

Personnel selection

Ivan T. Robertson* and Mike Smith

Manchester School of Management, UMIST, UK

The main elements in the design and validation of personnel selection procedures have been in place for many years. The role of job analysis, contemporary models of work performance and criteria are reviewed critically. After identifying some important issues and reviewing research work on attracting applicants, including applicant perceptions of personnel selection processes, the research on major personnel selection methods is reviewed. Recent work on cognitive ability has confirmed the good criterion-related validity, but problems of adverse impact remain. Work on personality is progressing beyond studies designed simply to explore the criterion-related validity of personality. Interview and assessment centre research is reviewed, and recent studies indicating the key constructs measured by both are discussed. In both cases, one of the key constructs measured seems to be generally cognitive ability. Biodata validity and the processes used to develop biodata instruments are also critically reviewed. The article concludes with a critical evaluation of the processes for obtaining validity evidence (primarily from meta-analyses) and the limitations of the current state of the art. Speculative future prospects are briefly reviewed.

This article focuses on personnel selection research. Much contemporary practice within personnel selection has been influenced by the research literature, but it is clearly not the case that there is a systematic linear flow from the research literature into the work of practitioners. The situation is much more complex. For example, assessment centres were designed originally to meet a clear practical need. Their original design was heavily influenced by psychologists. There was, however, relatively little research into some of the specific components of assessment centres, when they were first used for practical personnel selection decisions, in the armed services and in commercial settings. Research into the overall validity of assessment centres and into the validity, adverse impact and utility of many of the component parts of assessment centres followed from these highly practical beginnings. In turn, this research has informed the practice of contemporary assessment centres. Similarly, complex interplay takes place for all other selection methods. This article, then, as well as reflecting contemporary research interests, as far as personnel selection is concerned, will also inevitably reflect contemporary practice to some degree.

*Requests for reprints should be addressed to Prof. Ivan Robertson, Manchester School of Management, UMIST, PO Box 88, Manchester M60 1QD, UK (e-mail: Ivan.Robertson@umist.ac.uk).

The traditional model for selection and assessment practice has not changed for many years. Smith and Robertson (1993) indicated the major sequence of events involved in the design and validation of any personnel selection system. The traditional system involves the initial detailed analysis of the job. This analysis is then used to indicate the psychological attributes required by an individual who may fill the job effectively. In turn, personnel selection methods are designed with the goal of enabling those responsible for selection to attract and evaluate candidates' capabilities on these attributes. A validation process is used to assess the extent to which the personnel selection methods provide valid predictors of job performance, or other criterion variables such as absenteeism or turnover.

Probably the most significant change within the personnel selection research literature in the last decade or so has been the increased confidence that researchers have in the validity of most personnel selection methods. This increased confidence has arisen from the results obtained by investigators using meta-analysis (Hunter & Schmidt, 1990). Meta-analytic studies of a wide variety of personnel selection methods have indicated that when the artefactual effects of sampling error, range restriction and measurement unreliability are removed, the 'true' validity of personnel selection methods is much higher than originally believed. Many selection methods have been subjected to a detailed meta-analytical review. One of the best lists of meta-analyses of selection methods is contained in Schmidt and Hunter's (1998) article where they identify meta-analyses of 17 methods of selection. Figure 1 is based on Schmidt and Hunter's review and shows the validity, estimated by meta analyses, of many selection methods. The numbers on the right show the validities when overall job performance ratings—usually by superiors—are used as criteria. The figures on the left of the diagram show the validities obtained when progress during training is used as the criterion. The two sets of results are very consistent, even though there are fewer meta-analyses available for training criteria. The table is also interesting because it casts light upon the development of thought concerning criteria. In the mists of time (before 1947!), psychologists sought a single criterion against which they could calibrate selection instruments. Following Dunnette's advice in the 1960s to junk THE criteria, they sought to obtain data for a bank of diverse criteria. The use of multiple criteria had disadvantages: they were often impractical or costly to collect and often led to confusion because they produced significantly different validities. Some order was restored in the 1980s, when criteria were organized into three groups: production criteria, personnel data and judgements. Schmidt and Hunter's tables imply that, in practice, psychologists have combined production criteria and judgemental criteria (usually supervisory ratings) to produce two *de-facto* categories. The hegemony of supervisory ratings as a criterion has, if anything, been strengthened by the current emphasis on contextual and citizenship behaviours as an element of job performance (see later in this paper): supervisory ratings are one of the few ways that such occupational citizenship can be gauged.

As far as criteria are concerned, the most significant changes within personnel selection research concern the broadening of the construct of job performance

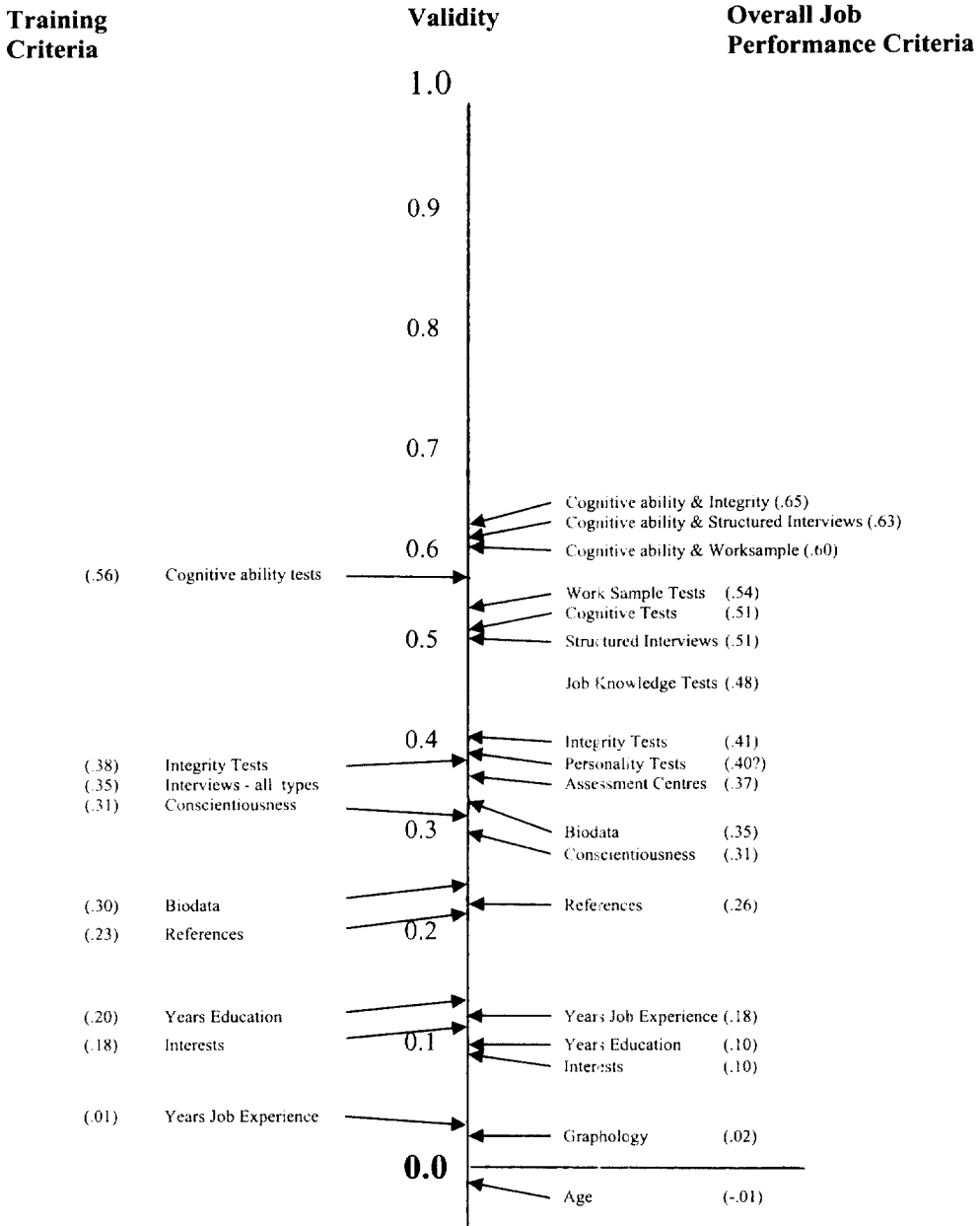


Figure 1. Accuracy of Selection Methods.

such that job performance includes not only effective performance of the relevant tasks but also contextual performance or organizational citizenship behaviour (Borman & Motowidlo, 1997; Coleman & Borman, 1999).

The area of personnel selection that has developed least and seems increasingly problematic is job analysis. The traditional role of job analysis within the personnel selection paradigm is to provide a fixed starting point for all subsequent steps in the process. Anyone who has given the remotest thought to contemporary organizational life will recognize that jobs are no longer anywhere near as stable as they were, even 10 or 15 years ago. At one time, the lifespan of a work-related technology and the career span of individual employees were reasonably well matched. Nowadays, technologies, work practices and even organizational forms come and go within the lifetime of an individual or even within a specific decade. This means that in many selection situations, the requirement to understand the job is made particularly complex and difficult, because the job in question is likely to be radically different in ways that are very difficult to predict within as little as 5 or maybe 10 years.

In their review of personnel selection, Hough and Oswald (2000) noted the importance of the changing nature of work and the difficulties that this presents for traditional job analysis. They indicate that, in recognition of the increasingly rapid changes that are taking place in the workplace, many researchers and practitioners now conduct analyses that focus on tasks and the cross-functional skills of workers, rather than traditional job analysis with its focus on more static aspects of jobs. In particular, they noted the use of O*NET as a flexible database that contains information about both work behaviours and worker attributes, including information on personality variables, cognitive variables, behavioural and situational variables (Petersen, Mumford, Borman, Jeanneret, & Fleishman, 1999). This modern approach to job analysis has many useful attributes but clearly cannot find a way of predicting the future requirements of jobs with any degree of certainty.

Major issues

This article is not intended to provide a comprehensive review of the research literature concerning personnel selection in the last decade or so. The brief from the editors of this special issue to the authors included a requirement to '... impose themes over what is powerful in the areas, synthesize *some* of the existing literature, make clear what we know and what we do not yet know'. This article has been written with these requirements in mind. Recent reviews of the personnel selection research literature provide a detailed account of the current state of the art. Hough and Oswald (2000) and Salgado (1999), have both provided detailed and comprehensive reviews of the personnel selection research literature. The review of Hough and Oswald (2000) covers the whole area of personnel selection from job and work analysis through to professional, legal and ethical standards. Salgado's (1999) review concentrates on personnel selection methods.

