

**FINAL REPORT ON  
A RESEARCH STUDY RELATING TO  
EFFECTIVE SELECTION OF  
STAFF FOR SENIOR POSTS IN THE  
CIVIL SERVICE**

**This report was prepared by  
Professor Clive Fletcher (Personnel Assessment Ltd)  
on behalf of the  
Office of the Civil Service Commissioners**

### **PROFESSOR CLIVE FLETCHER BA PhD FBPsS**

Clive Fletcher is the Managing Director of Personnel Assessment Ltd. He was formerly Professor of Occupational Psychology at Goldsmiths' College, University of London, where he still holds the title of Emeritus Professor after leaving to work in private practice. He is also Visiting Professor at Henley Management College. Prior to going into academia, he spent nearly seven years in Whitehall. Clive is a Chartered Occupational Psychologist and one of the relatively few psychologists to be elected to Fellowship of the British Psychological Society (BPS).

Clive has published nearly 200 books, chapters, journal articles and conference papers on psychological assessment in work settings (see [www.personnel-assessment.com](http://www.personnel-assessment.com) for a list of some of the more recent of these). Clive has acted in an advisory and consultancy capacity to many private sector organisations and to Government departments - he is currently an advisor to the Cabinet Office and the National Clinical Assessment Authority, and an independent member of the Defence Scientific Advisory Council. His knowledge and experience in the assessment field is recognised internationally. He has given invited papers at conferences in the USA and throughout Europe and he is a member of the Scientific Advisory Boards of companies in the USA and elsewhere.

# Contents

<b>Executive Summary</b>	<b>Pages 5 - 9</b>
*****	
<b>Background &amp; Aims</b>	<b>Page 10</b>
<ul style="list-style-type: none"><li>• The Requirement</li><li>• Structure of this Report</li></ul>	
<b>What to Assess</b>	<b>Pages 11 -12</b>
<b>The Effectiveness of Different Assessment Methods</b>	<b>Pages 12 - 23</b>
<ul style="list-style-type: none"><li>• Introduction</li><li>• The Selection Interview</li></ul>	Pages 12 -13 Pages 13 -16
<ul style="list-style-type: none"><li>○ Types of Interview</li><li>○ Validity of Interviews<ul style="list-style-type: none"><li>▪ Panel Interviews</li><li>▪ What are interviews good at assessing?</li></ul></li><li>○ Fairness and the Interview</li><li>○ Impact on the Candidate</li><li>○ Cost</li><li>○ Meta-analysis references</li></ul>	
<ul style="list-style-type: none"><li>• Psychological Tests and Questionnaires</li></ul>	Pages 16 -19
<ul style="list-style-type: none"><li>○ Cognitive Ability Tests</li></ul>	Pages 16 -18
<ul style="list-style-type: none"><li>○ Validity</li><li>○ Fairness</li><li>○ Impact on the Candidate</li><li>○ Cost</li><li>○ Meta-analysis references</li></ul>	
<ul style="list-style-type: none"><li>○ Personality Questionnaires and Related Measures</li></ul>	Pages 18 -19
<ul style="list-style-type: none"><li>○ Validity</li><li>○ Fairness</li><li>○ Impact on the Candidate</li><li>○ Cost</li><li>○ Meta-analysis references</li></ul>	
<ul style="list-style-type: none"><li>• Application Forms, References and Biodata</li></ul>	Pages 20 - 21
<ul style="list-style-type: none"><li>○ Validity<ul style="list-style-type: none"><li>▪ Application Forms</li><li>▪ References</li><li>▪ Biodata</li></ul></li><li>○ Fairness</li><li>○ Impact on the Candidate</li><li>○ Cost</li><li>○ Meta-analysis and review references</li></ul>	

• Assessment Centres and Work Sample Tests	Pages 21 - 23
○ Validity	
○ Fairness	
○ Impact on the Candidate	
○ Cost	
○ Meta-analysis references	
<b>Some overall conclusions on the literature review of different selection and assessment methods</b>	<b>Pages 23 - 26</b>
<b>The literature on top management selection</b>	<b>Pages 26 - 27</b>
<b>A brief survey of top management selection practices outside the Civil Service</b>	<b>Pages 27 - 30</b>
• Findings from overview bodies	Page 28
• Findings from organisations contacted	Pages 29 - 30
<b>Conclusions and Recommendations</b>	<b>Pages 30 - 33</b>
Appendix A: What is a competency-based interview?	Page 34
References	Pages 35 - 38

# EXECUTIVE SUMMARY

## Background and Purpose of Study

1. The requirement issued by the Office of the Civil Service Commissioners that gave rise to this project was specified as follows –

“To review up to date UK research findings on the most reliable, validated techniques for recruiting and selecting senior staff in the private, public and voluntary sectors, with particular reference to effective selection interviewing, the use of assessment centres and psychological and psychometric testing”

In addition to this literature review, it was also judged desirable to know what other organisations in the public and (particularly) the private sector do in relation to top level appointments. The second element of the work was therefore to contact a number of such organisations to conduct a small survey of the kind of selection methods they are using in making such appointments.

## Findings (a): What to Assess?

2. Attention is drawn to the work on ‘meta competencies’, or learning competencies and to other work on competencies for top level management positions. Although a competency framework already exists for the Senior Civil Service (SCS), it may be worth reviewing this as a basis for more clearly specifying the requirement when selecting to top SCS positions. It is noted that some organisations take a broader perspective on what to assess, and look beyond an individual candidate’s competencies and take into account other factors, such as the competency profiles and characteristics of the people they will be working closely with.

## Findings (b): The Effectiveness of Different Selection Methods

3. The literature on the effectiveness of the various assessment methods was reviewed against four criteria –

- *Validity* – the extent to which they can be shown to measure what they claim to measure, and specifically how well scores or judgements made on a particular measure correlate with performance measures. A correlation of 1.0 would represent perfect prediction (never obtained!); correlations around .20, .40 and .60 would reflect low (but possibly still useful), good and high validity respectively.
- *Fairness* – the extent to which the data they produce are free from unfair discrimination and bias
- *Impact on the Candidate* - the reactions and attitudes candidates tend to have when faced with a particular assessment method in a selection situation
- *Costs* – a discussion of the likely costs involved in applying different methods

4. The average validity achieved by the different methods reviewed is summarised in the Table below; each of the methods and the broader findings on them will be discussed in the paragraphs that follow:

Structured Interviews	.44-.56
Tests of Cognitive Ability	.40 -.50
Assessment Centres	.37-.43
Personality Questionnaires	.40
Work Sample Tests	.32
Biodata	.30-.35
References	.25
Unstructured Interviews	.14-.33

Up to a point, validity can be increased by combining assessment methods. For example, use of structured interviews and psychometric measures together has typically been found to achieve validity of around .63-.65.

**5. The Interview.** The main conclusions relating to the interview are -

- The more highly structured the interview type and (up to a point) the longer the interview, the higher the validity tends to be
- Prediction of performance ratings is better than prediction of length of job tenure
- Training of interviewers makes a difference irrespective of the amount of interview structure
- Panel interviews are no more or less effective than comparable one-to-one interviews done sequentially
- There is little or no evidence of unfair bias found in structured interviews, but unstructured interviews are more likely than any other assessment method to face legal challenges
- Candidate reactions to interviews are largely positive, though less so in the case of highly structured interviews

**6. Psychological Tests and Questionnaires: Cognitive Ability Tests.**

- Almost all the evidence collected, from many hundreds of studies done across hundreds of thousands of candidates and covering a huge range of jobs, consistently shows that cognitive ability is one of the best predictors of future job performance.
- The question of whether these findings hold for top level management is less clear cut. In general, of course, senior executives need a high degree of general mental ability – but how high? Beyond a certain ‘threshold’ measure, general mental ability may not add much to our prediction of managerial performance. Where that threshold lies will probably vary from job to job.
- Cognitive tests have a potential for producing adverse impact. Again, though, at top management level it is doubtful that group differences (based on ethnicity or gender) would have any impact on performance – because top management candidates are already a highly able, highly-selected group.

- Cognitive tests are not especially ‘popular’ with candidates, but much depends on their perceived relevance to the job being applied for and how the tests are presented.

## 7. Psychological Tests and Questionnaires: Personality Questionnaires.

- Personality questionnaire use is now widespread in the UK. The CIPD Recruitment, Retention and Turnover Survey for 2004 shows that 46% of the organisations covered used them – which compares with 43% using assessment centres and 53% using cognitive ability tests. However, the figures for public sector organisations only were 36%, 48% and 52% respectively. For the UK as a whole, test and questionnaire use has increased since 2002 while assessment centre use has decreased.
- The evidence shows that personality questionnaires do predict behaviour and performance, producing correlations that are useful, but less than those for cognitive ability tests. However, precisely which personality factors are the most important to assess is less clear cut and will vary across job roles.
- Whilst there is little indication of ethnic group differences on personality questionnaires, gender differences are found – these seem to reflect genuine differences in behaviour and self-perceptions.
- As with cognitive tests, candidate reactions depend largely on how they are presented. Candidate “faking” on personality measures does not seem to be a major problem in terms of their validity.

## 8. Application Forms, References and Biodata

- The value of Application Forms can be improved by adding some structure to them, principally in terms of including self assessment against competencies. Similarly, References structured around competency ratings provide greater validity. However, changes in the law are making references increasingly difficult to get.
- Biodata questionnaires can achieve good validity but are subject to concerns about fairness, and have been found to be unpopular with candidates.

