts of a 360 feedback report.

The theory helps explain some of the ‘do’s and don’ts’ of 360 feedback.

Design clear, relevant questions and feedback reports

Long, ambiguous and vague questions do not change thoughts or actions. They can be interpreted in many ways, and hence neither support nor disconfirm perceptions. Some examples:

- *Consults with his/her team as appropriate* invites different interpretations of the word ‘appropriate’;
- *Does not inhibit others from sharing ideas* contains a double negative, which tempts respondents to make a mistake;
- *Works hard at getting on with people and planning things in detail* is ambiguous; it asks about two different things.

Atwater *et al.* (1995) found clear, relevant questions to be absolutely critical.

Similarly, lengthy statistical feedback reports using graphs and averages can be read in different ways. And, their complexity adds to the workload of individuals receiving feedback. Concise, simple, visual feedback reports have much more impact; their implications for the individual’s perceptions are immediately clear. Figure 1 is a snippet from a 360 feedback report showing all the ratings given by each respondent on ‘Implementing best professional/technical practice’.

Visual impact is achieved by putting the lower ratings in red backgrounds and higher ratings in green backgrounds. In this case the individual (S) rates him/herself as average, whereas the direct reports (D) have given high ratings, except one who did not feel he knew the recipient well enough to answer the question. Of two colleagues (C), one has given a high score; the other did not answer the question. Of two managers (M), one has given a high score, and the other a low score. These ratings challenge the individual’s self perception and invite him or her to explore the meaning behind the feedback. Just showing the average score, even breaking it down for groups of respondents, would have hidden important information.

Provide a large sample of credible completed questionnaires

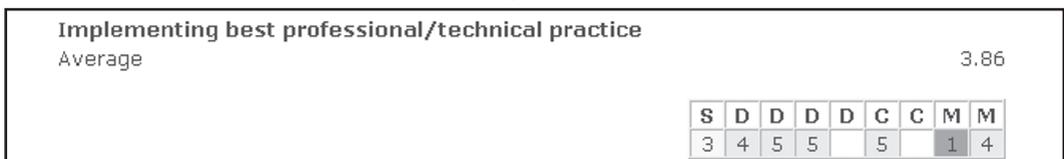
Research suggests that a feedback report based on more questionnaires is more credible (London & Smither, 1995). In practice that seems to be about 10 questionnaires. Reports based on just five or six questionnaires often lack sufficient information to disconfirm perceptions. For some individuals, especially technical specialists, with few if any direct reports, this may present a difficulty. However with today’s internet-based 360 systems it may be feasible to ask people outside the organisation for feedback, customers or suppliers for example.

Engineer some critical feedback

Unless the 360 system checks questionnaires, to ensure a proportion of critical ratings, only five per cent to seven per cent of answers will be critical. Nearly all the feedback will be mild praise, which will raise individuals’ self-esteem and reinforce existing behaviours. It will not disconfirm perceptions, and will not create change (for the research see Meyer, 1980).

Therefore, if you are using feedback to change skills, culture, or performance you will need a 360 system that ensures a minimum of critical feedback is given by those completing

Figure 1.



questionnaires. Our experience shows that where the 360 system requires a proportion of critical feedback, this critical feedback will not be given arbitrarily, or 'just for the sake of it'; those completing questionnaires will pick the poorer competencies and questions. So, such a system may sometimes exaggerate critical feedback somewhat, but it won't be scattered randomly. It also helps to take care with briefing people, clarifying their expectations before they come to complete the questionnaire. One should also use a rating scale which makes it clear that people are being rated on a competency in comparison with the other things that they do, rather than in comparison with other people.

Don't think feedback will change the worst performers

Goodge and Burr (1999) found that 360 feedback had more impact on below-average performers, yet recognise that the worst of performers rarely change as a result of feedback. How do we explain this?

Personal Construct Psychology suggests that perceptions don't exist in isolation, unconnected with one another, but are linked together in elaborate networks. And, that some of our self-perceptions are 'core' – they hold together many of our other perceptions about ourselves. Not surprisingly, core self-perceptions are difficult to change – so many other beliefs about ourselves depend on them.

The worst of performers seem often to be people who have been in their jobs for some years, and failed to change with the times. They frequently have fixed perceptions of their job, which are closely linked to their core self-perceptions. That may be why the poorest of poor performers don't respond to feedback. For them feedback may question the kind of person they believe themselves to be, and that's not easily changed for any of us!

Personal Construct Psychology theory also suggests what's important in the feedback process; it is subtle but easily engineered. The moment of real personal change is when an individual states that his/her interpretation of the feedback is inconsistent with their beliefs. It is a moment that is facilitated by exploring beliefs prior to seeing the feedback, and also inviting the individual to say what the feedback means.

Hence, it is enormously helpful to:

- a. Spend some time, perhaps 30 to 45 minutes, at the beginning of a feedback session exploring the individual's role, performance and concerns. That initial discussion not only informs the facilitator, but clarifies existing relevant perceptions for the individual.
- b. Encourage the individual to work out what the feedback means for themselves, aiming for simplicity (perhaps looking for just a few strengths and weaknesses).

Perhaps the thing to remember is that only an individual him/herself can disconfirm their own perceptions; others can not do it for them, or force it upon them. So, if you find yourself explaining or telling an individual what their feedback means, expect little or no change in perceptions or performance. A helpful facilitator says very little, but engineers a lot!