Both Hough and Oswald (2000) and Salgado (1999) provide convincing evidence of the earlier statement that the results of meta-analysis have provided strong evidence of good validity for many personnel selection methods. Several methods, including cognitive ability tests, personality questionnaires, interviews, assessment centres and biodata, have all been shown to have reasonably good validity. One major area that causes difficulties for both researchers and practitioners relates to

the fairness and adverse impact of personnel selection methods. Adverse impact occurs when members of one sub-group are selected disproportionately more or less often than the members of another sub-group. In the United States, this has caused problems for a number of years in relation to people from different ethnic minority groups. Similar problems have arisen in the United Kingdom and other countries. In general terms, cognitive ability creates most problems when it comes to adverse impact. Even when combined with methods that have a lower adverse impact, cognitive ability frequently creates adverse impact problems for selection systems (Bobko, Roth, & Potosky, 1999; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997). Some personnel selection methods that do not show an adverse impact, e.g. personality questionnaires (Ones & Visweveran, 1998) are being more widely used (Shackleton & Newell, 1997). Other methods, such as biodata, which show minimal adverse impact and reasonably good levels of validity, continue to be used relatively little (Bliesener, 1996; Shackleton & Newell, 1997).

For a number of years, the personnel selection research literature has been dominated by studies that have explored the validity of specific personnel selection methods. The development of meta-analysis and subsequent use of the technique to provide better estimates of the validity of a whole range of methods provided a significant step forward. Validity evidence concerning a wide range of methods is now reasonably stable and a number of topics such as those mentioned above, i.e. job and work analysis, criterion measurement, adverse impact and fairness, are beginning to be increasingly visible within the personnel selection research literature. They are also important within the practitioner domain. Some other issues, which are of growing importance in personnel selection research and practice, are: selection procedures that take account of the group within which the candidates will work (i.e. team member selection); selection for multi-national organizations, where recruits are required to work across different cultures; the reactions of applicants to personnel selection experiences and the criterion-related validity to be obtained from combining different selection methods. All of these issues are considered later in this article.

Describing jobs and worker characteristics

Traditionally, job analyses are divided into two main kinds: task-orientated job analysis and worker-orientated job analysis (see Sandberg, 2000).

Task analysis

During the period of this review, relatively little work has been conducted on job analysis in its strict sense. Hough and Oswald (2000) do not include a single reference to task analysis in their lengthy review. Salgado's (1999) review does include a section on job analysis, but it deals mainly with the second stage—worker characteristics. Sandberg (2000) maintains that task analysis has the advantage of identifying essential activities and giving concrete, detailed descriptions that can be readily applied. Contemporary research on task analysis is sparse. One empirical

study is Landis, Fogli, and Goldberg's (1998) use of future-orientated job analysis. They give a description of the steps needed to obtain an inventory of future tasks for three new entry-level positions in an insurance organization. Their method seems particularly relevant in an environment that is fast paced and electronic. They recommend that a panel of SMEs (Subject-matter Experts) should include both incumbents and non-incumbents because non-incumbents are better able to take a strategic view of future developments. Sanchez (2000) gives an example of 'strategic job analysis' with air-traffic controllers and mentions applications by Porter (1985) and Wright and McMahan (1992).

The paucity of empirical research on task analysis would imply either that the topic is unimportant or that we have reached a satisfactory state of knowledge in this area. For details of the psychometric properties of task analysis, we need to rely on older work such as that of Sanchez and Fraser (1992). Work on theories of job performance such as Borman and Motowidlo (1993) and Campbell (1994) that distinguish task performance and contextual performance is largely ignored in contemporary task analysis. In practice, there is still the tendency to focus upon specific, discrete tasks and ignore contextual aspects such as maintaining morale, courtesy and other citizenship behaviours listed by Viswesvaran and Ones (2000).

Person specification (worker-orientated analysis)

There has been more work in the area of Person Specification, although much of this has been conducted under the headings of competency determination or worker-orientated job analysis. The knowledge, skills and abilities (KSAs) that appear in a person specification are usually determined in the light of results from a task analysis. However, many practitioners go directly to KSAs by asking subject-matter experts to identify the competencies required for the job. Very little is known about the validity, reliability or other psychometric properties of this process.

A number of instruments has been developed to establish the personality requirements of a job. Rolland and Mogenet (1994) developed an ipsative system that identifies the most salient 'big five' personality factors for performance in a given job. Raymark, Schmidt, and Guion (1997) developed a Personality-related Position Requirements Form that also aids the identification of relevant personality traits. Hogan and Rybicki (1998) developed an instrument that creates a profile of job requirements that can be used in conjunction with the Hogan Personality Questionnaire. Westoby and Smith (2000) developed a 60-item questionnaire, completed by subject-matter experts, that indicates which 16 PF scales are likely to be important determinants of performance in a specific job. The extent to which these instruments identify personality traits that provide good predictions of subsequent performance has not been evaluated. These instruments are too new for an authoritative assessment of their psychometric properties. Their structured, systematic approaches using multiple informants *should* increase the reliability and validity of 'person specifications', and preliminary evidence suggests that this is so. A more accurate specification of the personality requirements of a job would mean

that irrelevant dimensions can be eliminated, and the 'average validity' of the remaining dimensions should be higher.

Models of work performance and criteria

The success of a selection system is gauged against criteria. Often, the choice of these criteria is decided by convenience. The attenuation or contamination arising from the use of poor criteria results in a systematic underestimation of the true validity of the selection methods. Murphy (2000) wrote 'validity coefficients are usually larger than, and more consistent than, a casual review of the literature . . . would suggest'. Problems with criteria can be mitigated in two main ways. First, they can be chosen more carefully, on the basis of a task analysis, as indicated above. Second, they can be chosen on the basis of models of work performance.

Viswesvaran and Ones (2000) gave an excellent overview of models of job performance, and their paper is worth considering in some detail. They give useful examples of Borman and Motowidlo's (1993) distinction between task and contextual performance. *Task performance* is defined as 'the proficiency with which incumbents perform activities that are formally recognised as part of the job; activities that contribute to the organization's technical core either directly by implementing a part of its technological process, or indirectly by providing it with needed materials or services'. Task performance is likely to be contained in most job descriptions and has, perhaps, been over-emphasized and over-used in developing criteria. *Contextual performance* is very similar to the concept of organizational citizenship, organizational spontaneity, extra-role behaviour and pro-social organizational behaviour. Contextual performance consists of behaviour that promotes the welfare of individuals or groups, and it includes components such as altruism, courtesy, civic virtue, conscientiousness, making constructive suggestions, protecting the organization, developing oneself and spreading goodwill.

Viswesvaran and Ones (2000) provided a very useful collation of the dimensions of job performance, which have been identified by some other researchers. The dimensions are organized into two groups: dimensions for single occupations and dimensions that are applicable to all occupations. These findings are summarized in Table 1. The lists of dimensions of work performance are remarkable in several ways. First, they contain a substantial number of tautological or vague phrases such as 'overall job performance'. Second, they show remarkable variety and little common ground. It would seem that we have a considerable way to go before research provides us with a common set of variables underlying work performance. Some of that difficulty may lie in deficiencies in the scales used to measure work performance (see Arvey & Murphy, 1998). Measurement of work performance invariably takes one of two forms: counts of output or other behaviours enshrined in organizational records or ratings by other people. Unfortunately, organizational records are often incomplete, irrelevant or seriously contaminated by artefacts. Ratings by other people are often unreliable and subjective. To make matters worse, job performance is usually measured as a static phenomenon, whereas work performance is dynamic in many ways. Error variance produced by these

Table 1. Dimensions of job performance from various studies

<i>Job performance dimensions for specific jobs</i>	
Entry-level service jobs: Hunt (1996)	Entry-level military jobs: Campbell, McHenry, and Wise (1990)
(1) Adherence to confrontational rules	(1) Core technical proficiency
(2) Industriousness	(2) Soldiering proficiency
(3) Thoroughness	(3) Effort and leadership
(4) Schedule flexibility	(4) Personal discipline
(5) Attendance	(5) Physical fitness and military bearing Borman and Motowidlo (1985)
(6) Off-task behaviour	(6) Allegiance (commitment and socialization)
(7) Unruliness	(7) Team work (socialization and morale)
(8) Theft	(8) Determination (morale and commitment)
(9) Drug misuse	Campbell (1990)
Managers: Conway (1999); Borman and Brush (1993)	(1) Job-specific proficiency
(1) Leadership and supervision	(2) Non-job-specific proficiency
(2) Interpersonal relations and communications	(3) Written and oral communication
(3) Technical behaviours, e.g. administration	(4) Demonstrating effort
(4) Useful behaviours, e.g. handling crisis	(5) Maintaining personal discipline
	(6) Facilitating help and team performance
	(7) Supervision
	(8) Management or administration
<i>Job performance dimensions for jobs in general</i>	Viswesvaran (1993)
Bernadin and Beatty (1984)	(1) Overall job performance
(1) Quality	(2) Productivity
(2) Quantity	(3) Effort
(3) Timeliness	(4) Job knowledge
(4) Cost-effectiveness	(5) Interpersonal competence
(5) Need for supervision	(6) Administrative competence
(6) Interpersonal impact	(7) Quality
	(8) Communication competence
	(9) Leadership
	(10) Compliance with rules

deficiencies in measurement can obscure underlying dimensions of work performance. Even worse, error variance can be mistaken for true variance and wrongly interpreted.

It would appear that, despite these measurement problems, about 50% of the variance in performance ratings is common, and there is a 'g' factor of work performance that is analogous to the 'g' factor of cognitive ability. If a 'g' factor of job performance does exist, research effort would be needed to develop ways in which it can be measured and to establish how and why it varies from person to person. Hunter, Schmidt, Rauchenberger, and Jayne (2000) suggest that a 'g' factor in work performance is determined by two characteristics: general mental ability and conscientiousness. The 'g' factor in work performance is very similar to the universal domain suggested by Smith's (1994) theory of the validity of predictors.