## 9. Assessment Centres and Work Sample Tests

- Assessment Centres (ACs) show higher validity (around .60-.65) when using career advancement criteria, eg promotion, than when using job performance criteria. Work Sample tests and simulations are usually a large part of an AC, but they may be used as stand-alone assessment tools also. When looked at against job performance criteria, such tests obtain correlations in the order of .32
- ACs have a good record on fairness, and are also usually perceived positively by candidates if they are run well. However, for top level posts, issues of confidentiality and political aspects of the selection situation make them more difficult to use.
- ACs are expensive to set up and apply, though individual Work Sample tests are much less expensive and offer a way of assessing cognitive ability and decision making.

### **Findings (c): Survey of External Practices in Top Management Selection**

**10.** A total of 15 organisations were approached, including the Institute of Directors and the Chartered Institute for Personnel and Development (CIPD). The observations from this part of the study can be summarised as follows –

- It was estimated that 85% of external organisations still use just interviews and references for top posts. Those with HQs in the USA or other close American links, organisations in the retail sector and Local Authorities were more likely to make use of psychometrics, work sample tests or assessment centres.
- In the private sector, the interviews are nearly all done on a sequential, one-to-one basis
- The amount of time spent on interviewing is typically considerably more than is usually the case in senior civil service appointments
- The interviewing is often done, in part at least, around the organisation's competencies

### **Conclusions and Recommendations**

**11.** How might the present approach to assessing and selecting candidates for top civil service posts be improved? Taking each stage of the process in turn-

- A more systematic analysis of the role/candidate requirements in terms of the SCS competency framework might be helpful, and indeed essential if other recommendations that follow are to be implemented. In addition, though, the SCS competency framework itself has been in existence for some while and might benefit from review. To some extent, this has already happened in the process of establishing the High Potential Development Scheme and its enhancement of the SCS competency set, but this review should be put on a more formal and systematic basis. More attention might also be paid to looking at each appointment in terms of the profile of the peers and direct reports the individual would be working with.
- Consideration might be given to including structured self assessment on the application forms, and the validity of References might also be improved by including an element of structure around competencies.
- Where Individual Psychological Assessments are used, more time and attention needs to be given to the briefing of the psychologist concerned, and to the use of the output (assessment reports) from this element of the process
- To increase the validity of the interview, it would be helpful to both include an element of structured, behaviourally focused interviewing, and to expose the candidates to longer interviewing time that is presently the case.

**12.** To deal with some specific issues -

- Given that the validity of cognitive ability and personality measures is good, the case for including them more routinely is strong. However, their use needs to be guided and targeted by careful briefing about the role being selected for and its context.
- At the present time, the psychometric measures that are most likely to be useful for top posts are personality and related (eg Leadership, Emotional Intelligence, etc) questionnaires.
- Cognitive ability tests might be valuable for selection to some top level posts, but careful consideration needs to be given as to whether there is any real danger of indirect discrimination arising. The situations where they might make the most valuable contribution are where the demands of the job place a special emphasis on intellectual capacity, and where the candidate field is varied in terms of its background and academic attainment.
- For the most part, it seems unlikely that Assessment Centres will be feasible alternatives for top management selection, on grounds of cost and candidate concerns. But Work Sample tests could be included in the assessment for top CS posts. It might be worth considering commissioning the development of one or more Individual Decision Making exercises that have generic relevance to SCS work.
- Although very highly structured interviews yield the best validity, they are less likely to be appropriate for the complexity of top level jobs. A competency-based interview offers a more suitable approach. It would be desirable to offer – if not insist on – training in competency based interviewing for all those likely to be involved in it.

**13.** To summarise; it is **recommended** that –

**The current approach to making top management appointments in the Civil Service be reviewed and revised to increase its potential validity, and that this should principally involve -**

- A. A more refined and systematic way of describing the job and the person specification, at least partly in terms of the relevant competencies**
- B. The use of assessment devices other than just the interview - specifically psychometric measures and, possibly, Work Sample tests; commissioning of the development of the latter for exclusive CS use might be considered**
- C. Including a more structured, behaviourally focused approach to at least part of the interview, and greater time being devoted to the interviewing**
- D. Training for interviewers in any new approach adopted**

\*\*\*\*\*

# MAIN REPORT: BACKGROUND AND AIMS

## **The Requirement**

The Civil Service Commissioners have recently initiated a project aimed at helping Civil Service departments improve the way they select and recruit staff at the most senior level. The Commissioners wish to ensure that the processes they regulate, and which are used by departments and agencies to recruit senior staff, are the most effective available in identifying the most meritorious candidate. They want to consider whether or not the standard approach to recruitment used by departments comprising long listing, short listing and a single board interview can be improved.

As part of that project they commissioned this piece of work, the first element of which consisted of a literature review of recent research findings on the most reliable, validated techniques for recruiting and selecting senior staff in the private, public and voluntary sectors. The requirement is specified as follows –

“To review up to date UK research findings on the most reliable, validated techniques for recruiting and selecting senior staff in the private, public and voluntary sectors, with particular reference to effective selection interviewing, the use of assessment centres and psychological and psychometric testing”

However, as much of what is known about selection at high levels is anecdotal or is reported in case-study style, it was also judged desirable to capture this and to know what other organisations in the public and (particularly) the private sector do in relation to top level appointments. The second element of the work was therefore to contact a number of such organisations to determine the kind of selection methods they are using in making such appointments.

## **Structure of this Report**

The first part of the report described the findings of the literature review, dealing first with some brief observations on what is assessed, then going on to look in detail at the effectiveness of the various assessment methods themselves. The latter is broken down into an outline of some general issues relating to the research in this field, followed by a consideration of the findings on (1) Selection Interviewing (2) Tests of Cognitive Ability (3) Personality Questionnaires and related measures (4) Application forms, References and Biodata (5) Assessment Centres and Work Sample Tests. An overall conclusion on this literature is presented before moving on the next section, which focuses on Top Management in particular.

The second part of the report covers the contacts with other organisations and the survey of top management assessment practices followed in them. The final section offers, as final sections do, a summary and list of conclusions, along with some recommendations.

## WHAT TO ASSESS?

There is a vast literature on competencies, and it is beyond the remit of this project to discuss this. In addition, a competency framework for the Senior Civil Service, along with a complementary one for assessing candidates on the High Potential Development Scheme, already exists. However, some of the wider literature on competencies has been devoted to identifying those competencies that are believed to be important in, or indicative of potential for, top management. Hall (Hall & Moss, 1998; Briscoe & Hall, 1999) puts forward the notion of “metacompetencies”, which are competencies so powerful that they affect an individual’s ability to acquire other competencies. Hall’s work, in turn, is influenced by McCall (Spreitzer, McCall & Mahoney, 1997), who pointed out that the skills needed tomorrow may not exactly the same as those needed today, and that the ability to learn from experience may be more important than existing competency levels. Thus, the basic idea behind Hall’s metacomptencies is that they describe the characteristics underpinning the ability to both acquire experience and to learn from it, which is the key to development generally. He proposes two of them, Identity and Adaptability -

The learning behaviours associated with Identity include -

- Accurate self assessment
- Seeking, hearing and acting on feedback
- Exploring, communicating and acting on personal values
- Being open to diverse people and ideas
- Engaging in personal development activity
- Being able to modify one’s self perception as one’s attributes change

The learning behaviours associated with Adaptability include –

- Being able to identify for oneself the qualities that are critical for future performance and being able to make the changes needed to develop them
- Eagerness to accept new challenges
- Exploration of new territory
- Comfort with turbulent change

These concepts have echoes in some of the work on self awareness (Fletcher, 1997a) and on emotional intelligence (Higgs & Dulewicz, 1999). The SCS competency framework includes “Learning and Improving” as one of six key competencies, and this contains some elements of the ideas behind learning competencies. For example, some of the associated behavioural indicators for this competency are -

- Aware of own strengths, weaknesses and motivations
- Applies learning from own and others’ experience
- Understands, values and incorporates different perspectives
- Seeks new or different opportunities to learn
- Readily shares idea or information with others
- Encourages experimentation and tries innovative ways of working.

It may be that this element of the SCS framework should be further strengthened.

Other literature that is relevant to any review of the SCS competencies is the competency framework developed for Directors on behalf of the Institute of Directors (Dulewicz, 1995; Dulewicz & Gay, 1997; Gay & Dulewicz, 1997).

A broader perspective on what to assess is the extent to which organisations look beyond an individual candidate's competencies and take into account other factors, such as the competency profiles and characteristics of the people they will be working closely with. Although there is little literature on this (other than psychometric measures of team make-up and functioning), there are examples of this kind of consideration entering into the selection process for top management in the private sector, as will be seen below (page 29) when the findings from the survey of outside practice are discussed.