References

- Atwater, L., Roush, P. & Fischthal, A. (1995). The influence of upward feedback on self- and follower ratings of leadership. *Personnel Psychology, 48*, 34–59.
- Goodge, P. & Burr, J. (1999). 360 degree feedback – for once the research is useful. *Selection & Development Review, 15*(2), 3–7.
- Fransella, F. (2005). *The essential handbook of personal construct psychology*. Chichester: John Wiley and Sons Ltd.
- London, M. & Smither, J. (1995). Can multi-source feedback change perceptions of goal accomplishment, self, evaluations and performance related outcomes? *Personnel Psychology, 48*, 803–839.
- Meyer, H.H. (1980). Self-appraisal of job performance. *Personnel Psychology, 33*, 291–295.

Peter Goodge and **Jane Coomber** are partners with the360.co.uk, a specialist provider of online 360 degree feedback solutions. Their careers span psychology, personal development, and IT. For more information, including a new handbook on 360 degree feedback, see www.the360.co.uk.

The case for coaching

DURING THE 1980s, the term 'coaching' was virtually non-existent within the business context, although I'm sure it was being performed indirectly. However, coaching is now very firmly a buzz word in many forward thinking companies and organisations.

Coaching is the art of accessing and then improving the performance, learning and development of another. Corporate coaching is one of the most popular areas of coaching. As one of the most effective tools to use in business to improve performance and productivity, it differs enormously from training and development programmes. This is because the individual being coached takes control of their actions and makes the changes required to improve.

Whilst only a coach can gauge how coaching impacts on the bottom line within their niche area, there are certain benefits that corporate coaching can give in general. It works so well because the coach can provide an objective viewpoint in any coaching discussion. Employees who do have any concerns over confidentiality can be assured that they are being coached as a benefit to their well-being first – and the company profits second. In particular, corporate coaching can achieve results in the following ways:

- Improved performance and productivity;
 - Staff development and self-reliant staff;
 - Improved learning;
 - Improved internal relationships;
 - Improved quality of life for individuals;
 - Greater flexibility and adaptability to change.
- It is now widely acknowledged that people are the most valuable resource of any company or organisation. However, whilst reports from the 2004 CIPD Training & Development Survey found that 99 per cent of respondents agree that coaching can deliver tangible benefits both to individuals and organisations, only 14 per cent claim that coaching skills training was compulsory for all staff that managed people. This statistic underlines the fact that many companies are missing out on the significant competitive advantage of a happy, motivated and valued workforce who are more productive.

John Toscano

Coaching, like most good things, can be a very simple yet highly effective process. It can also be a source of delight when outstanding results are achieved. I work with many private clients who value their own personal development and want to make changes in their lives. They want to become motivated, focused and to achieve their own personal goals – whatever they may be. They appreciate the results that a professional coach can help them achieve, the way their coach will help them to acknowledge their true values, remove their limiting beliefs and yet, never command them to do anything.

Similarly, I work with managers in organisations who are looking for more effective ways of working with colleagues. They are aware that the more traditional approaches to managing people are becoming less and less appropriate. They are, maybe, accepting that the methods they have adopted that made them successful managers in the past may now be halting their own progress – and that of their staff. The primary reason that coaching is so effective in improving a company's productivity and performance is because it is totally people focused. Unlike computers and machinery, human beings are unpredictable for a whole host of reasons. Coaching works because it gets to the root of these issues by being solely individual – centered and holistic. This is in contrast to the traditional instructor-led, 'one size fits all' and 'sheep dip' approach to learning and development.

It must also be stated that coaching is not purely a remedial process, and is invaluable for people at all levels. Many organisations have come to understand that on-going personal development through coaching is essential to fully utilise its 'human capital resource'. In a rapidly changing environment, even top performers will benefit from the chance to reflect upon what works well for them, how to remain positive and to sustain excellence.

Therefore, coaching really does provide a winning scenario for the employee and the company. In addition to the benefits discussed above, staff will certainly feel more valued, productive and understood, as well as having the benefit of a confidential and impartial 'sounding board'. The company benefits from a happier, more motivated and higher performing workforce. The outcome is a powerful combination that promotes trust, confidence and loyalty within any organisation. What is this worth to your company? When *Fortune* magazine surveyed 100 executives from 'Fortune 1000' companies in the US, who had received coaching, managers collectively described an average return of \$100,000 – many times what the coaching cost their companies. So next time you hear a company proclaim that 'people are their greatest resource' you will have no hesitation in believing it to be true!

Correspondence

contact@johntoscano.co.uk
Tel/Fax: +44 (0) 29 20401440.
Mobile: 07812 063675.

SDR

Spurilitarianism

MARK PARKINSON'S article on utility analysis (UA) (*SDR*, Vol. 21, No. 6) is a timely reminder that these tools have to be used with care to avoid spurious results. He points out correctly that the system is less easy to apply in multi-stage selection processes but does not address more fundamental objections to the way UA is calculated and currently used. Selling a psychological process either by promising net gains or, as he suggests, with an eye to the loss averse, by pointing out losses that will accrue if the system is not used, is only defensible if the gains or losses can be accurately quantified. This is not necessarily the case.

Convincing an employer of the usefulness of a particular test by promising a net gain arrived at by UA assumes: (a) that they have previously employed a wholly random means of employing staff; (b) that in future they would rely wholly on this suggested (single) method; and (c) that the validity coefficient of that method accurately indicates its likely effectiveness.