Table 2. Ratings of the importance of attributes in selection decisions according to one Scottish Consultancy Organization

Attribute	Rating (0–3)	Percentage of maximum
Honesty	2.89	96
Conscientiousness	2.77	92
General ability	2.72	91
Potential	2.65	88
Experience	2.56	85
Adaptability	2.55	84
Drive	2.45	82
Experience	2.34	78
Fit of values	2.39	80
Job knowledge	2.25	75
Social ability	2.19	73
Health	2.10	70
Professional qualification	2.03	68
Accident/appearance	1.99	66
Academic qualifications	1.78	59
Years with other firms	1.61	54
Similarity to future colleagues	1.50	50
Age	1.50	50
Outside interests	1.20	40

It is interesting to compare explicit models of job performance with the implicit models used by selectors. Table 2 shows data derived from a survey by Scholarios and Lockyer (1999) of the way in which small consultancy organizations in Scotland recruit professionals such as accountants, architects, lawyers and surveyors. Typically, each consultancy recruited two professionals each year. Scholarios and Lockyer presented interviewees with a list of candidate qualities. It can be seen that there is considerable correspondence between the explicit model of Hunter *et al.* and the implicit model. Both models stress the importance of conscientiousness/integrity and general ability. However, the explicit model places general ability higher than conscientiousness. Huang (2000) surveyed HRM practices in four multinational organizations in Taiwan and found that multinationals tend to select employees on the basis of job skills rather than their fit with the organizational culture.

Newell (2000) speculated that the qualities underlying job performance will change in the light of changes in the economy and business organization. She argued that Knowledge Management—the way in which an organization creates, utilizes and stores the expertise that underlies its products—is the current management fashion. Proponents claim that Knowledge Management will be the key to competitive advantage in an ‘information age’ in the same way that the management of capital and physical resources was the key advantage for the old

'smokestack' industries. In the 'Knowledge Era', jobs will change rapidly, and a key attribute will be the ability to form social networks that facilitate access to a huge pool of information held by other people. Some of these changes are also identified by Schmitt and Chan's (1998) excellent text 'Personnel Selection: A Theoretical Approach'. They suggest that in the future, there will be an increase in the speed of technical change, the use of teams, communication, globalization and service orientation. Such changes would increase the importance of team selection and expatriate selection. This might suggest that job analysis will be inappropriate and that the ability to share information will be a key determinant of job performance. However, job analyses are not necessarily static, and many incorporate a responsibility to develop, change and be flexible in response to environmental changes. It could also be that cognitive ability will be at a premium in a business world where huge quantities of information need to be assimilated and processed.

Attracting applicants

Once a job has been defined and the qualities of the ideal applicant specified, it is necessary to attract applicants. Recruitment remains an area that escapes the attention of many researchers. Indeed, Mathews and Redman (1998) claimed that the area of recruitment has been neglected.

Purcell and Purcell (1998) identify changes in the labour market such as outsourcing (sub-contracting specialist services from outside providers), in-sourcing (the use of agency staff and temporary employees) and the establishment of a cadre of core workers. These trends imply the development of separate labour markets that need different recruitment strategies. Mathews and Redman (1998) surveyed the views of 191 managers and executives. They discovered that 54% look at job adverts weekly, while 83% look at job adverts at least once a month. For 80% of managers, this browsing is not connected with job-seeking but is concerned with comparing salaries and keeping in touch with the marketplace. Mathews and Redman asked the sample to rate the importance of a menu of 21 items that might appear in an advert. The items achieving the highest rank were description of the job, salary, key responsibilities, career prospects, closing date, company details, location and experience needed. When applicant views were compared with the items that appear in actual adverts, only a moderate correlation was found. Actual adverts seem to under-play the importance of promotion prospects and closing date whilst over-emphasizing personal characteristics. The 'applicants' were also shown a list of personal characteristics that frequently appear in adverts, and they were asked to rate whether their inclusion would encourage or discourage an application. Adjectives that are most likely to discourage an application are: analytical, creative, innovative, energetic and interpersonal. It would be fascinating to know the connotations that give these words a less favourable impact. The survey also showed that managers prefer to apply using a CV (55%) or an application form (41%). The methods chosen by firms tend to under-represent these media (40% and 12%, respectively). Recommendations to

respond by making a telephone call are not popular with managerial applicants (3%).

Applicant perceptions

Applicant perceptions play a key role in recruitment since negative views will inhibit some people from putting themselves forward. Furthermore, negative attitudes might affect the motivation, and thus subsequent performance, of applicants, when they take tests or attend interviews. Applicant perceptions might also influence the degree to which they are prepared to give answers that are truthful. Research in this area has followed two main approaches: determining aspects of the selection procedure that tend to be disliked and attempting to explain why applicants develop negative views.

Considerable research, prior to the period of this review, attempted to determine applicants' views on selection methods (see Hough & Oswald, 2000, p. 647; Salgado, 1999, p. 29). Many authors such as Rynes and Connelly (1993), approached the topic at a macro-level by asking people (often students seeking jobs in their final year) to rate various selection devices on the degree to which they liked each selection method. As a gross over-simplification designed to incite further reading, the results of these studies show that candidates tend to like work samples and unstructured interviews but tend to dislike tests. Conway and Peneno (1999) compared applicant reactions to three types of interview questions. Moscoso (2000) suggested that job knowledge or job experience may be a possible moderating variable. Candidates with extensive knowledge or experience within a specific job may react more favourably to structured interviews because they are better equipped to answer these types of questions. Kroeck and Magnusen (1997) examined candidate and interviewer reactions to video conference job interviews. In general candidates preferred traditional face-to-face interviews. Tonidandel and Quinones (2000) explored reactions to adaptive testing and found that candidates like the capability provided by adaptive testing to 'skip' questions but thought that the way in which adaptive testing presents different candidates with different sets of questions was unfair.

Other studies approached the issue of candidate reactions at a micro-level. They tried to identify specific facets of selection methods that caused dislike. For example, some very early work focused on the impact of the non-verbal behaviour of interviewers. Again, as an over-simplification, interviewees liked interviewers who emit a high-level of positive non-verbal behaviour such as nods and smiles. Bretz and Judge (1998) examined the way in which information contained in a written realistic job preview influenced candidate perceptions. Not surprisingly, they found that negative information reduced the organization's attractiveness, whereas a high salary and the prospect of being invited to visit the organization for a second interview increased attractiveness.

Very influential papers by Gilliland (1993) and Schmitt and Gilliland (1992) focused attention upon the role that perceptions of fairness play in determining applicant attitudes. Subsequent investigators differentiated procedural justice from distributive justice and investigated the effects of both types in fostering positive

attitudes. Generally, the evidence suggests that feelings of injustice engender negative attitudes towards selection systems. Thorsteinson and Ryan (1997) investigated the effect of selection ratios and outcome (accepted or rejected) on applicant perceptions. They found that selection ratio had very little effect, while the outcome in terms of receiving a job offer was a significant variable. Elkins and Phillips (2000) investigated perceptions of fairness (both procedural and distributive) and job relatedness of a biodata instrument. The most important determinant of both kinds of fairness was the selection decision. If the selection decision is in the candidate's favour, the procedure is likely to be viewed as fair. If the selection decision goes against a candidate, it is likely to be viewed as unfair.

Ployhart and Ryan (1998) also used a laboratory simulation to investigate candidates' perceptions of fairness. The results again showed that the outcome, receiving a job offer, was the strongest determinant of perceptions of fairness. The influence of procedural violations such as too much or too little time to complete the test had an asymmetric effect. A rule violation that produced favourable treatment was not perceived as unfair, whilst a rule violation that produced unfavourable treatment was regarded as unfair. As a simplification, it would seem that an outcome that is favourable to the individual or a treatment that is favourable to the individual is seen as fair, while an unfavourable outcome or unfavourable treatment is seen as unfair. Notions of fairness or unfairness very much depend on the individual's view of the impact. It seems that the concepts of fairness and self-interest are closely intertwined!

Chan, Schmitt, Jennings, Clause, and Delbridge (1998) investigated perceptions of fairness, job relevance and a self-serving bias. They suggested that when self-esteem is threatened by rejection, candidates reduce the threat by perceiving the system as unfair. Chan *et al.*'s study is notable because it used genuine applicants for jobs.

Legal aspects of candidate perceptions

In an extreme case, an adverse reaction from applicants can lead to a legal challenge. Gilliland (1993) suggested that such challenges are less likely if candidates feel that the selection method has four characteristics: (1) job relatedness, (2) an opportunity for the candidate to demonstrate ability, (3) sympathetic interpersonal treatment, (4) questions that are not considered improper. Generally, about 64% of legal challenges are made by 'new hires' rather than someone who is already working for the organization.

Terpstra, Mohammed, and Kethley (1999) analysed 158 US Federal Court cases. They compared the proportion of legal challenges with the proportion that would be expected on the basis of the frequency with which a selection method is used. The methods most likely to be challenged in proportion to their use are tests of physical ability (350%), ability tests (230%) and unstructured interviews (200%) (figures in brackets represent the percentage of over-representation). These figures need to be interpreted with care. Unstructured interviews are used very frequently and, in absolute terms, are the method most likely to be challenged. However, work sample (20%), assessment centres (20%) and structured interviews (50%) were least

likely to be challenged. The number of challenges is not necessarily the most important factor since many challenges are defeated. Terpstra *et al.* (1999) found that all the challenges against assessment centres and structured interviews were dismissed. The large majority of challenges against work samples (86%) and ability tests (67%) were also unsuccessful. About half the cases against unstructured interviews (59%) and physical ability tests (58%) were resisted. These results show that employers who are solely motivated by avoiding trouble should not use tests of physical ability, tests of mental ability or unstructured interviews. Employers who are prepared to argue their case in court, however, should be most wary of physical ability tests and unstructured interviews.

Research results on candidates' perceptions need to be viewed in the light of the methodologies used. Most studies, with the notable exception of that of Chan *et al.* (1998), have used simulations and student participants. Furthermore, the measurement of key variables is less than ideal. Fairness, for example, is frequently operationalized by responses to questions such as, 'would you recommend this to a friend?' or 'would you accept a job offer?' The answers to these questions are heavily contaminated by factors such as the predispositions of a friend or alternative prospects of employment. Often, the scales used to measure fairness have few items, and the use of a scale with a small number of items can cause problems with reliability. Consequently, only broad findings should be interpreted. The one finding that emerges from most studies is that candidates' perceptions are determined by the offer, or not, of a job. An unkind critic might observe that the only reliable application from this genre of research is that organizations should improve candidate perceptions by offering more candidates a job.