## **THE EFFECTIVENESS OF DIFFERENT ASSESSMENT METHODS**

### **Introduction**

The evaluation criteria considered in this review are as follows –

- *Validity* – the extent to which a selection method can be shown to measure what it claims to measure. Traditionally, this can be established by either discovering the extent to which candidates' scores on a selection method correlate with either existing (known as concurrent validity) or future (known as predictive validity) measures of their performance. This is the most fundamental criterion, because a procedure that lacks validity is clearly unacceptable. A correlation of 1.0 would indicate perfect validity, an exact correspondence between the scores on some measure (eg a test) and the criterion measure of performance – a level of perfection that is never achieved. A correlation of 0 would indicate no relationship at all. Most correlations in the selection field range from around .10 to .60, the latter being considered high.
- Sometimes Reliability is used as a measure of how sound a method is – it refers to the extent to which it will produce the same or very similar results with a group of individuals when applied to them on two or more successive occasions with varying lengths of time in between. A reliable measure is one that produced consistent results (assuming it is seeking to measure a consistent quality of the individual). However, reliability will not be referred to separately in this report, and will be subsumed under validity, because it is impossible for a measure to be valid without also being reliable (though the reverse is not true – a measure can indeed be reliable without being valid).
- *Fairness* - the extent to which a measure is free from direct or indirect discrimination effects, and the extent to which it is perceived to be fair.
- *Candidate Impact* – the reaction candidates have to different selection methods and processes. This is an important criterion, as it can impact on the candidate's perceptions of the organisation and their willingness to accept a job if one is offered (Hausknecht, Day & Thomas, 2004).
- *Cost* – the financial costs of implementing various methods, and the way these do or do not relate to their effectiveness on other criteria.

Before plunging in to the findings on effectiveness, however, there is an important methodological point that has to be outlined. This is the way in which findings in the selection and assessment field are brought together and summarised. Because individual research studies often use different samples (in size and nature), varying procedures for collecting the data and alternative forms of statistical analysis technique, it used to be impossible to reflect a synthesis of the results from all the studies in anything but the most crude manner. Around 25 years ago, however, a technique called **meta-analysis** was devised that facilitated the statistical aggregation of research findings from many different studies and the assessment and expression of their findings in a common way. This technique has become the standard method of summarising such research. Accordingly, what follows in this report is based heavily on the relevant meta-analyses that have been carried out – some of which focus on specific methods and some of which cover a range of methods and make comparisons between them. A word of caution, though – not all meta-analyses produce the same results, because they too may vary slightly in methods used but also, more importantly, in the studies they have sought to bring together and summarise. Nonetheless, there is a very good level of convergence across meta-analytic studies.

To enhance ease of reading, the review of each assessment method presented below lists the relevant meta-analyses it rests on at the end of that section, rather than repeating the reference to them at every conclusion they might support. However, where an individual research study supports a specific issue being discussed, it will be referenced at that point in the text.

## **The Selection Interview**

### Types of Interview

Partly because of the inadequacies of early research methods and the way their findings were presented, in the past the Interview received a worse ‘press’ than it deserved. Whilst there is no room for complacency, the interview has been rehabilitated to some extent, not least because of the developments in the types of interview available. The descriptions used here are –

- *Unstructured:* Interviews that follow no set and predetermined pattern, typically varying in content from interviewer to interviewer and from candidate to candidate.
- *Structured:* Interviews that follow a pre-determined format, which may vary in terms of rigidity of structure and in the content or form of interviewing. The main types are –
  - Loosely structured, where there is a list of topic areas to be probed in the interview, though not very much structure beyond this, based on an analysis of the job and a resulting person specification.
  - Competency based interviews (also known as behaviourally focused interviews), which are structured around a set of competencies (typically 4-6 in a one hour interview), with the behavioural descriptors associated with each competency forming the basis of questions that ask the candidate to describe in some detail examples where they have demonstrated such behaviour in the past.

- Situational Interviews are developed on the basis of a careful analysis of a particular role and problem situations that typically have been found to arise in it. Approaches to dealing with these situations are identified and evaluated by people who have experience of managing them, and 'scored' according to the degree of effectiveness the varying approaches have shown in terms of solving the problem. The situations are then presented as interview questions and the candidate asked how they would deal with them, their answers being scored according to which of the response alternatives they most closely correspond. These are thus highly-structured interviews, almost akin to an oral questionnaire. A variant on it is known as the Behaviour Description Interview (BDI), which is similarly based on an initial critical incidents analysis, but is more focused on using this information to ask about how the individual has dealt with similar situations in the past (akin to a competency based interview) rather than how they would in the future (as in the situational interview).

The first two structured interview types mentioned above are also sometimes referred to as Conventionally Structured Interviews (CSI). There are numerous other variations on these themes.

### Validity of Interviews

There have been numerous meta-analyses of the interview, and the conclusions are consistent, namely that structured interviews achieve good levels of validity (though the findings are perhaps a little stronger on concurrent than on predictive validity). To summarise them -

- Unstructured interviews achieve validity correlations in the range of .14 to .33
- Structured interviews achieve validity correlations in the range of .44 to .56
- The more highly structured the interview type and (up to a point) the longer the interview, the higher the validity tends to be
- Prediction of performance ratings is better than prediction of length of job tenure
- Training of interviewers makes a difference irrespective of the amount of structure the interview has (Hufcutt & Woehr, 1999)
- Interview questions focusing on past behaviour have better validity than questions that focus on (anticipated) future behaviour (Salgado, 1999)

### *Panel Interviews*

The majority of interviews that the research literature is based on are done on a one-to-one basis. However, two meta-analyses have specifically looked at studies of panel interviews and compared them with studies of one-to-one interviews. Wiesner & Cronshaw (1988), whose review covers studies that had a combined candidate total in excess of 51,000, found no significant difference between the two approaches, though such trends as there were favoured the panel. Hufcutt and Woehr (1999) similarly found no improvement in validity was associated with using panel interviews. The basis of such comparisons should be treated with caution, though – what is the most appropriate comparison to make, a panel and sequential one-to-one interviews with an equal number of interviewers involved in each approach? Or should it be done on the basis of time with the candidate, so

that the panel time should equate with the total time taken by all the individual interviews? Or both? It should be noted that the studies covered by these analyses did not adequately control for such factors.

A rather different approach to the same question was taken by Bayne, Fletcher and Colwell (1983). They compared both types in a carefully controlled study using Civil Service Commission interviewers, and again found no substantial advantage of three individual, sequential interviews over panels (of three members) in terms of accurate prediction of a personality attribute or the amount and quality of information obtained.

The CIPD Recruitment Retention and Turnover Survey for 2004 (CIPD, 2004) showed that 85% of the public sector organisations covered used structured panel interviews compared to 45-46% in the private sector.

#### *What are Interviews Good at Assessing?*

Not all interviews are carried out to assess the same attributes – indeed, they are used to assess against a wide range of qualities that differ across jobs, levels and organisations. Robertson and Smith (2001) tentatively conclude that interviews measure social skills, experience and job knowledge and that general mental ability has only a moderate correlation with interview performance. Not surprisingly, the interview shows a reasonable ability to assess extraversion and emotional stability.

#### Fairness and the Interview

Empirical studies on adverse impact and the interview are not extensive (Moscoso, 2000). However, those that have been done are largely reassuring. For example, the meta-analysis by Huffcutt and Roth (1998) showed that interviews had little negative impact on ethnic minorities and that structured interviews were better in this respect than unstructured. Harris (1989) looking at age, gender and ethnicity as factors concluded that there was little evidence of discrimination in the interview. It should be noted that most of these studies were done in the US, where it has also been shown that not only are unstructured interviews much more likely to be legally challenged than structured interviews (Terpstra, Mohammed & Kethley, 1999), but they are also the most likely to be challenged than all the other main selection methods. In terms of perceived fairness on the part of the candidate, the interview is generally seen as fair, but like all other selection methods this is dependent on outcome – people perceive procedures to be less fair when they are rejected!

#### Impact on the Candidate

The interview is generally perceived to be acceptable and appropriate as a selection methods, and this seems to hold true across difference countries (Steiner & Gilliland, 1996; Salgado & Moscoso, 2000). But the evidence on the effects of different types of interview is much less clear-cut. Overall, it suggests that structured, behaviourally focused interviews receive a slightly less favourable candidate reaction, though the quality of studies and the methodology followed are not strong in many of the investigations done in this area.

On a slightly different tack, candidates prefer face to face interviews compared to video-conference interviews (Kroeck & Magnusen, 1997).

Finally, an older study (Keenan, 1978) showed that job applicants were more likely to accept an offer from a trained interviewer than one who was not. This does not mean that the applicants knew who was trained and who was not, but probably results from the trained interviewers adopting a more competent approach, which influenced the candidates' decisions.

### Cost

Interviews – especially if they are unstructured - are potentially one of the lowest cost selection tools, and this part of their attraction. However, developing a structured interview approach and training interviewers to use it entails quite considerable initial setting-up costs. Just how much will vary according to the nature of the interview process adopted. In addition, when selecting for top management levels, the cost of the interviewers' time has to be added into the equation, and at such levels the interviewers themselves will be very senior and hence more expensive. The panel approach may be more cost effective in one sense – less time may be needed for more people to have a view of the candidate. But this does not answer the question as to whether more time spent on sequential one to one interviews might not produce better validity – given that more interview time is associated with better validity.

### Meta Analysis References:

Weisner & Cronshaw (1988); McDaniel, Whetzel, Schmidt & Maurer, (1994); Moscoso (2000); Salgado & Moscoso (2001).