Credibility in the claims made for the tool has been increased over the years through developments of SDy calculation (Cascio & Ramos, 1986; Schmidt & Hunter, 1983) the addition of finance variables (Boudreau, 1983), techniques such as breakeven analysis ((Boudreau, 1984), relaxing the assumption of top down hiring (Murphy, 1986), determining the effects of recruiting efforts (Rynes & Boudreau, 1986) and allowing poor performers to be dismissed (De Corte, 1994). Despite this body of work, and apart from the first, few of these developments are commonly used in calculating UA, possibly because they all reduce the total gain shown by employing a selection device.

Perhaps fortunately, there is little evidence that business decisions are in fact based on the tool, and less to show that it can be presented to employers as a reasonably accurate measure of quantified profit. Irrespective of the particular factors listed above, UA estimates may still overstate the case, because current models only concern themselves with a single criterion variable, because efficacy is compared with random selection rather than an existing process, and because

Glen Fox
OPS Ltd.

the validity of the test, rather than its variance, is still being used in most equations as a percentage measurement of its effectiveness.

Although current models assume the validity coefficient of the test or process will indicate its efficacy if it is multiplied by the SDy gain, this may not bear examination. A validity coefficient is a correlation with a criterion, not a percentage indicator, and it is the variance which may more properly be used to indicate how much a process will add in quantifiable terms. This is not the case with a perfect process; a validity coefficient of 1 when multiplied with a notional gain of, say, £10,000 for appointing an individual required to be 1 SD above the mean, will of course yield a net gain of £10k if the perfect test is used on each person appointed. However, if the validity coefficient is 0.5 (variance 0.25), multiplying the SDy by 0.5 will produce an apparent profit per person appointed result of £5000, or 50 per cent, which is unlikely to be the case. Utilising the variance will more appropriately show a net gain of £2500.

More commonly voiced objections relate to the criterion employed in calculating the validity coefficient of the test, which may not be related to or relevant to performance and/or profit. A validity coefficient arrived at by calculating construct validity in relation to a similar test, or predictive validity in relation to, say, promotion rates within the Armed Forces, may not convincingly relate to performance/profit generation in sales executives in an fmcg (fast moving consumer goods) organisation.

Although most organisations gather information from multiple sources when recruiting or promoting staff, utility analysis measures only a single variable and compares its use with random selection. In reality, the more likely scenario adds a new selection tool to a given selection process, rather than random selection, so the incremental validity of the new selection tool

needs to be calculated, rather than assuming a comparison with a random process. In addition, few organisations limit their selection process by using a single variable, so the model needs to be expanded to take account of multiple predictors, such as interviews, references, application forms, tests, presentations and exercises, which are all routinely employed.

In reality, UA should only be used by computing the utility of previous methods used, then computing the utility of the proposed method(s), expanding the equation appropriately to take account of a set of several predictors, giving due weight to incremental validities, if, as is usual, multiple processes are employed, and to their weightings, and then subtracting that figure from the utility of the new method(s) (Sturman & Judge, 2006). This would allay fears that we are attaching excessive importance to tests/predictors which actually duplicate some or all of each other's value, or which are not all equally valued in the selection process.

Little research has been done on how much attention is actually paid to UA in making business decisions, and it would be interesting to know if the weaknesses of UA have affected its use in the business arena.

As occupational psychologists, we need to be as careful and accurate in our marketing as in our professional practice, and promising exaggerated monetary gains to clients does nothing to convince them of our concern for high professional standards. Utility analysis has its place in the toolkit, but it should be used with care as a tool for producing accurate estimates of gain rather than as yet another example of hyperbabbble.

References

- Boudreau, J. (1983). Economic considerations in estimating the utility of human resource productivity improvements. *Personnel Psychology*, 36, 551–557.
- Boudreau, J. (1984). Decision theory contributions to HRM research and practice. *Industrial Relations*, 23, 198–217.
- Cascio, W.F. & Ramos, R. (1986). Development and application of a new method of assessing job performance in behavioural/economic terms. *Journal of Applied Psychology*, 71, 20–28.
- De Corte, W. (1994). Utility analysis for the one cohort selection-retention decision with a probationary period. *Journal of Applied Psychology*, 79, 402–411.
- Schmidt, F.L. & Hunter, J.E. (1983). Individual differences in productivity: An empirical test of estimates derived from studies on selection procedure utility. *Journal of Applied Psychology*, 68, 407–414.
- Sturman, M. & Judge, T. (2006). *Utility analysis for multiple selection devices and multiple outcomes*. Working Paper Series, Centre for Advanced Human Resource Studies, New York.

Correspondence

Dr Glen Fox

Principal Psychologist,
OPS Ltd.

E-mail glen@opsLtd.com



The
British
Psychological
Society

The Special Group in Coaching Psychology

We are pleased to present further details of our 2006 Event Programme. These events will offer opportunities for further training and development in topics which have been carefully selected to complement coaching psychology. We really look forward to seeing you at one or more events.

30th June 2006:

'Managing Personal and Client Stress in Coaching Psychology'

Facilitator: Jenny Summerfield, CPsychol.

This one-day event will be held at the Work Foundation, 3 Carlton Terrace, London, SW1Y 5DG and will run from 10.00am to 5.00pm.

Registration from 9.30am.

7th September 2006:

'Positive Psychology in Coaching Psychology'

This one-day conference will be held at the BPS London Office and will run from 9.30am to 4.30pm. Registration from 9.00am.

For further details and information about these events see the 'News Page' of the BPS SGCP website on: <http://www.coachingpsychologyforum.org.uk>. For booking information please contact: Tracy White, Email: tracy@virtuallyorganised.com

18th & 19th December 2006:

'SGCP 3rd Annual National Conference', City University, London, UK.