Personnel selection methods

Since the very earliest research on personnel selection, cognitive ability has been one of the major methods used to attempt to discriminate between candidates and to predict subsequent job performance. During the 1980s, several meta-analytic studies of the criterion-related validity of cognitive ability tests produced conclusive results (see Schmidt & Hunter, 1998). These studies have produced clear findings concerning both the validity of cognitive ability and the extent to which cognitive ability is fair when used in testing people from different ethnic groups. The findings have shown that cognitive ability provides criterion-related validity that generalizes across more or less all occupational areas. The results concerning differential validity have also been reasonably conclusive, indicating that cognitive ability provides accurate predictions of subsequent work performance that are more or less equally accurate across different ethnic groups. In other words, cognitive ability testing does not provide differentially 'unfair' (Cleary, 1968) predictions for members of different ethnic minority groups. Of course, these scientific findings do not imply that it is wise to use cognitive tests for all selection purposes. As already noted, the problems of adverse impact are difficult to cope with, given that members of some minority groups obtain lower scores on such tests. This state of affairs, i.e. no differential validity but poorer scores for some groups is challenging for people involved in the design and validation of selection procedures. There is no simple solution.

A further conclusive finding has shown that the core dimension of cognitive ability (general mental ability, or 'g') is the key component in providing predictions of subsequent job performance. The use of specific abilities (i.e. sub-components of general mental ability) does not enhance the predictions provided by the use of 'g' alone (Olea & Ree, 1994; Ree, Earles, & Teachout, 1994). Traditional ability tests have focused on assessing specific competencies that have been considered since the early 1900s to underlie intelligence (see, for example, Carroll, 1993). These factors (now conceptualized as fluid intelligence, crystallized intelligence, visualization, retrieval and cognitive speed) still underlie the majority of cognitive ability tests used today.

One area of interest related to cognitive ability concerns the development of 'practical intelligence' (Sternberg & Wagner, 1986, 1995). For these authors, practical intelligence can be distinguished from the kind of intelligence that lies behind success in academic pursuits. Practical intelligence is unrelated to formal academic success but related quite directly to the abilities that people develop in seeking to attain their goals in every-day life. Although the ideas put forward by Sternberg and Wagner are interesting, there is, so far, little conclusive evidence that practical intelligence is any more effective at predicting subsequent job performance, or indeed provides anything that is significantly different from general mental ability. There are few published studies with reasonable size samples that have investigated the criterion-related validity for tacit knowledge, and where this has been done (e.g. Sue-Chan, Latham, Evans, & Rotman, 1997, cited in Salgado, 1999), the results have shown that validity is modest and provides little gain beyond what is already obtainable from tests of general mental ability.

Another related, but different, concept from practical intelligence is emotional intelligence. Emotional intelligence (Goleman, 1996) relates to the ways in which people perceive, understand and manage emotion. Amongst practitioners, the use of the term emotional intelligence is widespread, but a thorough search of the scientific literature failed to provide any studies that demonstrated the criterion-related validity of emotional intelligence for any specific occupational area.

Personality

Until quite recently, personality was not a popular personnel selection method. Indeed, as recently as 1965, Guion and Gottier concluded that it was impossible to conduct a review of the criterion-related validity of personality because too few studies were available in the literature. The 1990s have seen a huge growth in the use of personality assessment within personnel selection practice and research studies designed to evaluate and explore the role of personality within personnel selection (e.g. Barrick & Mount, 1991; Frei & McDaniel, 1997; Ones, Visweveran, & Schmidt, 1993; Salgado, 1998; Tett, Jackson, & Rothstein, 1991). All of these studies adopted a meta-analytic procedure and provided positive evidence for the criterion-related validity of personality. From an initial position of scepticism concerning the contribution that personality could make to effective personnel

selection, researchers and practitioners have moved to a position where there is confidence that personality can play a role. From this base, more refined research questions have begun to be investigated. These include several interesting questions such as: the level of analysis that should be used when utilizing personality for personnel selection and assessment purposes (e.g. the big-five level or more specific factors); the extent to which conscientiousness or a broad factor relating to integrity acts as the single best predictor for personality, in much the same way that general mental ability works in the cognitive ability domain; and the role of intentional or other forms of distortion in influencing candidate responses and the incremental validity provided by personality assessment over and above that which is provided by other more established methods of personnel selection, such as general mental ability.

The research focusing on the use of level of analysis best used for personality assessment is, in many ways, directly related to the extent to which broad factors such as conscientiousness provide most of the essential predictive power of personality. Several researchers have attempted to address the issue of the appropriate level of analysis when using personality assessment. One view, best reflected by Ones and Visweveran (1996) maintains that broad measures using the big five or similar frameworks provide the best level of analysis for personality assessment. Others (e.g. Schneider, Hough, & Dunnette, 1996) favour the narrower approach using more specific personality factors (see also Robertson & Callinan, 1998). There is no simple solution to this debate. The key question seems to focus on the specific areas of job performance that the personality assessment is designed to predict. Deniz Ones and her collaborators have shown, quite convincingly, that when it comes to the prediction of overall job performance, particularly when data are aggregated over large samples, broad measures such as conscientiousness or integrity produce good validity coefficients. However, other investigators (e.g. Robertson, Baron, Gibbons, MacIver, & Nyfield, 2000) have shown that for particular occupational areas and particular job performance factors, broad measures, such as conscientiousness, do not provide convincing levels of validity.

Most of the research concerning the effects of impression management or intentional or unintentional distortion on the validity of personality assessment has provided results that indicate that in practical terms, there are relatively few problems. There is evidence that applicants do distort their responses when personality assessment is used in a selection procedure (see Hough, 1998). Despite this evidence, the research concerning the impact of matters such as motivational distortion, self-deception and impression management usually shows that there is no detrimental influence on validity (e.g. Barrick & Mount, 1996; Christiansen, Gothin, Johnson, & Rothstein, 1994; Hough, 1998). Some studies have found small effects, but it has also been shown that intentional distortion can be minimized if applicants are warned of the consequences of such distortion. Although distortion by candidates does not appear to create major problems for criterion-related validity, it may still be valuable to include 'social desirability' scales in personality instruments. This is currently common practice and does provide a safeguard against some forms of impression management by candidates.

Interviews

As always, there has been considerable research into interviews as a selection method. Salgado's (1999) review gives an excellent account of work to date. Probably the most comprehensive review of interviews was conducted by McDaniel, Whetzel, Schmitt, and Maurer (1994). A more recent review of interviews has been provided by Moscoso (2000).

Predictive validity and structure of interviews. Probably the most consistent finding is that interviews are improved by using a structure. Typical corrected validity coefficients, quoted by Salgado, are 0.56 for highly structured interviews and 0.20 for interviews with very little predetermined structure. The two main ways of structuring interviews are situational interviewing and behaviour description interviewing. It would seem that situational interviews obtain higher validities than behaviour description interviews (0.50 vs. 0.39). Other notable findings from Salgado's review are that past-orientated questions have a higher validity than future-orientated questions (0.51 vs. 0.39) and that the concurrent validity of interviews is rather higher than the predictive validity.

Construct validity of interviews. Unlike cognitive ability or personality tests, interviews do not focus on specific constructs—they are designed to assess many different candidate attributes. Recent work has focused upon the construct validity of interviews and determining what interviews actually measure. Highly structured and job-related interviews could be measuring cognitive factors such as cognitive ability (Huffcutt, Roth, & McDaniel, 1996; Hunter & Hirsch, 1987), tacit knowledge (Harris, 1998) or job knowledge, while unstructured interviews may be measuring social skills and aspects of personality. Schmidt and Hunter (1998) and Schmidt and Rader (1999) consider that interviews measure a *mélange* of experience, cognitive ability, specific abilities and aspects of personality such as conscientiousness.

Table 3 presents a range of correlations gathered from a number of sources. They are arranged primarily by the characteristic measured and the correlations, where available, with interview performance.

Table 3 cannot lead to incontrovertible conclusions because it does not comprehensively cover all relevant studies. Furthermore, the correlations presented may not be totally comparable; some may be based on small samples and other on larger ones; some correlations may be subject to restriction of range, whilst others may have been corrected. Nevertheless, some tentative conclusions may be drawn. Such conclusions fly in the face of the notion that interviews measure cognitive ability plus conscientiousness. The data in the table suggest that interviews are primarily measuring social skills, experience and job knowledge. General mental ability has only a moderate correlation with interview performance, and the contribution of conscientiousness seems to be quite small. Extroversion and emotional stability would seem to make small, but notable, contributions. Agreeableness and openness to experience also seem to make only a small contribution.

New interview methods. Most research on interviews has focused upon the traditional, unstructured interview and the various forms of structured interview such as

situational interviews and behaviour patterned descriptive interviews. Other forms of interviews are possible, especially with technological advances. Schuler (1989) developed a multimodal interview that is divided into four parts: self-presentations, vocational questions, biographical questions and situational questions. A study with 306 subjects suggests that self-presentation and situational questions were highly correlated with social skills. Silvester, Anderson, Haddleton, Cunningham-Snell, and Gibb (2000) compared face-to-face interviews with the telephone interviews of 70 applicants to a multinational oil corporation. Applicants received lower ratings for telephone interviews. However, those who were interviewed by telephone and then face to face improved their ratings more than candidates who were interviewed face to face and then by telephone. This result probably arose because telephone interviewers focused only upon the verbal content of replies, but face-to-face interviewers added credit for other aspects such as non-verbal behaviour.

Assessment centres

The criterion-related validity of assessment centres has been established for some time. In their reviews, both Hough and Oswald (2000) and Salgado (1999) noted the generally good evidence for criterion-related validity and also the indications that assessment centres create a low adverse impact (Baron & Janman, 1996). Although the criterion-related validity for assessment centres is well established, there has been significant concern concerning the constructs that are measured by assessment centres. Repeated factor analytic studies have indicated that the key factors that emerge from an analysis of assessment centre data are related to exercises rather than the dimensions or psychological constructs that are being assessed. Hough and Oswald (2000) noted several features that might be used to improve the psychometric quality of assessment centre ratings. These are: (a) having only a few conceptually distinct constructs; (b) using concrete job-related construct definitions; (c) using frame of reference assessor training with evaluative standards; (d) using cross-exercise assessment; and (e) using several psychology-trained assessors.