## **Psychological Tests and Questionnaires**

The word Test implies a right or wrong answer, so the term Psychological Tests is usually applied to cognitive ability measures (such as tests of verbal or numerical reasoning ability, and so on). With personality questionnaires and similar devices, however, there are no right or wrong answers and they should not be referred to as tests – though they often are. Psychometrics is a blanket term often used to cover both types. In this section, Cognitive Ability Tests will be dealt with first, followed by Personality Questionnaires and Related Measures.

### **Cognitive Ability Tests**

#### Validity

Almost all the evidence collected, from many hundreds of studies done across hundreds of thousands of candidates and covering a huge range of jobs, consistently shows that cognitive ability is one of (if not the) best predictors of future job performance. Whilst specific cognitive abilities – like verbal ability or spatial ability – may be important in specific roles, the overall conclusion is that it is an underlying general mental ability factor, usually referred to as 'g', that is the key predictor and which underlies performance on all such tests. Measures of specific ability seem to add little or no incremental predictive power.

Typically, meta-analyses have shown that where job performance criteria are used, tests of general mental ability correlate around .40 to .50, and where the criterion is performance in training this rises to .50 to .60; while most of these analyses have been done in the US, the findings from Europe are much the same (Salgado, Anderson, Moscoso, Bertua & de Fruyt, 2003).

There is little disputing the findings on cognitive ability and work performance, but it has to be noted that they do, as indicated, reflect studies across a very wide range of jobs, some of which are at graduate and managerial levels but also many of which are at unskilled or semi skilled in nature. The question of whether the findings hold for top level management is less clear cut. In general, senior executives do of course need a high degree of general mental ability – but how high? When Goleman (1996) presented his case for the importance of ‘Emotional Intelligence’ (EI), he claimed that IQ only accounted for about 20% in predicting managerial performance, with the other 80% being down to personal factors – of which he proposed EI was the major component. Leaving aside the basis for Goleman’s specific claims – some of which are rather thin – he was in fact echoing observations made long before, namely that one can have too much of this particular ‘good thing’. At very high levels mental ability the correlation with managerial performance measures may fall and even become negative (ie, the more intelligent you are beyond a certain point, the less likely you are to perform well on performance measures). Thus, beyond a certain “threshold” measure, general mental ability may not add much to our prediction of managerial performance. Where that threshold lies will probably vary from job to job and role to role.

### Fairness

Tests of general mental ability consistently show subgroup differences, principally between Afro-Caribbean and white samples, with the former performing up to one standard deviation lower than the latter. But beyond this, it is difficult and dangerous to generalise. Women are also sometimes found to perform a little less well than men – and sometimes better than them. Other ethnic groups (eg Asian or Chinese) can outperform whites, and educational and socio-economic background levels have an impact on the findings, as does the amount of test-taking experience. At top management level, it is doubtful that group differences would have any impact on performance – because top management candidates are already a highly-able, highly-selected group.

### Impact on the Candidate

Cognitive tests are not especially ‘popular’ with candidates – apart from anything else, they are of course rather demanding and difficult, as they are intended to be – but much depends on their perceived relevance to the job being applied for and how the tests are presented. If the purpose of the tests and how they are being used is explained, and feedback on performance is offered, then attitudes towards them become more positive (Fletcher, 1997b; Anderson, Born & Cunningham-Snell, 2001). For ethnic minority candidates, giving practice test items is also important (though probably less so at more senior levels).

There is a separate issue of Test Anxiety – some individuals habitually under-perform on cognitive ability tests because instead of focusing on performing the task, they dwell on non-task relevant thoughts, such as the consequences of failure, how other people in the room are doing, and so on. If given the same tests in a non-evaluative setting, they perform at a significantly higher level.

### Cost

Costs of testing vary according to how the tests are given, by whom and to how many candidates. In themselves, cognitive tests cost only a few pounds per candidate, and – when taking account of their predictive power - represent the most cost-effective assessment method available. But if the tests are administered in-house, then the costs of training people to obtain the British Psychological Society Level A Certificate of competence in occupational testing – a pre-requisite for ordering tests from reputable test publishers – is likely to run to around £1,000 per head. If the tests are given by an external consultant, then they will be an integral element of the fees charged.

### Meta Analysis References

Hunter & Hunter, (1984); Schmidt & Hunter, (1998); Salgado & Anderson (2001); Salgado, Anderson, Moscoso, Bertua & de Fruyt, (2003)

## **Personality Questionnaires and Related Measures**

Personality questionnaire use is now widespread in the UK. The CIPD Recruitment, Retention and Turnover Survey for 2004 (CIPD, 2004) shows that 46% of the organisations covered used them – which compares with 43% using assessment centres and 53% using cognitive ability tests. However, the figures for public sector organisations only were 36%, 48% and 52% respectively. For the UK as a whole, test and questionnaire use has increased since 2002 while assessment centre use has decreased.

### Validity

There is now little serious contention of the evidence showing personality questionnaires do predict performance criteria, correlations falling around .40 on average – certainly useful, but less than those for cognitive ability tests. However, precisely which personality factors are the most important to assess is less clear, and in particular whether they should be broad personality traits such as the ‘big 5’ (Extraversion, Neuroticism, Agreeableness, Openness to Experience and Conscientiousness) or more specific constructs, such as Locus of Control. Measures of personality attributes such as Integrity or Need for Achievement can and do show important correlations with work behaviour, but the choice of what measures are used is probably best focused on the nature of the role being selected for.

Under this broad heading, there are many other questionnaire measures that are *personality related* – such as measures of motivation, leadership, emotional intelligence, team contribution and so on. The findings here are more variable and

dependent on the specific measure and the criterion used to judge it. Hence, some considerable caution is needed in employing such questionnaires.

### Fairness

Most of the research in this respect has focused on gender differences, and they have often been found – for example, women scoring higher on anxiety, conscientiousness and gregariousness than men, and men scoring higher on assertiveness. Whilst these may well represent ‘real’ differences rather than measurement artefacts, they do contain the potential for producing indirect discrimination unless used carefully. Again, much of the research has been done in the US, but Ones and Anderson (2002) looked at three widely-used personality questionnaires with a large UK sample and failed to find any large gender differences. Nor did such ethnic group differences as they found achieve a magnitude that would make adverse impact a concern.

### Impact on the Candidate

Much the same can be said of personality measures as was said earlier about the impact of cognitive tests on candidates – they may not elicit a very positive reaction if applied insensitively, and much depends on their perceived relevance and how they are presented/explained to the candidate. However, Test Anxiety as such is not a problem in relation to personality measures.

Another issue that does differ from cognitive testing, and which is often raised in relation to the use of personality questionnaires, is faking or ‘social desirability’ effects – candidates seeking to indulge in impression management. Actually, the evidence is largely reassuring on this, and it does not seem to be a major problem for validity (Robertson & Smith, 2001). In other words, some candidates may do it, but overall it does not detract significantly from the effectiveness of the measures. On an anecdotal level, it might be observed that top management candidates are often sufficiently self confident to project themselves honestly and directly.

### Cost

Again, most of what was said about cognitive tests applies here, except that the relevant qualification for using and ordering reputable personality measures is the BPS Level B certificate in competence. Thus, to use both cognitive and personality measures, an individual needs to acquire both certificates of competence. The cost effectiveness of personality questionnaires is not quite as good overall as that for cognitive tests, because they do not have equivalent predictive power and can cost slightly more per head.

### Meta Analysis and Survey References

Barrick & Mount, (1991); Salgado, (1998); Ones & Viswesvaran, (2001)

## **Application Forms, References and Biodata**

### **Validity**

*Application forms* are often the first sifting mechanism. They are usually judged necessary to collect basic information about those applying, but they are often extended beyond this to collect self-assessment data on competencies. Evidence suggests that certain forms of structured self assessment have been shown to correlate with assessments made by referees and may have some value in selection (Jones & Fletcher, 2002, 2004). However, it is difficult to arrive at a validity figure for application forms alone.

*References*, as predictors of subsequent performance, have a poor record when subject to research. Only the use of structured references, as opposed to the typical reference letter, seem to have a modest amount of validity, at best reaching correlations averaging around .25. But, paradoxically, references are perhaps increasingly important as independent verification that the candidate is who they say they are and possesses the qualifications they claim (failure to do so might leave the employer open to a charge of negligent hiring in the event of some subsequent problem), and at the same time increasingly difficult to obtain, a fact that is perhaps reflected in the declining use of employment references in the UK (CIPD, 2004). The latter almost certainly reflects changes in the law which allow candidates to see what is written about them and to challenge it if they wish – if a reference contains false or unsubstantiated statements that damage an individual's reputation, they may sue. Given this, and the fact that employers are not legally obliged to provide references, many are now less willing to do anything other than to confirm that the individual worked for them during a specified period.

*Biodata* is a term given to the statistical prediction of future performance on the basis of basic biographical information about the candidate. This breaks down into 'hard' biodata, which is factual in nature (eg type of school attended, their academic attainments) and 'soft' biodata, which relates to statements candidates make about themselves on a structured questionnaire (eg about their likes, interests etc). The latter may become almost indistinguishable from a personality questionnaire. To create an effective Biodata measure, the information has first to be collected and then statistically analysed to show which elements of it are related to subsequent performance; this is obviously a fairly sophisticated exercise, and there is no guarantee of success. However, Biodata has a reasonable record in terms of validity, around .30-.35 typically. There is some uncertainty, though, about how well Biodata measures generalise – in other words, does one have to create a new Biodata measure for each job?