This 'not to be missed event' will offer a stimulating and thought-provoking mix of keynote speeches, individual papers, workshops, and round-table discussions, delivered by national and international leaders in the field.

Parallel research paper and poster sessions will also provide some of the latest thinking and research in Coaching Psychology.

Further details will soon be announced on the SGCP website.

The 2006 membership fee to join SGCP is £3.50. BPS members can join and immediately benefit from member rates at events.

Making Waves – Saville Consulting Wave Styles questionnaires

THE SAVILLE CONSULTING (SC) WAVE STYLES questionnaires have been developed by Saville Consulting, an entirely new business founded by Professor Peter Saville. The vision of the company is to transform assessment and the first tranche of new products have now been launched including a range of aptitude tests and culture assessment tools.

The SC Wave Styles questionnaires are central to the integrated development strategy. The individual measures (the styles questionnaires) of the SC Wave suite and the organisational measures (culture and climate) have been developed from one integrated model.

This article discusses the concepts and development that formed the backbone of the individual measures: the SC Wave Styles questionnaires.

SC Wave Styles – An overview

SC Wave Styles presents a uniquely integrated model of personality, competency, culture and motivation. The concepts of motive and talent are integral to the structure and have real implications for individual development, career planning and performance management as well as selection.

Rather than taking a paper questionnaire and putting it on the internet, SC Wave Styles questionnaires are an innovative suite of self-report measures developed with the opportunities and challenges of the internet at the heart of their design. They operate as both trait and type instruments and rely on a new hierarchical model of work performance that we have developed. Our model is aligned to the Big Five personality factors and the Great Eight competencies but provides more information than either model.

The development of the questionnaires has benefited from a performance driven or valida-

R. MacIver, P. Saville, R. Kurz,
A. Mitchener, K. Mariscal, G. Parry,
S. Becker, W. Saville, K. O'Connor,
R. Patterson & H. Oxley
Saville Consulting Group

tion centric methodology to maximise the validity of the questionnaires by selecting the most valid items from our item pools.

A new dynamic online format integrates rating and ranking responses and results in a combined profile that highlights differences between the ipsative and normative scores on the profile. This new scaling technology also allows unprecedented levels of detail to be tapped, yet with radically reduced completion times.

The research matching the questionnaire to the preferred culture, environment and job demands allows individuals (and their managers) to gain new perspectives on what they can take from their job and what will motivate them.

SC Wave Styles questionnaires are different in several respects from established psychometric questionnaires. Below we provide an overview of the new features and approach before taking a brief look at the reliability, validity and applications of the questionnaires.

The motive talent concept

SC Wave Styles questionnaires have been developed to separate out talents from underlying predispositions or motives. For every work construct measured there is one motive item and one talent item.

The Expert Report profile indicates where 'motive-talent splits' occur. This allows the user to identify whether the individual is motivated to

develop in a particular area or needs their talent in an area supported or encouraged because there is a lack of underlying motivation or interest.

The approach has profound implications for the development of individuals. Because the model matches every talent item with an underlying motive item, it is much simpler to understand the impact of motivation on work performance

Performance driven content

SC Wave Styles questionnaires have been developed (and continue to be developed) using a variety of development strategies, but at the core is a validation centric strategy.

As Burisch (1984) points out mixed approaches, which include validity data as part of scale development, are surprisingly rare in questionnaire development.

'Actually this is rarely done, particularly the combination of deductive scale writing and external information for item analysis.'

For the initial development of the Professional Styles questionnaires 214 work constructs were written (each with separate motive and talent components. 428 work constructs in total). 108 of these constructs (facets) made it into the final questionnaire with item selection based first and foremost on criterion validity. Items were correlated with external ratings on relevant work behaviour competencies as well as overall job proficiency and potential for promotion.

The SC Wave Styles questionnaires are, therefore, based on the work constructs which are the best indicators of performance (i.e. performance driven).

The facet approach to measurement ensures that each item in the questionnaires measures a different work construct to help avoid the feeling of 'needless repetition' that respondents can experience when completing questionnaires.

Clearer interpretation

One of the criticisms that can be fairly levelled at self report questionnaires is that despite their reliability and validity there is a degree of subjectivity in their interpretation.

Despite users being trained prior to questionnaire use, subject matter experts in assessment believe that poor interpretation is a significant source of error in the use of personality questionnaires (Smith & Foley, 2006). A lack of consistency between interpreters is much more

likely where an aspect of work performance is predicted by a complex combination of predictor scales which is the situation with many multi-scale self-report personality instruments.

Is inconsistency a given? Can we do anything about it? With the performance driven approach, the work constructs that best predict a work competency are brought together to form a scale. This largely removes the need to look around the profile for what scales relate to a particular competency (i.e. we move from predictor centric models to criterion centric models of work performance and have to work less hard to join the dots).

The feedback provider also no longer has to guess what an average score overlooks, where there are differences in scores underpinning a dimension these are highlighted in the profile.

Better interpretation, inevitably leads to improved validity in decision making based on questionnaire data (i.e. better decision making, fewer selection errors and better identification of development needs).

Dynamic normative-ipsative format

There are advantages and disadvantages to both normative and ipsative response formats. At the practitioner level, it is useful to have both sources of information. The online SC Wave Styles questionnaires firstly present a group of six normative statements. The system then calculates their ipsative rankings based on the order the statements have been rated (this saves the individual time by calculating the majority of rankings for an individual). Where an individual ties their ratings the system immediately represents the tied items to be placed in rank order.