Scholz and Schuler (1993, cited in Salgado, 1999) conducted a meta-analysis of assessment centre data attempting to explore the key constructs that are measured in the overall assessment rating. They found that the overall assessment rating was highly correlated with general intelligence (0.43), achievement motivation (0.4), social competence (0.41), self-confidence (0.32) and dominance (0.30). These results suggest that the primary construct measured within assessment centres relates to general mental ability (see also Goldstein, Yusko, Braverman, Smith, & Chung, 1998). Taken together, these findings raise two key questions about the role of assessment centres within personnel selection and assessment. The first question concerns the extent to which assessment centres provide utility in the personnel selection process. They are clearly an expensive resource and require large numbers of assessors and extensive updating and monitoring of exercises. If the predictive value obtained from assessment centres could be equally well obtained from cheaper methods, such as psychometric testing,

Table 3. Correlates of interviews

Study	Type of interview	Characteristic	<i>r</i>
Sue-Chan, Latham, Evans, and Rotman (1997)	Situational	Self-efficacy	.56
Sue-Chan, Latham, Evans, and Rotman (1997)	BDI	Self-efficacy	.55
Cook, Vance, and Spector (1998)		Locus of control	
Schuler and Funke (1989)	Multi-modal	Social skills	.60
Salgado and Moscoso (2000)	Behaviour SI	Social skills	.54
Salgado and Moscoso (2000)	Conventional SI	Social skills	.38
Hunter and Hirsch (1987)	Unstructured	Social skills	
Huffcutt, Roth, and McDaniel (1996)		Social skills	
Salgado and Moscoso (2000)	Behaviour SI	Experience	.54
Conway and Peneno (1999)	BDI	Experience	.43
Conway and Peneno (1999)	Situational	Experience	.29
Salgado and Moscoso (2000)	Conventional SI	Experience	.26
Schmidt and Rader (1999)		Experience	
Schmidt and Hunter (1998)		Experience	
Salgado and Moscoso (2000)	Behaviour SI	Job knowledge	.50
Burroughs and White (1996)	BDI	Job knowledge	.39
Maurer, Solamon, and Troxtel (1998)	Situational	Job knowledge	.34
US Office of Personnel Management (1987)	BDI	Job knowledge	.23
Harris (1998)	Structured	Job knowledge	
Harris (1998)	Structured	Tacit knowledge	
Huffcutt, Roth, and McDaniel (1999)	Unstructured	Fundamental ability	.50
Salgado and Moscoso (2000)	Conventional SI	General mental ability	.43
Huffcutt, Roth, and McDaniel (1999)	Structured	General mental ability	.35
Salgado and Moscoso (2000)	Behaviour SI	General mental ability	.26
Schuler, Moser, Diemand, and Funke (1995)	Multi-modal	General mental ability	.21
Harris (1998)	Structured	Abilities and skills	
Hunter and Hirsch (1987)	Structured	General mental ability	
Schmidt and Rader (1999)		Ability (specific to job)	
Schmidt and Rader (1999)		General mental ability	
Huffcutt, Roth, and McDaniel (1999)		General mental ability	
Huffcutt, Roth, and McDaniel (1999)		General mental ability	
Sue-Chan, Latham, Evans, and Rotman (1997)		General mental ability	
Schmidt and Hunter (1998)		General mental ability	
Salgado and Moscoso (2000)	Conventional SI	Extraversion	.56
Salgado and Moscoso (2000)	Behaviour SI	Extraversion	.15
Schuler (1989)	Multi-modal	Extraversion	
Salgado and Moscoso (2000)	Conventional SI	Emotional stability	.54
Salgado and Moscoso (2000)	Behaviour SI	Emotional stability	.09
Conway and Peneno (1999)	General question	Emotional stability	

(Continued)

Table 3. Continued

Study	Type of interview	Characteristic	<i>r</i>
Schuler (1989)	Multi-modal	Emotional stability	
Cook, Vance, and Spector (1998)		Emotional stability	
Salgado and Moscoso (2000)	Conventional SI	Agreeableness	.21
Salgado and Moscoso (2000)	Behaviour SI	Agreeableness	.20
Conway and Peneno (1989)	General question	Agreeableness	.17
Salgado and Moscoso (2000)	Conventional SI	Conscientiousness	.25
Salgado and Moscoso (2000)	Behaviour SI	Conscientiousness	.13
Schmidt and Rader (1999)		Conscientiousness	
Schmidt and Hunter (1998)		Conscientiousness	
Salgado and Moscoso (2000)	Conventional SI	Open to experience	.26
Salgado and Moscoso (2000)	Behaviour SI	Open to experience	.04
Salgado and Moscoso (2000)	Conventional SI	Grade Point Average	.15
Salgado and Moscoso (2000)	Behaviour SI	Grade Point Average	.14
Caldwell and Burger (1998)		Personality in general	
Huffcutt, Roth, and McDaniel (1999)		Personality in general	
Schuler (1989)	Multi-modal	Achievement motivation	
Cook, Vance, and Spector (1998)		Achievement motivation	
Harris (1998)	Structured	Achievement motivation	

BDI=behaviour description interviews. Behavioural SI=situational interviews similar to those developed by Janz. Conventional SI=situational interviews similar to those developed by Latham and Saari. Multimodal=interviews developed by Schuler—see section, 'New interview methods'.

the cost-effectiveness of assessment centres in the personnel selection and assessment process should be questioned. The second concern relates to the extent to which the information provided from assessment centres can be used to indicate strengths and weaknesses in candidates and can provide a basis for further development. Concern over the construct validity of the dimensions assessed in assessment centres raises questions over the validity and reliability of the assessment of specific competencies, derived from assessment-centre scores.

Biodata

Although biodata are used far less frequently than other selection methods such as the interview and psychometric tests, they have attracted considerable research attention. The attention may have been prompted by Salgado's (1999) conclusions that biodata have substantial and generalizable criterion validity and that construct validity is well established. Bliesener's (1996) authoritative meta-analysis suggested that the validity of biodata scales was 0.30. However, several factors appeared to moderate this finding. Concurrent validity studies yielded a higher figure of 0.35. The type of criterion used in a study appeared to have a significant effect: studies

using training criteria obtained validities of 0.36, and studies using objective criteria obtained validities of 0.53. Interestingly, the validity of biodata for females is higher than the validity for males (0.51 and 0.27, respectively). More recently, Mount, Witt, and Barrick (2000) demonstrated that empirically keyed biodata scales had incremental validity over a combination of tests of general mental ability and personality.

Biodata have been applied to a wide range of occupations such as clerical jobs in the private sector (Mount *et al.*, 2000), accountants (Harvey-Cook & Taffler, 2000), mechanical equipment distributors (Stokes & Searcy, 1999), hotel staff (Allworth & Hesketh, 1999), civil servants (West & Karas, 1999), managers (Wilkinson, 1997) and naval ratings (Strickler & Rock, 1998). An attempt has even been made to use biodata to predict people's vocational interests (Wilkinson, 1997) and ability to deal with other people from a wide range of backgrounds and ethnic groups (Douthitt, Eby, & Simon, 1999). Research has not placed any emphasis on whether items should be 'hard' biodata items, where a candidate's response can, in principle, be subject to external verification, or whether items should be 'soft' and rely upon a subjective response. Many studies used 'soft' items that resemble questions found in tests of personality. Indeed, it is suspected that if a 'blind' trial were to be conducted, most people would be unable to differentiate many biodata questionnaires from personality questionnaires.

Typically, biodata questionnaires were designed to measure success in a job and yielded one score of overall suitability. Examples of this still exist and can be seen in a paper by Harvey-Cook and Taffler (2000), who employed biodata to predict the success of trainee accountants in professional exams. However, the majority of studies now use biodata to produce scores on dimensions that can then be combined to make predictions. Scales that have been constructed range from: emotional stability (Strickler & Rock, 1998); family and social orientation—work orientation (Carlson, Scullen, Schmidt, Rothstein, & Erwin, 1999), to money management and interest in home repairs (Stokes & Searcy, 1999). These examples serve to show the breadth of the dimensions measured by biodata. It is tempting to believe that research could distil a smaller and more universal set of dimensions.

Publications concerning biodata raise two main issues: the method of 'keying' used in the construction of biodata forms, and the accuracy and generalizability of biodata.

Construction of biodata scales (keying). Archetypal biodata scales were constructed on the basis of *empirical* weights. The main alternative to the empirical method was the *rational* method, where a group of experts would assemble a number of items that they believed would be relevant to success in the job.

The main methods used by the papers included in this review lie between the rational method and the empirical method. Typically, experts assemble a group of items for each trait to be measured. A draft questionnaire is completed by a large sample. The items for an individual trait are then factor-analysed, and any items that do not have a significant loading on the main factor are eliminated. The procedure is repeated for each of the other traits (see Allworth & Hesketh, 1999; Karas & West, 1999). This procedure ensures that scales are unidimensional, but it

is less than ideal because, for example, the items eliminated from each analysis might, in sum, constitute an additional unsuspected factor. An ideal method would be to enter all questions on a biodata form in one large factor analysis. However, this is rarely possible. A biodata form might contain 150 items, and reliable analysis would require a sample of at least 450. Stokes and Searcy (1999) constructed scales using all three methods and compared their accuracy in predicting overall performance and sales performance in two samples of mechanical equipment distributors. They found that the rational scales were as predictive as the empirical scales.

The generalizability of biodata is an important issue. If they are not generalizable, a form needs to be re-standardized in each situation where it is used. The continual re-standardization involves considerable extra effort and resources that would deter many organizations. Initial concerns about the lack of generalizability (e.g. Dreher & Sackett, 1983) were allayed to an extent by Rothstein, Schmidt, Erwin, Owens, and Sparks (1990), who found that a biodata form for supervisors had a validity of about 0.32 across organizations and groups. Carlson *et al.* (1999) provided further reassurance. They developed a biodata form for managers which focused upon five factors. The form was developed in a single organization, and performance ratings were used as criteria. A validity study with a large sample of managers in one organization (Carlson *et al.*, 1999) achieved an observed validity of 0.52. Subsequently, data were accumulated for 7334 managers from 24 varied organizations. The level of progression within the organization was used as the criterion. Meta-analysis produced a mean observed validity across organizations of 0.48. This clearly demonstrated that a biodata key, created in a single organization generalized across organizations and industries (see also Rothstein *et al.*, 1990).

Résumés, CV's and application forms

After interviews, résumés, CVs and application forms collectively constitute the second most frequently used method of selection. Usually, they are the first contact that a candidate makes with a potential employer, and errors at this stage will have a disproportionate effect. It is therefore amazing that they have been so neglected by researchers. The neglect is still more amazing because it has been noted by a string of reviewers from Stephens, Watt, and Hobbs (1979) to Brown and Campion (1994). Bright and Hutton (2000) provided an excellent overview of recent research which seemed to focus upon two themes: the use of competency statements and physical attractiveness.