### **Fairness**

There is little of substance specifically relating to fairness and application forms or references for top management positions (other than to note the point made above about the legal implications of unsubstantiated references), though gender and ethnic biases can enter into the content and evaluation of either. Biodata can become a vehicle for bias and for perpetuating an existing culture if care is not taken. Thus, if the current situation in a particular organisation is that going to a public school and

Oxbridge is associated with making rapid progress in that organisation, then building these ‘predictors’ into a Biodata instrument to be used in selection will simply produce more of the same. Interestingly, though, biodata has been shown to be more valid for women than for men (Bliesener, 1996).

### Impact on the Candidate

Managerial candidates prefer to apply for jobs by completing application forms or sending CVs (Mathews & Redman, 1998), and probably expect to give references. However, they react negatively to biodata use, as has been indicated by a number of studies (Anderson, Born & Cunningham-Snell, 2001).

### Cost

Completed application forms and references are relatively inexpensive to obtain (especially with small numbers of candidates) and the former can yield a great deal of useful information if designed correctly. For this to be the case, though, means allowing for a degree of tailoring for the specific post. Designing and validating a Biodata questionnaire is a potentially high-cost option and usually takes some while to achieve – it is only likely to be worthwhile where large candidate numbers are involved.

### Meta Analyses and Review References

Bliesener, (1996); Bright & Hutton, (2000); Robertson & Smith, (2001).

### **Assessment Centres and Work Sample Tests**

The term assessment centre (AC) should be reserved for a process which involves the assessment of a group of candidates by a team of trained assessors using a variety of integrated assessment techniques chosen or devised for their relevance to the attributes being measured, which themselves should have been arrived at by a careful and systematic analysis of the requirements of the role being selected for. They are likely to take not less than a day for each candidate group, and sometimes two or even three days. Unfortunately, the term AC is often applied to procedures that are not worthy of that name, because they are haphazardly-constructed, use untrained assessors and rest on no clear person specification. What follows is, as far as possible, related to ‘true’ ACs.

One of the key elements of an AC is simulation exercises (ACs may or may not also contain psychological tests and questionnaires, and interviews). Typically, these may be individual decision making exercises (like an in-tray exercise) or group decision making exercises (such as a business game between teams of candidates, or a task requiring the candidate group to analyse information and come to a decision). Depending on how they are constructed, these simulation exercises may be called Work Sample tests – because they may actually represent a sample of the work that would be involved if the person is appointed.

## Validity

There have been many studies over a long period of time that have demonstrated the validity of assessment centres, and in this context perhaps most notably the 30-year follow-up of CSSB entrants by Anstey (1977), which showed a correlation of .66 between the assessments made at CSSB and the grade within the CS subsequently attained. Most studies of ACs using job performance criteria rather than career advancement criteria, however, produce rather more modest correlations in the range of .37-.43. It should be noted that career advancement criteria are more objective (people either make a grade or they don't), though perhaps less relevant in terms of selection directly into top management positions. The majority of the studies on ACs have looked at the relationship between the overall assessment rating made at the end of the AC and some subsequent performance measure of those selected. Recently, Arthur, Day, McNelly and Edens (2003) have argued that this may underestimate the predictive power of the method, and that looking at the relationship between the competency assessments made in the AC and subsequent performance might be more important.

As indicated earlier, Work Sample tests and simulations are usually a large part of an AC, but they may be used as stand-alone assessment tools also. When looked at against job performance criteria, correlations in the order of .32 are obtained, but with specific and (usually) lower level jobs, correlations may rise to .40-.50.

## Fairness

Whilst research on ACs generally is positive in terms of them showing little or no adverse impact in terms of gender or – in terms of the UK at least - ethnicity (Baron & Jarman, 1996), it has been argued that ACs often enshrine a male-dominated perspective on leadership in terms of the underlying model of leadership they represent and the way it is described in terms of competencies and behaviours.

Work sample tests are also usually found to be fair, and in some cases may offer an alternative to traditional tests of cognitive ability – for example, an in-tray or similar individual decision-making exercise (the Policy exercise at CSSB would fall into this category also), may be constructed to assess both intelligence and the ability to apply it in a practical way to examples of real issues and problems in the organisation concerned.

## Impact on the Candidate

ACs have a high level of 'face-validity' – that is, they look as though they are measuring relevant job behaviours, both to assessors and to candidates. Generally speaking, ACs are viewed positively by candidates, partly because of this and partly because they are seen to be thorough. However, they can have longer-term detrimental psychological effects on unsuccessful candidates (Fletcher, 1991). In terms of the top management group specifically though, ACs may have less appeal, not least because candidates at this level frequently do not wish their candidature to be widely known (which taking part in a group process is likely to undermine). In addition, in practical terms, the small number of candidates for a top post may be too few to mount an AC.

Work Sample tests, because they are just that, similarly have credibility and acceptability with candidates. Whereas an AC may not be suitable for top level management candidates for the reasons indicated, an individual decision making exercise might be more acceptable as it does not involve any more exposure of the candidate to a wider group than would a psychological test.

Cost

ACs represent the highest cost in terms of assessment methods, although just how expensive they are clearly varies according to the nature of the AC and its duration. They are usually only used where the numbers of candidates make the cost worthwhile. But there are some advantages offered by ACs that are difficult to quantify in terms of simple costs – their face validity with both candidates and assessors, the useful development information they yield for the candidates, the favourable external impression of the organisation they can give, etc etc.

Work Sample tests used as stand alone elements in a selection procedure involving interviews and, perhaps, personality questionnaires, may be much less expensive, especially if they are bought off the shelf (as many simulation exercises can be).

Meta Analysis References

Gaugler, Rosenthal, Thornton & Bentson, (1987); Schmidt & Hunter, (1998); Hough & Oswald, (2000); Arthur, Day, McNelly & Edens, (2003).

**SOME OVERALL CONCLUSIONS ON THE LITERATURE REVIEW OF DIFFERENT SELECTION AND ASSESSMENT METHODS**

The average validity achieved by the different methods reviewed is summarised in the Table below –

Structured Interviews	.44-.56
Tests of Cognitive Ability	.40 -.50
Assessment Centres	.37-.43
Personality Questionnaires	.40
Work Sample Tests	.32
Biodata	.30-.35
References	.25
Unstructured Interviews	.14-.33

As far as possible, these figures represent validities obtained where job performance rather than other criteria were used, and tend to focus more (but by no means exclusively) on managerial and graduate populations. They also reflect statistical corrections for various distorting factors that might otherwise lead to an underestimation of the true predictive power of the different methods.

However, it is of course rare for a single assessment method to be used in selection – even at the simplest level it is usual to have at least two (interview and references). Combining assessment methods can enhance the validity achieved, though it is not automatically the case. Thus, based on a number of studies, here are some average validities obtained from different combinations of assessment methods –

Tests of Cognitive Ability and Personality (Integrity) Measure	.65
Structured Interview and Tests of Cognitive Ability	.63
Tests of Cognitive Ability and Work Sample Test	.60

The above is illustrative only. For example, the first of the three relates to studies where a single personality-type measure focusing on Integrity was used – such a measure would only be relevant to certain selection situations, and mostly in jobs at a fairly low level. Indeed, none of these take into account job level. However, they do show how, up to a point, combining different methods can often lead to an increment in validity of assessment.

Turning now to a fundamental issue. Almost all the studies covered in the meta-analyses from which the above conclusions are drawn do not necessarily deal with very senior appointments – indeed, mostly they relate to selection situations where larger numbers of applicants and jobs (eg Graduate or middle-management selection) are involved. This is partly because there is organisational interest in large-scale recruitment and selection, but also because there are fewer methodological problems in doing research of this kind. *There are considerable difficulties in doing good quality research on selection to individual high-level posts* – because to do validation research generally requires an adequate sample size that can be followed up and evaluated using statistical techniques, a situation that does not usually arise for top level management positions. In the latter, there are usually only 3-5 candidates for a specific job, and only one will be appointed! The result is that there is extremely little research published on selection to top management.

Does this mean that meta-analysis findings are not relevant to top management selection? Fortunately, no, because what meta analyses have shown is that validity is *generalizable* from one group to another. By establishing that particular techniques have repeatedly been shown to have validity across such widely-differing candidate groups, countries, job roles and levels, they provide the justification for saying that in the majority of cases they are likely to be similarly valid for any new group or setting where they are applied (assuming of course that they are relevant to that situation). For the most part, then, there is every reason to believe that the validities established hold good for top management levels also – with some caveats, the principal one of which relates to cognitive ability.

As indicated earlier, where a group is already very highly selected on a particular attribute such as intelligence, adding further measures of it to the selection process may yield little additional information. One would be trying to make an increasingly fine discrimination across a very narrow range, and it is very doubtful in the case of intelligence that beyond a certain threshold level any increment will have much positive effect (though it might have a negative effect!). Thus, the value of using tests

of cognitive ability with a group who all hold good university degrees and have very successful track records may be small. However, where the candidate group is more varied in background and academic achievements, or where the nature of the job is, even for a top level group, especially intellectually demanding, then cognitive tests may have something more to contribute.

The other context in which tests of cognitive ability are important is where potential is being assessed. At middle and senior management levels, an individual may be coping well enough but actually be operating at the ceiling of their intellectual ability. Further promotion can lead to them struggling. Thus, for grades just below the top level, it is important to ensure that those being considered for progression have the intellectual power to make the next step. But, as indicated, this is for candidates who are currently outside the top management bracket.