Where there are differences between normative and ipsative scores these are highlighted on the profile to allow the user to explore the reasons for the difference (which of the two scores is most representative of the true score and which is more distorted). The user can then focus on specific areas where socially desirable responding (or overly self critical responding) may have occurred.

New levels – new lessons

SC Wave is a model with several levels in its hierarchy. The new scaling technology allows much more to be assessed in less time.

At the top of the Professional Styles model are four clusters which also provide the basis of the Saville Consulting Types model which profiles People Type and Task Type.

Each cluster breaks down into three sections giving a total of 12 sections. Each of the Big Five and Great Eight constructs has a counterpart with one of these sections and four further areas are covered.

Each of the 12 sections covers three themes giving 36 dimensions which is the fidelity level of many occupational personality measures.

The newly developed scaling allows SC Wave to go down one more level with each of the 36 dimensions being composed of three facets giving a total of 108 facets.

The 216 questions are usually answered in around 40 minutes. The ultra-compact Wave Focus questionnaires measure 36 facets selected for their strong validity in just 15 minutes.

The new dynamic 'Ra-Ra' (Rate-Rank) format combining ipsative and normative scores has profound implications for measurement – it is simply not necessary to have six or eight items in a scale when two items can do as well or better (i.e. strong validity, strong reliability, good spread of scores, etc). More items are required

to achieve this with conventional five-point Likert items or ipsative-only questionnaires (MacIver, 1996).

The profile highlights within any of the 36 dimensions where there is a significant 'facet range', i.e. one or more of the three underlying facets having markedly different scores from the others.

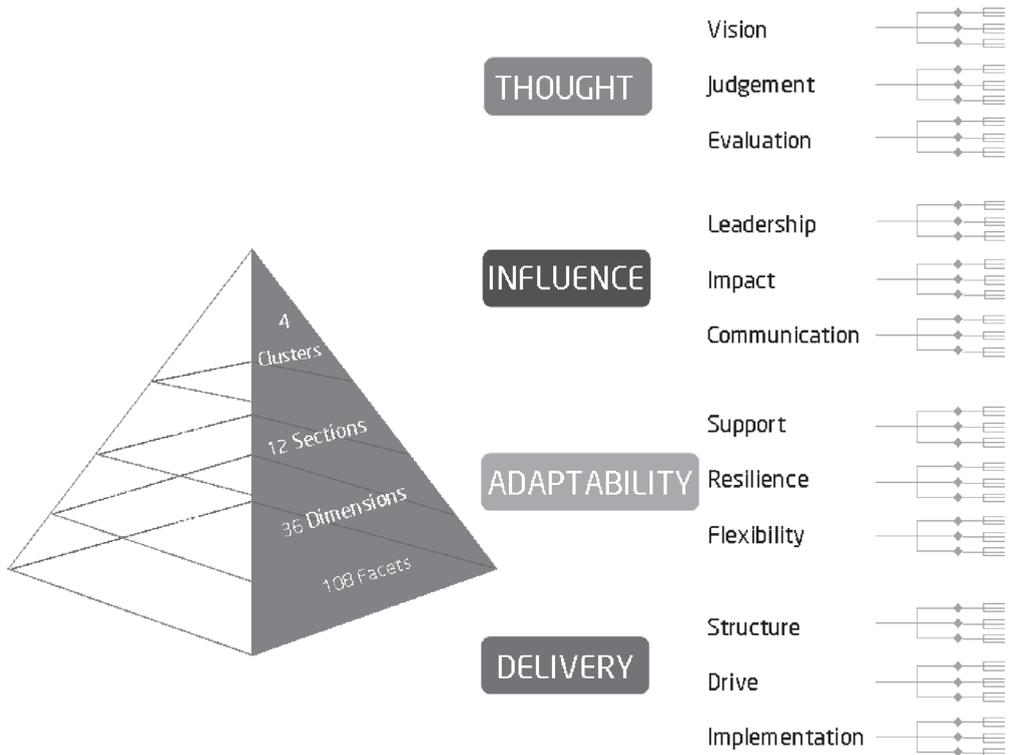
As well as many more individual scales being measured this provides the user with much more detailed insight into the 'uniqueness' of the individual.

Configurable competency reporting

The level of detail that SC Wave Styles achieves also enables a more detailed match with client competency models allowing for fast configuration of output reports to predict client competencies and even their competency indicators.

This detailed configuration has also enabled the Entrecode® model of successful entrepreneurs developed by Professor David Hall and his associates to be available as a separate report.

Figure 1: The Saville Consulting Wave Work Hierarchy.



Enhanced security

The internet offers great convenience in allowing individuals to respond at great geographical distance (without an administrator present) by sending a link to the questionnaire directly to an e-mail address. This so called 'controlled' mode does present security concerns. For example, not being sure the questionnaire is completed honestly by who you have sent it to (and not say, by a group of the candidate's friends one evening). We believe strongly that as well as 'controlled' (or 'invited access' forms) a self report questionnaire (particularly questionnaires that can be used for selection or other decision making processes) should have a separate supervised secure form. Therefore, two parallel forms of SC Wave are available for 'Invited Access' and more secure 'Supervised Access' administrations.

Culture match

The SC Wave development programme has also developed measures of work culture that are parallel to the SC Wave Styles model. This empirical research allows us to indicate the preferred culture/environment and job demands that would suit an individual based on completion of SC Wave Styles questionnaires.

From the perspective of Positive Psychology, Seligman (2003) has argued that work can be changed to suit the employee (rather than just finding an employee that fits the job or trying to develop the individual to better match/meet job demands). Assessment can be constructed to support this approach.