Competency statements are self-evaluations made by candidates. They cannot be easily verified by a selector in the same way that qualifications or job history can be verified. A typical competency statement might read, 'I am highly motivated with a proven track record in achieving goals and targets'. An investigation by Earl, Bright, and Adams (1998) indicated that the inclusion of competency statements in CVs increased the probability of producing an invitation to an interview. Furthermore, competency statements did most to improve the chances of candidates who were thought to be poor in other respects. Bright and Hutton (2000) looked at this effect in more detail. They chose four good CVs and inserted into them, on a

systematic basis, competency statements. Sometimes, these were related to the job, and sometimes, the competency statements were placed in a letter of application that preceded the CV. It is important to note that the materials in this investigation were drawn from actual CVs submitted in response to an actual vacancy. The CVs were judged by a panel of recruitment consultants, human-resource managers and line managers with expertise in evaluating CVs. Some of the panel were given job descriptions, whilst others were only given the advertisement. The results were clear-cut. The higher the number of competency statements, the higher the evaluation of the CV. It did not matter whether the competency statements were related to the job or if they were placed in the introductory letter.

A similar methodology was adopted by Watkins and Johnston (2000), but they used students to judge CVs. Watkins and Johnston simulated a vacancy for a graduate trainee manager and the CVs of two female applicants, who differed in course grades, work experience and positions of responsibility etc., so that one CV was clearly superior to the other. The CVs included, on a systematic basis either no photograph, a photograph of an attractive candidate or a photograph of an 'average' candidate. The students were asked to imagine they were the recruiting officer for a company, and they were required to evaluate the CV in terms of suitability, probability of an invitation to an interview and the likely starting salary. The results revealed a complex but consistent pattern. The inclusion of a photograph, whether attractive or average, made little difference to good CVs which consistently evoked higher evaluations, more invitations to interviews and higher indications of starting salary. However, the inclusion of an attractive photograph did help the evaluation of an average CV and did increase the reported probability of an interview offer. But, the inclusion of an attractive photograph with an average CV did not significantly alter the estimates of likely starting salary.

Validation of selection procedures

The research literature on personnel selection methods generally focuses on one specific indicator of validity, the criterion-related validity coefficient. This is given prominence above all other indicators of validity. Clearly, in many ways, this emphasis on the extent to which personnel selection procedures can adequately predict work criteria is appropriate. The whole purpose of a personnel selection process is to identify candidates who are most or least suited to the occupational area in question. Although the current and historical focus on criterion-related validity as the major quality standard for personnel selection methods seems appropriate, difficulties arise when considering the interpretation of the evidence, concerning criterion-related validity of personnel selection methods. Meta-analysis has provided a statistical mechanism for giving a clear indication of the criterion-related validity for personnel selection methods. The positive contribution to the research literature of meta-analysis should not be underestimated. It is clear, however, that some misapplications and misinterpretations of meta-analytic results have been unhelpful. Hermelin and Robertson (in press) have provided a comparative analysis of the data concerning meta-analytic results for many of the major personnel selection methods.

In meta-analytic studies, a specific indicator for the validity of a personnel selection method is usually produced. In common with any other correlation coefficient, this point estimate for validity would have a confidence interval around it. The confidence interval indicates the likely lower and upper boundary for this estimate. A simple comparison of the mean validities hides the fact that the lower boundary for the validity of one method may be lower than the lower boundary for the validity of the other, despite the fact that the mean estimates of validity are the other way around. In other words, the key message here is that comparing the point estimates for validity derived from meta-analytic studies without looking at the relevant confidence intervals is inappropriate and may lead to misleading conclusions.

The second point concerns the way in which individual investigators have conducted meta-analytic studies. As Hermelin and Robertson (in press) show, when investigators focus on the same personnel selection method, they do not use a consistent set of correction factors in order to derive a meta-analytic mean validity coefficient. Thus, some studies correct for factors such as range restriction in the predictor, and others do not. Furthermore, for example, when correcting for unreliability, studies differ in the correction factors that are applied, even for the same personnel selection methods. These different practices lead to conclusions that are not comparable.

Finally, when considering the validity evidence for different personnel selection methods, the issue of construct validity needs to be considered. Only two personnel selection methods (mental-ability testing and personality testing) are directly associated with specific constructs. These two selection methods are defined by the constructs that they measure. Other selection methods are not defined in this way. They are defined by the procedures adopted and not the specific constructs measured. For example, the International Personnel Management Association (Task Force on Assessment Center Guidelines, 1989) specifies 10 essential elements that define an assessment centre. These include a variety of criteria such as the fact that the assessment centre must be based on job analysis, that multiple assessment techniques must be used and that multiple assessors must observe the performance of each assessee. In fact, all of the 10 criteria relate to the procedures and structure of the assessment process. None is related to the construct or constructs that are measured by the process. Thus, when comparing validities for assessment centres, structured interviews, cognitive-ability tests and personality tests, we are not comparing similar approaches. Much more needs to be known about the constructs that are measured within specific assessment methods. Without such information, comparative evaluation of validity is almost meaningless.

The meta-analytically derived information on the validity of personnel selection methods is, nevertheless, useful and has provided researchers and practitioners with a clear indication of the validities associated with different methods. Unfortunately, the current technology of meta-analysis and the database on which investigators may draw does not allow for a thorough evaluation of the extent to which selection methods may be combined to provide incremental validity. In general, meta-analytic studies have focused on the validity of one particular method. In practice, of course, personnel selection procedures frequently use many methods. In this

context, the key question for people who are designing personnel selection procedures concerns the extent to which different methods provide unique and non-overlapping information concerning candidates' likely performance. The issue of the incremental validity provided by different methods is something that is being actively explored by personnel selection researchers. Within the last 5 years or so, a number of articles has appeared, attempting to assess the extent to which combinations of methods are useful or, by contrast, provide overlapping and hence redundant information. Studies concerning the incremental validity of interviews, cognitive ability and personality (e.g. Cortina, Goldstein, Payne, Davison, & Gilliland, 2000) biodata and personality (e.g. Mount *et al.*, 2000) provide an indication of the kind of research that is currently being conducted. The article by Cortina *et al.* (2000), for example, used a meta-analytic approach to assess the relationships between cognitive ability, conscientiousness and interviews. These results were then combined with results concerning criterion-related validity from previous meta-analysis to provide estimates of the extent to which each of the individual methods provided unique validity. The results suggested that interview scores helped to predict job performance, beyond the information provided by cognitive ability and personality (conscientiousness). In particular, highly structured interviews were shown to contribute substantially to the prediction of job performance. As more research of this kind is conducted and published, investigators should develop a clearer view of the best ways in which personnel selection methods may be combined to provide optimal selection procedures.

Future prospects

The advances in the theory and practice of selection and assessment in the last 50 years have been enormous. We now know, with some certainty, the accuracy and validity of most methods of selection. We have a much clearer conceptual grasp of fairness and the nature of job criteria. There has also been a significant, but not very fruitful, investigation of selection from the candidates' perspective. Developments in several areas mentioned in this article are likely to be important. It is also possible that future advances will be made in new areas, not so far mentioned in this article. Two new areas of particular interest are: the use of physiological measures and the benchmarking of selection systems.

Physiological measures

Current research in selection and assessment appears to have overlooked advances in the wider realm of psychology, which suggest that physiological measures may be useful as assessment tools. For example, Shafer (1982, 1985) and Shafer and Marcus (1973) investigated several indices derived from EEG records. One index, 'the neural adaptability index' measures the degree to which the amplitude of brainwaves decreases when a stimulus is repeated. It was hypothesized that high-IQ individuals would habituate more quickly and thus conserve neural resources. It was found that neural adaptability had a corrected correlation of 0.82 with scores

from the Weschler intelligence scale—a correlation comparable to that found between two established tests of intelligence. Similarly, Eysenck and Barrett (1985) investigated the complexity of brainwaves (average evoked potentials) engendered by a standard stimulus. They obtained a correlation of 0.83 between the complexity of a person's average evoked potential and full scale scores on the Weschler test. There is some suggestion that certain aspects of personality such as emotional stability and extraversion also have neurophysiological correlates. It is too early to say whether these developments will have any advantage over the measures customarily used in selection. However, it is surprising that their potential has not been investigated in applied settings.

Benchmarking selection systems

Practitioners in the field of selection and assessment often need to compare (benchmark) their systems against the systems used by leading organizations. Practitioners in other fields, such as production managers, can use a number of methods to benchmark their operations against the production operations of leading companies. A production manager can have his or her methods externally audited and obtain a global score and an indication of those facets of his or her selection system that fall below the standards of best practice. In many organizations, the selection function is subject to similar pressures, and it is likely and desirable that methods of auditing selection systems are developed in the near future.

Explanatory models

It seems likely that, now the validity of selection measures has been established to a reasonable degree of satisfaction, attention will turn to explaining why measures are valid (or not valid). A beginning has been made by Hunter *et al.* (2000) and Schmidt and Hunter (1998), who suggest that measures of 'g' are predictive because general intelligence allows people to acquire job knowledge, which in turn has a direct effect upon work performance. Earlier in this article, research on the construct validity of interviews and assessment methods was reviewed. A grasp of the constructs being assessed by specific methods is a prerequisite for understanding the reasons behind criterion-related validities. However, a deeper level of analysis than that delivered by current methods of research and theory is needed.

Meta-analyses are likely to yield high validities when a characteristic is related to work performance in a substantial majority of occupations. Smith (1994) called such characteristics 'universals' and suggested the existence of three such characteristics: 'intelligence', 'vitality' and the proportion of their 'life space' an individual is prepared to devote to his or her work. Measures of a candidate's 'vitality' and the proportion of their life space devoted to work are rarely included in current selection systems—although they may be measured indirectly by interviews and biodata. Of course, the fact that a characteristic has broad, generalizable validity across a wide range of occupations does not mean that it is the most important

factor in any given occupational area, or specific job. This point is well illustrated by the research concerning the big-five personality factor, conscientiousness. In meta-analytic studies, conscientiousness has been shown to be important, with a similar level of validity across many areas (Barrick & Mount, 1991; Salgado, 1998). It is difficult to imagine many jobs in which it is not advantageous for incumbents to be dependable, self-disciplined and likely to meet deadlines. Although such attributes are an asset in many jobs, they may not be the main factors in determining high levels of performance. Depending on the job, other factors may have a much more influential role in determining the extremes (high or low) of performance.

Smith (1994) also identified a second domain of characteristics relevant to specific occupations. Perhaps the most important of these 'occupational characteristics' is job knowledge, which can emerge as a valid predictor for generic meta-analysis because, although job knowledge varies, it is easily recognized and classified. Other variables, e.g. specific personality factors, may be relevant only to certain occupations. For example, extraversion may be relevant to managerial and social occupations, while introversion may be relevant to precision work, such as electronics or 'back-room' research workers. A generic meta-analysis of these characteristics would probably yield a mean validity close to zero (with a wide variance), but a moderator analysis that 'coded' for the type of occupation could yield useful validities.