The situation in respect to cognitive test use is likely to change in the future, though. In recent years, the number of “good” degrees has increased, and this criterion has thus become less effective as a means of discriminating between candidates. Employers recruiting graduates have become less willing to rely on degree class or A levels (which have suffered similar inflationary effects) and have resorted to using cognitive ability tests as their way of benchmarking intellectual power. Indeed, the universities themselves are looking at using objective tests as an adjunct if not a replacement for A level grades as a means of selecting students. Whilst this does not impact on top management appointments at the present time, some years from now it may well do – because the same problem of evaluating the academic ability of those coming forward may exist. In which case, testing (in some form) for cognitive ability may become more desirable at that point.

Leaving aside the validity of the different assessment methods, what of the other criteria they might be evaluated on? Fairness is the most important, and a ‘League Table’ might look like this, starting with the most fair at the top –

Assessment Centres  
Structured Interviews  
Work Sample Tests  
Personality Questionnaires  
Biodata  
References  
Cognitive Tests  
Unstructured Interviews

Again, though, this is in very general terms. For top management positions, it is unlikely that Cognitive Tests would come any lower than Work Sample tests. And whilst Biodata may well be a vehicle for bias, if constructed well it may be completely fair. With all the methods, it is to some extent a question of how they are used – even the best can be misused and applied in a way that produces inequity. The same goes for their impact on candidates; whether this is positive or negative largely depends on how the assessment method is presented. Candidates respond most favourably when –

- they can see (or have explained to them) the relevance of the measure to the job

- they are told how it will be used in the selection process
- they know it will be not be a pass-fail mechanism, used in isolation from the rest of the selection procedure
- they are offered feedback on their performance
- they can see it is being used/presented in a professional manner

In terms of cost, the least practical methods for top level appointments are Assessment Centres and Biodata, because of the small number of candidates involved, the political aspects of the situation (in relation to using ACs) and the expense of devising either for a one-off exercise. Although it is possible to buy off-the-shelf ACs or Biodata questionnaires, unless custom-designed they are likely to lack both relevance and validity in making top level appointments.

Having reviewed and discussed the general evidence on the effectiveness of selection methods, the next section will look at the literature specifically on top management assessment.

## **THE LITERATURE ON TOP MANAGEMENT SELECTION**

The difficulties of doing selection validation research for top management referred to a little earlier were reflected in the paucity of material found when various literature searches on top management assessment were conducted. In truth, virtually no genuine research articles of this kind were identified. What follows is a review of other material located, most of which is of somewhat tangential relevance.

There are a number of practitioner pieces in more general management journals which give their authors' or contributors' views on what practices should be followed. For example, a recent article in the *Director* magazine (published by the Institute of Directors) made the point that psychometric tests were little used at Board level in contrast to the situation for levels below this; it asserted that top management still went more by 'gut feel' on selection to the Board, and advocated a more systematic approach using psychometrics (Leslie, 2004). In similar vein, an article in *Directors and Boards* magazine urges the use of competency-based, behaviourally-focused interviews, with each interviewer covering 2-3 competencies, and the administration of psychometrics to "validate" the interview findings (Zwill, 1997). While such articles are interesting, they do not in themselves offer any proof of the effectiveness of the approach they propose.

Of the more scientific journal articles, one by Sessa and Taylor (2000) is interesting in that it was based on a systematic investigation which involved interviewing 494 executives in the top three layers of organizations about their experiences of selection at their levels and the success achieved of those selected. As the authors point out, top level executives often have little expertise in selection, but despite this they frequently do not seek to call on those who do have expertise, and indeed do not follow the kind of decision making process that they would in other fields of activity. The chief finding from the study was that selection based on group decisions rather than individual decisions had significantly better outcomes. Moreover, they found that a wider mix of membership on the selection committee also enhanced outcomes. Where this included input – direct or indirect – from subordinates or customers, this further improved selection. It is of course not surprising that having a degree of diversity in

the selection committee is a positive factor, as many psychological studies show that one of the strongest predictors of interview outcome is candidate-interviewer similarity; the tendency to recruit 'like self' at all levels is a strong one that needs to be counteracted, both for fairness and to ensure a healthy diversity of management resource is available to the organisation.

One of the best articles that just creeps into the top management category is that referred to earlier in the section on What to Assess, by Spreitzer, McCall & Mahoney (1997), which deals with identifying top management potential. However, as this chiefly dealt with the qualities to be assessed rather than selection per se, it will not be discussed further here.

Although they are not exclusively about top management, there are also some articles of relevance because they deal with a process that is often used in selection at this level, namely what is termed Individual Executive Assessment or Individual Psychological Assessment. This is the process where an occupational psychologist (usually but not exclusively external to the organisation) gives a battery of tests and an in-depth interview to the short listed candidates for senior posts, and writes an assessment report for those conducting the final interviews. This deals with an important point about the use of selection methods at this level. Much of what has been established about the validity of different assessment techniques rests on their application to a large number of candidates, a very different situation to looking at 3-5 candidates for a single senior post. A psychologist using a test or questionnaire knows what its general validity is, but in the context of an Individual Psychological Assessment he or she is not using the tests in a statistical way – they are being used as an input to a wider judgement process about the candidate and the appointment which seeks to combine all the information together to reach a decision. This is what is referred to as a clinical judgement process, which rests not simply on the quality of the tests, but also on the experience and competence of the psychologist and on the quality of information they have available from the organisation.

There is little good quality research on Individual Psychological Assessment processes (largely for the same reasons as there is little research on top management selection as a whole). However, what there is shows the need for caution (Highhouse, 2002). As indicated above, a great deal depends on what the psychologist assessor is given to work with in terms of the assessment briefing. What seems never to have been researched is how Individual Psychological Assessment Reports are actually used as an input to making the assessment decision. In the majority of cases, they are simply left to the interviewer(s) to consider, and only rarely is the psychologist asked to brief the interviewers directly or to respond to queries about the report. How effective is this if the interviewers have little knowledge or experience of the nature of the assessment process they are based on?

## **A BRIEF SURVEY OF TOP MANAGEMENT SELECTION PRACTICES OUTSIDE THE CIVIL SERVICE**

To add another dimension to the study, a small survey of top management selection practices in other organisations was included. Initially, this was intended to cover just 6-7 organisations in the time available, but it grew to something larger than that, though it should be emphasised that it is not a full, systematic survey.

The organisations approached fell into two categories –

- professional and other bodies who might have an overview of what practices were being adopted
- a diverse group of large organisations whose selection practices might stand as examples of what is typically done in this field externally

### **Findings from Overview Bodies**

In this category, the following organisations were contacted, and all replied –

- Institute of Directors (Policy Research Unit and Information Services)
- Chartered Institute of Personnel and Development (CIPD)
- Society of Chief Personnel Officers (SOCPO)
- Centre for Board Effectiveness, Henley Management College
- Odgers Ray & Berndtson (Executive Search Company)

The first thing to say is that none of them had conducted a formal survey on this topic. The nearest to having that kind of breadth of view was the Executive Search company, Odgers, who have extensive experience of top management selection practices across a wide range of organisations and were kind enough to share their observations. Their assessment was that 85-90% of the organisations they deal with used just interviews and references. Those with HQs in the USA or other close American links (eg Microsoft or IBM), and organisations in the retail and consumer sector (eg Kingfisher) were more likely to make use of psychometrics, work sample tests or assessment centres. They felt there were two influences that might work to increase the sophistication of the selection methods used. One was the influence of the adoption of the Higgs report on Non-Executive Directors (Higgs, 2003), which required the formal appraisal of Directors on an annual basis – Odgers representatives felt this might have a knock-on effect and persuade companies to be more professional in how they went about selection at this level. The other influence was that search agencies themselves were increasingly building-in psychometric and other assessment methods into their own evaluation of candidates. This latter point is interesting, as it raises the question of objectivity, as will be seen in one of the organisational examples below.

The National Talent Management Lead for the Society of Personnel Officers reported that in terms of local government, no common assessment process for top appointments exists. He reported, however, that it is now quite common is for an assessment centre to be specifically designed and tailored for top level appointments. This is undertaken by either the HR department or (now) more commonly by consultants. Otherwise, apart for the normal interview the most common tools used were personality questionnaires and work sample tests.

The other overview bodies, where they were able to comment at all, offered observations much in line with those of Odgers. It was pointed out that (executive) Director level appointments were often made from within the company rather than through external sourcing.

## **Findings from the Organisations Contacted**

In this category, the following organisations were contacted and all replied (though for various reasons, Unilever, BP and the BBC all replied too late to be included in the short time frame of the study) –

- BP plc
- Pearson plc
- N M Rothschild
- UBS
- Unilever plc
- BBC
- Shell
- EMI Music
- KPMG
- Freemantle Media

The input from these organisations did not significantly contradict the impression formed from the overview bodies. However, they did add considerably more detailed information and some interesting experiences and examples. The main conclusions from this part of the survey are that -

- In the private sector, the interviews are nearly all done on a sequential, one-to-one basis
- The amount of time spent on interviewing is typically considerably more than is usually the case in senior civil service appointments
- The interviewing is often done, in part at least, around the organisation's competencies
- Frequently the references are supplemented with informal 'soundings' about the candidates
- In some cases (eg Investment Banking) there is hard evidence on performance that can be used
- Majority of (Executive) Directors are appointed from within – and may have been through other assessment processes earlier in their career
- 'Fit with Board' is probably a higher priority than openness and fairness in the assessment process

To take the first three points, it is common for candidates to have 4-6 interview sessions (apart from any with the executive search company) each lasting an hour or two; small wonder that private sector candidates short listed for senior CS posts often comment with surprise at the selection process being so brief. However, one company approached said they sometimes put the candidates in front of the MD first, because if he did not like them there was no point carrying things further!