With our unique model (See Figure 2) which ties together motive, talent, competency and culture, we can help individuals understand what work demands (culture, job and environment) they are most likely to favour.

Armed with this understanding it becomes easier to discuss what enhances or inhibits individuals' performance at work and therefore facilitate constructive discussions about how a job could better reflect a person's motives and talents.

This approach can also help managers think about how to tailor work to suit individual employees to keep them satisfied and motivated.

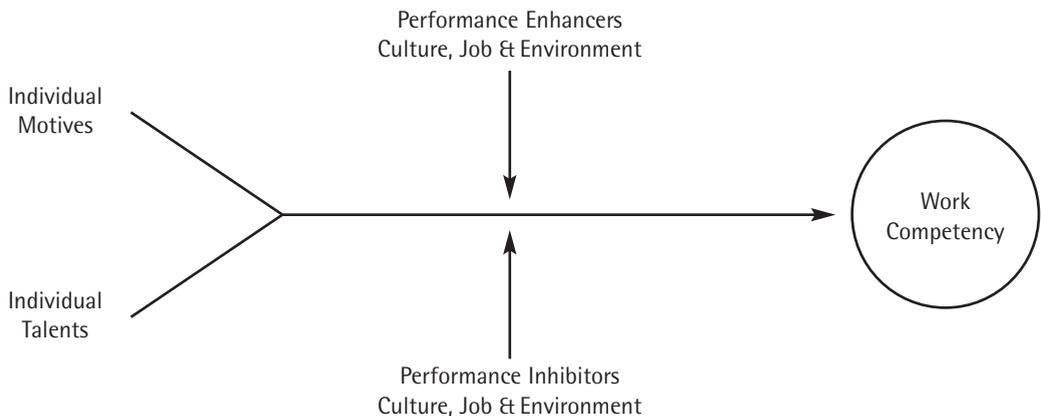
Reliability

A development goal of the SC Wave Styles questionnaires was to have alternate form and test retest reliability estimates as high as possible. In contrast, the SC Wave Styles questionnaires were designed to have moderate (0.6 to 0.8) rather than high internal consistencies at the dimension level (as they are made up of six different work constructs).

The alternate form reliability average across the two forms of the Professional Styles questionnaires is 0.86 with the lowest reliability 0.78 and the highest 0.93 ($N=1153$).

The highest correlation with a scale in one form was with its respective twin in the other form (e.g. Inventive in Invited Access form correlates most highly with Inventive in Supervised Access form not with any other scale).

Figure 2: The Saville Consulting Model of Work Performance Effectiveness.



The internal consistencies of the dimensions of the two Professional Styles questionnaires range from 0.58 to 0.87 with an average of 0.76 ($N=1153$).

Test retest average is 0.79 across the dimensions of the two normative Professional Styles questionnaires with an interval of one month. The lowest reliability was 0.71 and the highest ranged up to 0.91. ($N=112$).

Validity

SC Wave has already been validated on over 1500 people globally during its development, standardisation and initial use. Gathering validity on every item continues on an ongoing international basis.

The validation results from the development trial were cross validated in the standardisation trials of the questionnaire. As each dimension has been designed in the development trial to predict one of 36 work competencies, there were clear a priori hypotheses for the standardisation trial of which SC Wave Style dimension will correlate with which work competency.

In the standardisation 34 out of the 36 dimensions had significant correlations ($p<0.05$, one-tailed) with their matched criteria across the two forms. The combined ipsative and normative dimensions predict external ratings (manager or colleague) of effectiveness on the work competency that each was designed to predict with an average of 0.39 for the Invited Access form and 0.39 for Supervised Access. These figures are corrected for criterion unreliability but no other corrections have been applied, e.g. restriction of range, predictor unreliability, etc. ($N=556-658$).

Validity figures based on multiple regression equations cannot be used as an unbiased estimate of the validity of the questionnaire when they are not cross validated. The mean validity of the composite equation of SC Wave Styles dimensions in relation to the 36 work competencies is an average of 0.46 for both forms. This was based on developing the equation on one half of the standardisation data and cross validating the equation to the second half of the standardisation sample. These figures are corrected for criterion unreliability but no other corrections have been applied, e.g. restriction of range, predictor unreliability, etc. (Hold out Sample $N=252-316$).

Reporting

The profiles are not simply static text describing the content of a scale. Instead the text in the pro-

file feeds back the individual facets (giving different verbal descriptions for the different score (Sten) positions). This reduces the thinking time for the person giving feedback as they do not have to think how to explain the score. It also means that the recipient of the feedback understands the content of the profile much faster with less need for explanation.

Two of the main reports provided from SC Wave Styles are the Expert and the Types report. Each report comes with a free personal report for the respondent.

The Expert Report:

Summary of Individual's Response Styles;
Full Wave Psychometric Profile (with Motive Talent Splits, Ipsative-Normative Splits and Facet Ranges);
Culture/Environment Prediction Report;
Competency Potential Report.

Types Report:

The Types report brings together perspectives on teams, leadership and management in one straightforward individual differences model of performance at work. These are performance driven types that predict external work performance as they are based on our performance driven methodology.

Firstly, individuals are classified by their People Type which can be one of four: Individualist, Influencer, Adaptor or Transformer. Secondly they are classified as having a Task Type, again one of four: Preserver, Thinker, Doer or Transactor.

This leads to their overall Saville Consulting Type which is made up of an individual's People and Task type combined, e.g. Influencer Doer. This is followed by a bullet point summary of how they will tend to lead, work in teams and manage change.