Smith (1994) identified a third domain concerning the characteristics that help a person relate to a specific work setting or a specific employer. For example, an intelligent (universal) lawyer with a great deal of legal experience (occupational) might succeed in a neighbourhood law centre but might fail in a slick city law firm. The difference in outcome might be in the degree to which the incumbent's values coincide with the values of colleagues and the host organization. Characteristics of this kind might be termed 'relationals' because they determine an individual's fit to specific employers. A model of this kind would help identify the need for new types of measures such as vitality or life space devoted to work. Measures of 'relational' characteristics will be the most difficult to employ because they would require assessment of both the candidate and their organization. Such a model would also call for a more sophisticated meta-analysis. In effect, a model of this kind would reflect a growing concern in individual differences research, which involves attempting to combine the influences of both person and situation factors in models of behaviour causation.

Predictive models

As earlier sections of this article have shown, the construct validity of some selection methods (e.g. interviews, assessment centres and biodata) is not well understood. There are compelling reasons for researchers to explore construct validity issues more extensively. On the scientific front, it is important to understand the reasons for relationships between predictor variables and criteria related to performance (e.g. supervisory ratings, promotions, organizational citizenship) attachment (e.g. turnover, absenteeism, commitment) and well-being (e.g.

job satisfaction). The identification of key predictor constructs assessed by different selection methods is important in understanding the key attributes linked with criteria. In the same vein, criterion variable constructs need more conceptual and empirical clarification. The three major categories suggested above, performance, attachment and well-being, provide a broader set of criterion areas than those used in most contemporary personnel-selection research. These broader areas are important issues for organizations, individuals and psychological researchers. It would be helpful to see more studies emerging which use a number of predictors and broad criteria. Such studies would provide a basis for better understanding of predictor and criterion constructs and the relationships between them. Such studies would also provide good primary data for subsequent meta-analyses.

These subsequent meta-analyses could then provide clarification of predictor-criterion relationships when more than one predictor is used. Although we now have a reasonable grasp of the validity of many selection methods when used as single predictors, we need to know much more about how to use predictors in combination. A beginning has been made here by Schmidt and Hunter (1998), but much more needs to be done. The meta-analytic work is limited by the availability of primary data from studies of the kind mentioned above.

References

- Allworth, H., & Hesketh, B. (1999). Construct-oriented biodata: capturing change-related and contextually relevant future performance. *International Journal of Selection and Assessment*, 7, 97-111.
- Arvey, R. D., & Murphy, K. R. (1998). Performance evaluation in work settings. *Annual Review of Psychology*, 49, 141-168.
- Baron, H., & Janman, K. (1996). Fairness in the assessment centre. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (Vol. 11, pp. 61-113). London: Wiley.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, 44, 1-26.
- Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, 81, 261-272.
- Bernadin, H. J., & Beatty, R. W. (1984). *Performance appraisal: assessing human behaviour at work*. Boston: Kent.
- Bliesener, T. (1996). Methodological moderators in validating biographical data in personnel selection. *Journal of Occupational & Organizational Psychology*, 69, 107-120.
- Bobko, P., Roth, P. L., & Potosky, D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology*, 52, 1-31.
- Borman, W. C., & Brush, D. H. (1993). More progress towards a taxonomy of managerial performance requirements. *Human Performance*, 6, 1-21.
- Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations*. San Francisco, CA: Jossey Bass.
- Borman, W. C., & Motowidlo, S. J. (Eds.) (1997). Organizational citizenship behavior and contextual performance. *Human Performance*, 10, 69-192.
- Bretz, R. T., & Judge, T. A. (1998). Realistic job previews: a test of the adverse self-selection hypothesis. *Journal of Applied Psychology*, 83, 230-337.
- Bright, J. E. H., & Hutton, S. (2000). The impact of competences statements on resumes for short listing decisions. *International Journal of Selection and Assessment*, 8, 41-53.

- Brown, B. K., & Campion, M. A. (1994). Biodata phenomenology: recruiters' perceptions and use of biographical information in resume screening. *Journal of Applied Psychology, 79*, 897-908.
- Burrows, W. A., & White, L. L. (1996). Predicting sales performance. *Journal of Business and Psychology, 11*, 73-83.
- Caldwell, D. F., & Burger, J. M. (1998). Personality characteristics of job applicants success in screening interviews. *Personnel Psychology, 51*, 119-136.
- Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organisational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology*, Vol 1, 687-732. Palo Alto: Consulting Psychologists Press.
- Campbell, J. P. (1994). Alternative models of job performance and the implications for selection and classification. In M. G. Rumsey, C. B. Walker, & J. H. Harris (Eds.), *Personnel selection and classification*. Hillsdale, NJ: Earlbaum.
- Campbell, J. P., McHenry, J. J., & Wise, L. L. (1990). Modelling job performance in a population of jobs. *Personnel Psychology, 43*, 313-333.
- Carlson, K. D., Scullen, S. E., Schmidt, F. L., Rothstein, H., & Erwin, F. (1999). Generalisable biographical data validity can be achieved without multi-organizational development and keying. *Personnel Psychology, 52*, 731-755.
- Carroll, J. B. (1993). *Human cognitive abilities: a survey of factor-analytic studies*. Cambridge: Cambridge University Press.
- Chan, D., Schmitt, N., Jennings, D., Cause, C. S., & Delbridge, K. (1998). Applicant perceptions of test fairness: integrating justice and self-serving bias perspectives. *International Journal of Selection and Assessment, 6*, 232-239.
- Christiansen, N. D., Gohin, R. D., Johnson, N. G., & Rothstein, M. G. (1994). Correcting the 16PF for faking: Effects of criterion-related validity and individual hiring decisions. *Personnel Psychology, 47*, 847-860.
- Cleary, T. A. (1968). Test bias: prediction of grades of negro and white students in integrated colleges. *Journal of Educational Measurement, 5*, 115-124.
- Coleman, V., & Borman, W. (in press). Investigating the underlying structure of the citizenship performance domain. *Human Resources Research Rev.*
- Conway, J. M. (1999). Distinguishing contextual performance from task performance for managerial jobs. *Journal of Applied Psychology, 84*, 3-13.
- Conway, J. M., & Peneno, G. M. (1999). Compare structured interview question types: construct validity and applicant reactions. *Journal of Business and Psychology, 13*, 485-505.
- Cook, K. W., Vance, C. A., & Spector, P. E. (2000). The relation of candidate personality with selection-interview outcomes. *Journal of Applied Social Psychology, 30*, 4, 867-885.
- Cortina, J. M., Goldstein, N. B., Payne, S. C., Davison, H. K., & Gilliland, S. W. (2000). The incremental validity of interview scores over and above cognitive ability and conscientiousness scores. *Personnel Psychology, 53*, 325-351.
- Douthitt, S. S., Eby, L. T., & Simon, S. A. (1999). Diversity of life experiences: the development and validation of graphical measures of receptiveness to dissimilar others. *International Journal of Selection and Assessment, 7*, 112-125.
- Dreher, G. F., & Sackett, P. R. (1983). *Perspectives on staffing and selection*. Homewood, IL: Irwin.
- Earl, J., Bright, J. E. H. R., & Adams, A. (1998). 'In my opinion': what gets graduates resumes short-listed? *Australian Journal of Career Development, 7*, 15-19.
- Elkins, T. J., & Phillips, J. S. (2000). Job context, selection decision outcome, and perceived fairness of selection tests: Biodata as an illustrative case. *Journal of Applied Psychology, 85*, 479-484.
- Eysenck, H. J., & Barrett, P. (1985). Psychophysiology and the measurement of intelligence. In C. R. Reynolds & P. C. Wilson (Eds.), *Methodological and statistical advances in the study of individual differences* (pp. 1-49). New York: Plenum Press.
- Frei, R. L., & McDaniel, M. A. (1997). Validity of customer service measures in personnel selection: A review of criterion and construct evidence. *Human Performance, 11*, 1-27.
- Gilliland, S. W. (1993). The perceived fairness of selection systems: an organizational justice perspective. *Academy of Management Review, 18*, 694-734.

- Goldstein, H. W., Yusko, K. P., Braverman, E. P., Smith, D. B., & Chung, B. (1998). The role of cognitive ability in the subgroup differences and incremental validity of assessment centre exercises. *Personnel Psychology, 51*, 357-374.
- Goleman, D. (1996). *Emotional intelligence*. London: Bloomsbury.
- Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology, 18*, 135-164.
- Harris, M. M. (1998). The structured interview: what constructs are being measured? In R. Eder & M. Harris (Eds.), *The employment interview: theory, research and practice*. Thousand Oaks, CA: Sage Publications.
- Harvey-Cook, J. E., & Taffler, R. J. (2000). Biodata in professional entry-level selection: statistical scoring of common format applications. *Journal of Occupational & Organizational Psychology, 73*, 103-118.
- Hermelin, E., & Robertson, I. T. (in press). A critique and standardization of meta-analytic validity coefficients in personnel selection. *Journal of Occupational and Organizational Psychology*.
- Hogan, J., & Rybicki, S. L. (1998). *Performance improvement characteristics job analysis*. Tulsa, OK: Hogan Assessment Systems.
- Hough, L. M. (1998). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance, 11*, 209-244.
- Hough, L. M., & Oswald, F. L. (2000). Personnel selection: looking toward the future—remembering the past. *Annual Review of Psychology, 51*, 631-664.
- Huang, T. (2000). Human resource management practices at subsidiaries of multinational corporations and local firms in Taiwan. *International Journal of Selection and Assessment, 8*, 22-33.
- Huffcutt, A. I., Roth, P. L., & McDaniel, M. A. (1996). A meta-analytic investigation of cognitive ability in employment interview evaluations: moderating characteristics and implications for incremental validity. *Journal of Applied Psychology, 81*, 459-473.
- Hunt, S. T. (1996). Generic work behaviour: an investigation into the dimensions of entry-level hourly job performance. *Personnel Psychology, 49*, 51-83.
- Hunter, J. E., & Hirsch, H. R. (1987). Applications of meta-analysis. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology, 2*, 321-357.
- Hunter, J. R., & Schmidt, F. L. (1990). *Methods of meta-analysis: correcting error and bias in research findings*. Newbury Park, CA: Sage.
- Hunter, J. E., Schmidt, F. L., Rauchenberger, J. M. R., & Jayne, M. E. A. (2000). Intelligence, motivation and job performance. In C. L. Cooper & E. A. Locke (Eds.), *Industrial and organizational psychology: linking theory with practice*. Oxford: Blackwell.
- Karas, M., & West, J. (1999). Construct-oriented biodata development for a selection to a differentiated performance domain. *International Journal of Selection and Assessment, 7*, 86-96.
- 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.
- Landis, R. S., Fogli, L., & Goldberg, E. (1998). Future-oriented job analysis: a description of the process and its organizational implications. *International Journal of Selection and Assessment, 6*, 3, 192-197.
- McDaniel, M. A., Whetzel, D. L., Schmidt, M. L., & Maurer, S. (1994). The validity of employment interview: a comprehensive review and meta-analysis. *Journal of Applied Psychology, 79*, 599-617.
- Mathews, B. P., & Redman, T. (1998). Managerial recruitment advertisements—just how market oriented are they? *International Journal of Selection and Assessment, 6*, 240-248.
- Maurer, T., Solaman, J., & Troxstel, D. (1998). Relationship of coaching with performance in situational employment interviews. *Journal of Applied Psychology, 83*, 128-136.
- Moscoso, S. (2000). Selection interview: a review of validity evidence, adverse impact and applicant reactions. *International Journal of Selection and Assessment, 8*, 237-247.
- Mount, M. K., Witt, L. A., & Barrick, M. R. (2000). Incremental validity of empirically keyed Biodata scales over GMA and the five factor personality constructs. *Personnel Psychology, 53*, 299-323.
- Murphy, K. R. (2000). Impact of assessments of validity generalisation and situational specificity on the science and practice of personnel selection. *International Journal of Selection and Assessment, 8*, 194-215.