The fourth and seventh points perhaps reflect a difference between the values and constraints of the public sector compared to the private sector. Indeed, Higgs (2003)

noted a “high level of informality” surrounds the process of appointing NEDs and that almost half were recruited through personal contacts and friendships, with only 4% having a formal interview. He was favourably impressed by the open and transparent procedures in the public sector (Smerdon, 2004).

There are other variations specific to particular companies or sectors which will not be reported here as they have little relevance to the Civil Service. But there were a couple of interesting examples that are worth describing. The first of these was from a company that had experienced some problems with Director level appointments in recent years. They have now taken to forming a project group for similar appointments and include within this the coach who has been working with the level immediately below that of the person to be appointed – so that the strengths, weaknesses and style of the supporting team could be considered in the appointment process. They felt this had worked very well and provided a vital input. They also used an executive search company that did a psychometric and work sample-based assessment of candidates as part of their process.

This last point raises an issue that was alluded to earlier, and which is exemplified by another company contacted. They also had experienced a significant failure of a senior appointment, and similarly had used the executive search company to carry out the psychometric assessment of the candidates they fielded. In these circumstances, the question may be asked (and was by the company concerned) as to whether the search agents are the most detached and objective people to assess and report the outcome of the psychometrics? Is it a matter of Due Diligence that these two elements should be kept separate?

Finally, as an example of an unusually thorough process, the selection of Partners for KPMG is largely from within the company. It involves candidates having made the requisite progress through a number of earlier grades to the point where they may be considered to be in the field. To progress further, they need to be sponsored (in effect, recommended) by a Partner to go forward, after which the process consists of –

- a personality questionnaire and in-depth interview with a psychologist
- attending an assessment centre

For those getting through this hurdle, they are (up to a year) later required to make a presentation to a Partners’ meeting and go through one further interview. If they are successful, they can become Partners.

## **CONCLUSIONS AND RECOMMENDATIONS**

What are the implications of the findings of this study for the selection of candidates to the most senior Civil Service posts? The starting point for answering this question is to consider what is typically done when making such appointments at the present time; broadly, the process followed is -

- Drawing up of job and person specification
- Advertisement/Application
- Executive Search Company (ESC) create long list
- Department/OCSC representatives and ESC decide on short list
- References taken up (some may have been at the long list stage)
- In some cases, an Individual Psychological Assessment of the short list candidates is carried out, reports going to the Interview Panel

- Interview Panel, lasting up to an hour, and in some cases including a brief presentation by the candidate. The interview is usually loosely structured rather than being a competency-based, behaviourally focused one.

Commissioning of an ESC is only likely to take place where the post is open to external candidates. Individual Psychological Assessments are only commissioned in some competitions. Where they are used, they are generally done on all the short listed candidates, though there are occasions where this has not been the case.

It appears, then, that universal elements of the assessment process, references and a loosely structured interview, are those that have been demonstrated to have low or modest validity (in the sense that as interviews become less structured, they become less valid). Moreover, the amount of time given to interviewing the candidate is rather short by comparison with external organisations, and interview time is related to validity (up to a point, as one rises so does the other).

How might the present approach be improved? Taking each of the present stages of the process in turn (excluding Advertising, which is outside the scope of this study) -

(1) It is not possible here to comment in detail on the amount of time or methods used in drawing up the job and person specification – no systematically obtained information is available to the author on this. However, based on personal observation of the output of the process in the form of the statement of the job details and the candidate brief, they often seem to be couched in general terms with limited reference to the SCS competency framework, and – in the opinion of the author of this report – tend to fall into the trap of failing to differentiate between what is essential in terms of candidate attributes and what would simply be desirable. A more systematic analysis of the role/candidate requirements in terms of the SCS competency framework might be helpful, and indeed essential if other recommendations that follow are to be implemented. In addition, though, the SCS competency framework itself has been in existence for some while and might benefit from review. To some extent, this has already happened in the process of establishing the High Potential Development Scheme and its enhancement of the SCS competency set, but this review should be put on a more formal and systematic basis.

Apart from describing the role and the person requirement, more attention might also be paid to looking at the appointment in terms of the profile of the peers and direct reports the individual would be working with.

(2) Consideration might be given to including structured self assessment on the application forms

(3) The use of ESCs is helpful and appropriate in most instances. Given some of the points raised earlier in this report, it would seem most sensible to keep any Psychological Assessment of candidates independent of the ESCs' contribution.

(4) Short listing would also be better informed with a clearer analysis of the role and person requirements

(5) The validity of References is improved by including an element of structure around competencies. However, formal references may become increasingly difficult to obtain at all.

(6) Where Individual Psychological Assessments are used, more time and attention needs to be given to the briefing of the psychologist concerned, and to the use of the output (assessment reports) from this element of the process. Comments on the types of test that might be included are given below

(7) To increase the validity of the interview, it would be helpful to both include an element of structured, behaviourally focused interviewing, and to expose the candidates to longer interviewing time that is presently the case. It is not necessary, though, for *all* of the time devoted to interviewing to be competency-based.

All these improvements are consistent with the literature reviewed and with good practice elsewhere. However, there are several other issues and potential changes that need to be considered:

*\* Should Psychometric measures should be used, and if so, which?*

Given that the validity of cognitive ability and personality measures is good, the case for including them more routinely is strong. However, there are a number of important caveats. First, they need to be presented to candidates in a careful manner, and in such a way as to minimize less positive reactions. Second, their use needs to be guided and targeted by careful briefing about the role being selected for and its context. Third, at the present time, the psychometric measures that are most likely to be useful for top posts are personality and related (eg Leadership, Emotional Intelligence, etc) questionnaires. It would not be appropriate to seek to specify which ones in particular, as these would vary according to the role requirements and also according to the approach of the psychologist concerned. Fourth, cognitive ability tests might be valuable for selection to some top level posts, but careful consideration needs to be given as to whether there is any real danger of indirect discrimination arising. The situations where they might make the most valuable contribution are where the demands of the job place a special emphasis on intellectual capacity, and where the candidate field is varied in terms of its background and academic attainment.

*\* Are there other assessment methods that should/could be used?*

For the most part, it seems unlikely that Assessment Centres will be feasible alternatives for top management selection, on grounds of cost and candidate concerns. However, it might be worthwhile looking further into the claims that ACs are used often in local authority top management selection to see how the potential disadvantages are overcome (if indeed they are) in those settings. But Work Sample tests could be included in the assessment for top CS posts. It might be worth considering commissioning the development of one or more Individual Decision Making exercises that have generic relevance to SCS work. This could be in terms of broad fields of work or in terms of common problems that arise across fields and departments. Exercises of this kind would have the advantage of offering an alternative to cognitive ability tests, as they would not only test this quality but also its application in decision making. In addition, the high face validity of such exercises increases their acceptability to candidates, and it is less likely to generate indirect

discrimination than are cognitive ability tests. If a Work Sample test were given, it could be as part of an Individual Psychological Assessment or as part of a separate meeting at the hiring department.

*\* What kind of structured interviewing should be used?*

Although very highly structured interviews yield the best validity, they are less likely to be appropriate for the complexity of top level jobs, and also more likely to produce negative attitudes in the candidates. A competency-based interview – as it presently used in assessing candidates for the High Potential Development Scheme (HPDS) – offers a more suitable approach. Essentially, this takes the form of questions based on the behavioural descriptions of competencies. See Appendix A for an example, which is drawn from a section of the Guidance for Interviewers issued to Permanent Secretaries and NEDs who were serving as interviewers on the HPDS assessment process. Although it would be straightforward to develop similar but perhaps more general guidance notes for departments, it would be desirable to offer – if not insist on – training in competency based interviewing for all those likely to be involved in it. A similar approach was used when the Civil Service Selection Board went over to this kind of interviewing a few years ago.

None of these changes will be achieved without incurring additional costs, of course. However, in selection and assessment as in other fields, there is no quick fix, and any increase in time and money expended has to be viewed against the potential costs and damage resulting from making senior appointments that turn out to be less than successful. It is important, though, that any changes made be evaluated. There are various ways of doing this, but these are beyond the scope of the present report.