Applications of Saville Consulting Wave

Saville Consulting Wave Styles are designed to be applied throughout the employee's lifecycle. Below are some of the applications of SC Wave Styles with an insight into where they make a difference.

Selection

SC Wave Styles are designed to create a platform for much better decision making from a self report questionnaire, leading to increases in the calibre of employees. As a feed into interview or

as a source of data to cross reference with other data and pull together a more coherent picture of an individual, SC Wave gives more valid data in less time. The detail in SC Wave also allows bespoke reports to be created with a better content match to client competency models.

Career planning

In thinking about how to manage the future of an individual's career it is useful for the individual to understand what areas they are interested in developing (insight from motive-talent splits) and what type of culture will enhance or inhibit their success and motivation (culture prediction report). By giving this unique perspective SC Wave allows an individual a perspective on what they want from their work in the future.

Coaching and development

SC Wave Styles provides insights which are useful to the person being coached and provides a clear link to understanding the impact of their personal style (motive and talent) on their performance at work. Facet splits provide more detail and interesting contrasts that lead to a precise understanding of the individual's approach to work. In development it is also extremely useful to know where an individual has an internal incentive or motive to develop or has little internal incentive or motive to improve performance (which is provided by motive-talent splits).

Self-selection

Self-selection can be aided by highlighting to individuals where they have more or less fit with job demands and culture. It may be that self selection happens before a formal application is made by candidates, or as part of the selection process itself. By supplying each candidate with a culture prediction report we can enable individuals to have their own perspective on whether they feel they are suited to the role in question.

Individualised induction

SC Wave Styles can be used in induction to give an individual a view of what they would like to get out of their work. Despite new starters tending to be satisfied in general, their satisfaction ratings are relatively weak when they are asked about the feedback of (or lack of feedback

of) assessment data collected during their selection (Miles, 2006). The culture prediction report provides new starters with a picture of what is most likely to enhance their performance at work and can help them consider how best to make the most of their talents.

Team development

The Saville Consulting Types Model helps members of teams see how they complement each other (e.g. Thinker Influencers are complemented by Adaptor Doers). Group profile reporting is available for team building to explore how two or more people are likely to interact. And the detail provided on the full SC Wave Expert reports can also provide deeper insight into how people interact with each other.

Organisational talent audit/benchmark

SC Wave provides a vehicle for benchmarking groups in terms of their perceived motives and talents. This information can be combined with our organisational surveys of preferred and actual culture to give a unique insight into how the motives and talents of employees are aligned to the organisational culture which can help inform future organisational development strategy.

Conclusion

Saville Consulting Wave Styles provide a different approach to assessment. The new methods of scaling have led to super short and valid scales. The presence of ipsative and normative responses helps to make faking more difficult and allows the user to focus in on potential areas of distortion (socially desirable over rating and self critical under rating).

The SC Wave matched model of performance aids better interpretation and decision making. It brings together motives, talents, competency and culture in one integrated model, providing the 'Expert User' with clearer links between an individual's motivation, work environment and performance at work.

The culture prediction component helps the individual being assessed (and their manager or colleagues) have a perspective on how the job could alter to enhance their performance and satisfaction at work.

References

- Baron, H. (1996). Strengths and limitations of ipsative measurement. *Journal of Occupational and Organisational Psychology*, 69, 49–56.
- Burisch, M. (1984). Approaches to Personality Inventory Construction: A comparison of merits. *American Psychologist*, 39(3), 214–227.
- Kurz, R. & Bartram, D. (2002). Competency and individual performance: Modelling the world of work. *Organisational Effectiveness: The Role of Psychology*, 10, 227–255.
- MacIver, R. (1997). A shorter OPQ? *Selection & Development Review*, 13(1), 13.
- Miles, A. (2006). Issues surrounding the engagement of new starters. *Proceedings of the British Psychological Society Division of Occupational Psychology Conference*, 11–13 January, Glasgow.
- Seligman, M. (2003). *Authentic happiness: Using the new positive psychology to realise your potential for lasting fulfilment*. London: Nicholas Brealey Publishing.
- Saville, P. & Willson, E. (1991). The reliability and validity of normative and ipsative approaches to the measurement of personality. *Journal of Occupational Psychology*, 64, 219–238.
- Smith, M. & Foley, M. (2006). The specific nature of error variance in selection. *Proceedings of the British Psychological Society Division of Occupational Psychology Conference*, 11–13 January, Glasgow.

Correspondence

Rab MacIver

Saville Consulting Group,
Harley House, 94 Hare Lane,
Claygate, Surrey KT10 0RB.
Tel: +44 (0) 1372 475700
Fax: +44 (0) 1372 475701
Mobile: +44 (0) 7747470864
E-mail: rab.maciver@savilleconsulting.com

Letter to the Editor

What high performance?

DOUGHTY, GERAGHTY AND SCULLARD produce some interesting findings on the relationship of motivation and personality to performance in British Gas (*SDR*, Vol. 21, No. 6). Sadly their evidence is flawed by their failure to define high performing Customer Service Advisers. Occupational Psychologists involved in OD work with companies, and most of us as consumers of call centres, know that clarity here is essential.

High performance for the company can be measured in terms of speed of response, brevity of time per customer and lack of repeat calls. High performance for the customer can be measured in terms of response to individual needs, ability to have a 'one-stop shop' and satisfaction

with the outcome. These are far more subjective measures – and arguably more important for the business in terms of customer satisfaction and retention. The 'are you satisfied with this call?' question often asked at the end of a less than satisfactory conversation produces false evidence. 'Yes' is the preferred answer to prevent wasting the customer's time with any follow-up questions!

Customers of British Gas would be interested to know how they perceive high performance. Readers of *SDR* would like to know just what it was that Doughty, Geraghty and Scullard were investigating.

Francis Cattermole

E-mail: francis.cattermole@btopenworld.com

Media Training Courses 2006

Working with the media? Want to gain some valuable tips and experience?

Whether you are a complete beginner or looking to update your skills, you will find our training sessions stimulating and enjoyable.

An Introduction to Working with the Media

A one-day immersion in the media – newspapers, magazines, radio and television – with lots of hands-on experience. This course is designed to give a general introduction to how the media operates, as well as introducing some of the skills necessary in media liaison. e.g. press release writing and interview techniques.

Members' Cost: £120 (inc. VAT) Dates available: 18 September 2006; 4 December 2006.

Broadcast Interview Skills

A one-day course that covers everything required for speakers to feel confident about taking on broadcast interviews. It will focus on radio interviews, but will also cover TV interview techniques.

Delegates will be provided with plenty of practical opportunity to get in front of the microphone and to gain experience of actually being interviewed.

Members' Cost: £170 (inc. VAT) Dates available: 21 August 2006;
11 December 2006.

All courses take place in London and include lunch and course materials.

Registration form and further details from:

Dawn Schubert

Administrator, Publications & Communications Directorate

Tel: 0116 252 9581; E-mail: mediatraining@bps.org.uk



The
British
Psychological
Society





The
British
Psychological
Society

PSYCHOLOGICAL TESTING CENTRE (PTC)

APPOINTMENT FOR SUB-EDITORS FOR PTC WEBSITE

The Society's Psychological Testing Centre is looking for Chartered Occupational or Chartered Clinical/Health Psychologists who enjoy writing to work as sub-editors for its website (www.psychtesting.org.uk).

Responsibilities will include:

- soliciting, editing and drafting content for the relevant parts of the site;
- ensuring appropriate guidance information and related documents are on the site;
- suggesting relevant links to the site;
- making a contribution to the strategy of the site;
- working with other sub-editors, Society staff and the editor; and
- answering enquires from the public and members of the Society as necessary.

Experience of editing websites would be an advantage.

The commitment required is approximately 12 days a year. This will mainly be taken up with the responsibilities of sub-editor although it may also include attendance at meetings. The appointment will be for one year renewable yearly up to a maximum of three years. The sub-editors will be financially compensated for their time (details of rates available on request from the PTC Office).

Informal enquiries can be made to **Jan Bogg**, Editor of the site, telephone number 0151 794 5526 or e-mail jbogg@liverpool.ac.uk. If you are interested in applying, please contact **Julie Mrowicki** (0116 252 9524 or e-mail, Julie.Mrowicki@bps.org.uk for a statement of interest form and a copy of the proposed contract.

Journals 2006



The
British
Psychological
Society

British Journal of Psychology

Volume 97 (2006). Quarterly. ISSN 0007-1269

www.bpsjournals.co.uk/bjp

British Journal of Clinical Psychology

Volume 45 (2006). Quarterly. ISSN 0144-6657

www.bpsjournals.co.uk/bjcp

Psychology and Psychotherapy: Theory, Research and Practice

Volume 79 (2006). Quarterly. ISSN 1476-0835

www.bpsjournals.co.uk/paptrap

British Journal of Health Psychology

Volume 11 (2006). Quarterly. ISSN 1359-107X

www.bpsjournals.co.uk/bjhp

British Journal of Educational Psychology

Volume 76 (2006). Quarterly. ISSN 0007-0998

www.bpsjournals.co.uk/bjep

British Journal of Developmental Psychology

Volume 24 (2006). Quarterly. ISSN 0261-510X

www.bpsjournals.co.uk/bjdp

British Journal of Mathematical and Statistical Psychology

Volume 59 (2006). Biannual. ISSN 0007-1102

www.bpsjournals.co.uk/bjmstp

British Journal of Social Psychology

Volume 45 (2006). Quarterly. ISSN 0144-6665

www.bpsjournals.co.uk/bjsp

Legal and Criminological Psychology

Volume 11 (2006). Biannual. ISSN 1355-3259

www.bpsjournals.co.uk/lcp

Journal of Occupational and Organizational Psychology

Volume 79 (2006). Quarterly. ISSN 0963-1798

www.bpsjournals.co.uk/joop

View online sample copy, content pages and abstracts free of charge at:
www.bpsjournals.co.uk

Discounted subscription rates for BPS members (includes online access)
2006 journal rates: £19 member / £14 student member (per journal per year)

For further information, please contact:

Subscriptions Department, The British Psychological Society, St Andrews House,
48 Princess Road East, Leicester LE1 7DR, UK.

Tel: +44 (0)116 252 9537; Fax: +44 (0)116 247 0787;

E-mail: subscriptions@bpsjournals.co.uk

www.bpsjournals.co.uk



The
British
Psychological
Society

The best psychology, to your inbox - free!



Sign up for The British Psychological Society's free, fortnightly e-mail packed with the latest research and links to the A-level syllabus.

www.researchdigest.org.uk

St Andrews House, 48 Princess Road East, Leicester LE1 7DR, UK
Tel 0116 254 9568 Fax 0116 247 0787 E-mail mail@bps.org.uk www.bps.org.uk

© The British Psychological Society 2006
Incorporated by Royal Charter Registered Charity No 229642



The
British
Psychological
Society