To summarise; it is **recommended** that –

**The current approach to making top management appointments in the Civil Service be reviewed and revised to increase its potential validity, and that this should principally involve -**

- E. A more refined and systematic way of describing the job and the person specification, at least in terms of the relevant competencies**
- F. The use of assessment devices other than just the interview - specifically psychometric measures and, possibly, Work Sample tests; commissioning of the development of the latter for exclusive CS use might be considered**
- G. Including a more structured, behaviourally focused approach to at least part of the interview, and greater time being devoted to the interviewing**
- H. Training for interviewers in any new approach adopted**

## APPENDIX A

### What is a Competency-Based Interview?

Not surprisingly, it is an interview based on the behavioural descriptions of each competency. For example, taking one of the High Potential Development Scheme competencies -

Focusing on Delivery	<ul style="list-style-type: none"><li>- Shows speed and flexibility in adjusting to new circumstances</li><li>- Provides the focus and cohesion to enable staff/colleagues to deliver effectively</li><li>- Takes personal accountability for managing outcomes and risks</li><li>- Demonstrates a record of improved results</li><li>- Exceeds customer expectations</li><li>- Identifies and makes best use of diverse resources to deliver goals</li></ul>
----------------------	---

The kind of questions that can be developed from this are –

- Tell me about an occasion recently when you had to demonstrate particular speed and flexibility to adjust to new circumstances – what did you learn from it?
- How do you think you help your staff or your colleagues to deliver more effectively?
- Do you feel you have taken any risks at work – how did you feel about that, and what responsibility did you have for the outcomes?
- What evidence can you bring forward to show how you have achieved a record of improved results in your division?
- Can you give me an example of how you have exceeded customer expectations in your present role?
- Have you had a situation where you had to make use of diverse resources to deliver your goals? Please tell me about it.

These are examples only – obviously there could be differences in wording or variations on these themes. Indeed, the behaviours listed under a competency heading are not meant to be an exhaustive list, but simply indicators of the kind of behavioural domain covered by that competency. The point is to seek behavioural evidence to make an assessment of the individual on that competency.

## REFERENCES

Anderson, N., Born, M., and Cunningham-Snell, N. (2001) Recruitment and Selection; Applicant Perspectives and Outcomes. In N. Anderson, D. Ones, H. Sinangil & C. Viswesvaran (Eds) International Handbook of Work and Organizational Psychology. Sage: London

Arthur, W., Day, E. A., McNelly, T. L., & Edens, P. S. (2003) A meta-analysis of the criterion related validity of assessment centre dimensions. *Personnel Psychology*, 56, 1251-54.

Anstey, E. (1977) A 30-year follow up of the CSSB procedure, with lessons for the future. *Journal of Occupational Psychology*, 50, 149-159.

Baron, H., & Jarman, K. (1996) Fairness in the assessment centre. In C., L. Cooper and I. T. Robertson (Eds) International Review of Industrial and Organizational Psychology, Wiley: London

Barrick, M. R., & Mount, M. K. (1991) The big 5 personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.

Bayne, R. Fletcher, C. & Colwell, J. (1983) Board and Sequential Interviews in Selection: An Experimental Study of their Comparative Effectiveness. Personnel Review, 12, 14-19.

Bliesener, T. (1996) Methodological moderators in validating biographical data in personnel selection. *Journal of Occupational and Organizational Psychology*, 69, 107-120.

Bright, J. E., & Hutton, S. (2000) The impact of competences statements on resumes for short-listing decisions. *International Journal of Selection and Assessment*, 8, 41-53

Briscoe, J. P. & Hall, D. T. (1999) Grooming and picking leaders using competency frameworks: Do they work? An alternative approach and new guidelines for practice. *Organizational Dynamics*, Autumn, p37.

CIPD (2004) Recruitment, Retention and Turnover Survey for 2004. Chartered Institute of Personnel and Development, London.

Dulewicz, V, & Gay, K (1997). Personal Competences for Board Directors: The main Dimensions and Role Comparisons. *Competency Journal*, 4, 3, 37 - 44.

Dulewicz, V (1995). Personal qualities for the selection, appraisal & development of directors. In Taylor, B (Ed) The Director's Manual. London: Institute of Directors

- Fletcher, C. (1991) 'Candidates' Reactions to Assessment Centres and Their Outcomes: A longitudinal study'. *Journal of Occupational Psychology*, 64, pp. 117–127.
- Fletcher, C. (1997a) 'Self-Awareness – A Neglected Attribute in Selection and Assessment'. *International Journal of Selection and Assessment*, 5, pp. 183–187
- Fletcher, C. (1997b) The impact of psychometric assessment: Fostering positive candidate attitudes and reactions. *Selection and Development Review*, 13, 8-11.
- Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., & Bentson, C. (1987) Meta-analysis of assessment centre validity. *Journal of Applied Psychology*, 72, 493-511
- Gay, K., & Dulewicz, V (1997). Personal Competences for Board Directors: An Investigation into the requirements of International Companies. *Competency Journal*, 4, 4, 35 - 44.
- Hall, D. T., & Moss, J. (1998) The new protean career contract; Helping organizations and employees adapt. *Organizational Dynamics*, Winter, pp 22-37.
- Harris, M. M.(1989) 'Reconsidering the Employment Interview: A review of recent literature and suggestions for future research'. *Personnel Psychology*, 42, pp. 691–726.
- Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004) Applicant reactions to selection procedures; an updated model and meta-analysis. *Personnel Psychology*, 57, 639-683.
- Highhouse, S. (2002) Assessing the candidate as a whole; A historical and critical analysis of individual psychological assessment for personnel decision making. *Personnel Psychology*, 55, 363-397
- Higgs, D. (2003) Review of the Role and Effectiveness of the Non-Executive Director. Department of Trade and Industry.
- Higgs, M., & Dulewicz, V. (1999) Making Sense of Emotional Intelligence. NFER-NELSON: Windsor.
- Hough, L. M., & Oswald, F. L. (2000) Personnel selection: Looking forward to the future. *Annual Review of Psychology*, 51, 631-664.
- Huffcutt, A. I., & Roth, P. L. (1998) Racial group differences in employment interview evaluations. *Journal of Applied Psychology*, 83, 179-189.
- Huffcutt, A. I., & Woehr, D. J. (1999) Further analysis of employment interview validity: A quantitative evaluation of interviewer-related structuring methods. *Journal of Organizational Behaviour*, 20, 549-560.
- Jones, L. & Fletcher, C. (2002) Self assessment in a selection situation: An evaluation of different measurement approaches. *Journal of Occupational and Organizational Psychology*, 75, pp.145-161

- Jones, L., & Fletcher, C. (2004) The Impact of Measurement Conditions on the Validity of Self-Assessment in a Selection Setting. *European Journal of Work and Organizational Psychology*, 13, 101-111.
- Keenan, A. (1978) Selection interview outcomes in relation to interviewer training and experience. *Journal of Social Psychology*, 106, 249-260
- Kroeck, K. G., & Magnusen, K. O. (1997) Employer and job candidate reactions to video conference job interviewing. *International Journal of Selection and Assessment*, 5, 137-142.
- Leslie, N. (2004) Director know thyself. *The Director*, September, 66-71
- Mathews, B. P., & Redman, T. (1998) Managerial recruitment advertisements – just how market oriented are they? *International Journal of Selection and Assessment*, 6, 240-248.
- McDaniel, M. A., Whetzel, D. L., Schmidt, M. L., & Maurer, S. (1994) The validity of the employment interview: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79, 599-617
- Moscato, S. (2000) Selection interview: A review of validity evidence, adverse impact and applicant reactions. *International Journal of Selection and Assessment*, 8, 237-247
- Ones, D. S., & Anderson, N. (2002) Gender and ethnic group differences on personality scales in selection; Some British data. *Journal of Occupational and Organizational Psychology*, 75, 255-276.
- Ones, D. S., & Viswesvaran, C. (2001) Integrity tests and other criterion focuses occupational personality scales used in personnel selection. *International Journal of Selection and Assessment*, 9, 31-40
- Robertson, I. T., & Smith, M. (2001) Personnel selection. *Journal of Occupational and Organizational Psychology*, 74, 441-472.
- Salgado, J. F. (1998) Big five personality dimensions and job performance in army and civil occupations; a European perspective. *Human Performance*, 11, 271-288
- Salgado, J. F. (1999) Personnel selection methods. In C., L. Cooper and I. T. Robertson (Eds) *International Review of Industrial and Organizational Psychology*, Wiley: London
- Salgado, J. F., & Moscoso, S. (2000) Comprehensive meta-analysis of the construct validity of the employment interview. Unpublished paper
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., & de Fruyt, F. (2003) International validity generalization of GMA and cognitive abilities: A European Community meta-analysis. *Personnel Psychology*, 56, 573-605
- Schmidt, F. L., & Hunter, J. E. (1998) The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.
- Sessa, V. I., & Taylor, J. J. (2000) Choosing leaders: More cooks make better broth. *Counselling Psychology Journal: Practice and Research*, 52, 218-225.
- Smerdon, R. (2004) *A Practical Guide to Corporate Governance*. Thomson, Sweet & Maxwell: London
- Spreitzer, G. M., McCall, M. W., & Mahoney, J. D. (1997) Early identification of international executive potential. *Journal of Applied Psychology*, 82, pp 6-29

Steiner, D. D., & Gilliland, S. W. (1996) Fairness reactions to personnel selection techniques in France and the United States. *Journal of Applied Psychology*, 81, 134-141.

Terpstra, D. E., Mohammed, A. A., & Kethley, R. B. (1999) An analysis of Federal Court cases involving nine selection devices. *International Journal of Selection and Assessment*, 7, 26-33.

Weisner, W. H., & Cronshaw, S. F. (1988) A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. *Journal of Occupational Psychology*, 61, 275-290

Zwill, M (1997) How to hire the right CEO. *Directors and Boards*, 22, 4, 22-27