- Newell, S. (2000). Selection and assessment in the Knowledge Era. *International Journal of Selection and Assessment*, 8, 1–6.
- Olea, M. M., & Ree, M. J. (1994). Predicting pilot and navigator criteria: not much more than g. *Journal of Applied Psychology*, 79, 845–851.
- Ones, D. S., & Visweveran, C. (1996). *What do pre-employment customer service scales measure? Explorations in construct validity and implications for personnel selection*. Presented at Annual Meeting Society Industrial and Organizational Psychology, San Diego, CA.
- Ones, D. S., & Visweveran, C. (1998). Gender, age, and race differences on overt integrity tests: results across four large-scale job applicant data sets. *Journal of Applied Psychology*, 83, 35–42.
- Ones, D. S., Visweveran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: findings and implications of personnel selection and theories of job performance. *Journal of Applied Psychology*, 78, 679–703.
- Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P. R., & Fleishman, E. A. (1999). *An occupational information system for the 21st century: the development of O*NET*. Washington, DC: American Psychological Association.
- Ployhart, R. E., & Ryan, A. M. (1998). Applicants' reactions to the fairness of selection procedures: the effects of positive rule violation and time of measurement. *Journal of Applied Psychology*, 83, 3–16.
- Porter, M. E. (1985). *Competitive advantage*. New York: Free Press.
- Purcell, K., & Purcell, J. (1998). In-sourcing, out-sourcing and the growth of contingent labour as evidence of flexible employment strategies. *European Journal of Work and Organizational Psychology*, 7, 39–59.
- Raymark, P. H., Schmit, M. J., & Guion, R. M. (1997). Identifying potentially useful personality constructs for employees' selection. *Personnel Psychology*, 50, 723–736.
- Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance: not much more than g. *Journal of Applied Psychology*, 79, 518–524.
- Robertson, I. T., Baron, H., Gibbons, P., MacIver, R., & Nyfield, G. (2000). Conscientiousness and managerial performance. *Journal of Occupational and Organizational Psychology*, 73, 171–180.
- Robertson, I. T., & Callinan, M. (1998). Personality and work behaviour. *European Journal of Work and Organizational Psychology*, 7, 321–340.
- Robertson, I. T., & Kinder, A. (1993). Personality and job competences: the criterion-related validity of some personality variables. *Journal of Occupational and Organizational Psychology*, 66, 225–244.
- Rolland, J. P., & Mogenet, J. L. (1994). *Manuel d'application. Systeme D5D d'aide à l'évaluation des personnes*. Paris: Les Editions du Centre de Psychologie Appliquée.
- Rothstein, H. R., Schmidt, F. L., Erwin, F. W., Owens, W. S., & Sparks, P. P. (1990). Biographical data in employment selection: can validities be made generalisable? *Journal of Applied Psychology*, 75, 175–184.
- Rynes, S. L., & Connelly, M. L. (1993). Applicant reactions to alternative selection procedures. *Journal of Business and Psychology*, 7, 261–277.
- 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 & I. T. Robertson (Eds.), *International Review of Industrial & Organizational Psychology*. New York: Wiley.
- Salgado, J. F., & Moscoso, S. (2000). *Construct validity of employment interview*. Under review. Quoted by Moscoso.
- Sanchez, J. L. (2000). Adapting work analysis to a fast-paced and electronic business world. *International Journal of Selection and Assessment*, 8, 207–215.
- Sanchez, J. L., & Fraser, S. L. (1992). On the choice of scales for task analysis. *Journal of Applied Psychology*, 77, 545–553.
- Sandberg, J. (2000). Understanding human competence at work: an interpretative approach. *Academy of Management Journal*, 43, 9–25.
- Schmidt, F. L., & Hunter, J. F. (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.
- Schmidt, F. L., & Rader, M. (1999). Exploring the boundary conditions for interview validity: meta-analytic validity findings for a new interview type. *Personnel Psychology*, 52, 445–464.

- Schmitt, N., & Chan, D. (1998). *Personnel selection: a theoretical approach*. Thousand Oaks, CA: Sage Publications.
- Schmitt, N., & Gilliland, S. W. (1992). Beyond differential prediction: fairness in selection. In D. Saunders (Ed.), *New approaches to employee management*. Greenwich, CT: JAI Press.
- Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings, D. (1997). Adverse impact and predictive efficiency of various predictor combinations. *Journal of Applied Psychology, 82*, 719-730.
- Schneider, R. J., Hough, L. M., & Dunnette, M. D. (1996). Broad-sided by broad traits, or how to sink science in five dimensions or less. *Journal of Organizational Behavior, 17*, 639-655.
- Scholarios, D., & Lockyer, C. (1999). Recruiting and selecting professionals: contexts, qualities and methods. *International Journal of Selection and Assessment, 7*, 142-169.
- Scholz, G., & Schuler, H. (1993). Das nomologische Netzwerk des assessment centers: eine metaanalyse [The nomological network of the assessment centre: a meta-analysis]. *Zeitschrift für Arbeits- und Organisationspsychologie, 37*, 73-85.
- Schuler, H. (1989). Construct validity of a multi-model employment interview. In B. J. Fallon, H. P. Pfister, & J. Brebner (Eds.), *Advances in Industrial and Organizational Psychology*. New York: North Holland.
- Schuler, H., & Funke, U. (1989). The interview as a multimodal procedure. In E. W. Eider and G. R. Ferris (eds.) *The employment interview: Theory research and practice*. Newbury Park, California: Sage.
- Schuler, H., Moser, K., Diamond, A., & Funke, U. (1995). Validität eines Einstellungsinterviews zur Prognose des Ausbildungserfolgs. *Zeitschrift für Pädagogische Psychologie, 9*, 45-54.
- Shackleton, V., & Newell, S. (1997). International assessment and selection. In N. Anderson & P. Herriot (Eds.), *International handbook of selection and assessment*. Chichester, UK: Wiley.
- Shafer, E. W. P. (1982). Neural adaptability: a biological determinant of behavioural intelligence. *International Journal of Neuroscience, 17*, 183-191.
- Shafer, E. W. P. (1985). Neural adaptability: a biological determinant of g factor intelligence. *Behavioural and Brain Sciences, 8*, 240-241.
- Shafer, E. W. P., & Marcus, M. M. (1973). Self-stimulation alters human sensory brain responses. *Science, 181*, 175-177.
- Silvester, J., Anderson, N., Haddleton, E., Cunningham-Snell, N., & Gibb, A. (2000). A cross-modal comparison of telephoned and face-to-face selection interviews in graduate recruitment. *International Journal of Selection and Assessment, 8*, 16-21.
- Smith, M. (1994). A theory of the validity of predictors in selection. *Journal of Occupational and Organizational Psychology, 67*, 13-31.
- Smith, M., & Robertson, I. T. (1993). *Systematic personnel selection*. London: Macmillan.
- Stephens, D. B., Watt, J. T., & Jobbs, W. S. (1979). Getting through the resume preparation maze: some empirically based guidelines for resume format. *The Vocational Guidance Quarterly, 27*, 25-34.
- Sternberg, R. J., & Wagner, R. K. (1986). *Practical intelligence*. Cambridge: Cambridge University Press.
- Sternberg, R. J., & Wagner, R. K. (1995). Testing common sense. *American Psychologist, 50*, 912-927.
- Stokes, G. S., & Searcy, C. A. (1999). Specification of scales in Biodata from development: rational versus empirical and global versus specific. *International Journal of Selection and Assessment, 7*, 72-96.
- Strickler, L. J., & Rock, D. (1998). Assessing leadership potential with a biographical measure of personality traits. *International Journal of Selection and Assessment, 6*, 164-184.
- Sue-Chan, C., Latham, M. G., Evans, M. G., & Rotman, J. I. (1997). *The construct validity of the situation and patterned behaviour description interviews: cognitive ability, tacit knowledge and self-efficacy as correlates*. Unpublished manuscript, Faculty Management, University of Toronto, Canada.
- Task Force on Assessment Center Guidelines (1989). Guidelines and ethical considerations for assessment centre operations. *Public Personnel Management, 18*, 457-470.
- 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.
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: a meta-analytic review. *Personnel Psychology, 44*, 703-742.
- Thorsteinson, T. J., & Ryan, A. M. (1997). The effect of selection ratio on perceptions of the fairness of a selection test battery. *International Journal of Selection and Assessment, 5*, 159-168.
- Tonidandel, S., & Quinones, M. A. (2000). Psychological reactions to adaptive testing. *International Journal of Selection and Assessment, 8*, 7-15.

- Trepesta, D. E., & Rozell, E. J. (1993). The relationship of staffing practices to organizational level measures of performance. *Personnel Psychology, 46*, 27–48.
- US Office of Personnel Management (1987). *The structured interview*. Washington, DC: Office of Examination Development, Division of Alternative Examining Procedures.
- Viswesvaran, C. (1993). *Modeling job performance: Is there a general factor?* Unpublished PhD. Iowa City: University of Iowa.
- Viswesvaran, C., & Ones, D. S. (2000). Perspectives of models of job performance. *International Journal of Selection and Assessment, 8*, 216–225.
- Watkins, L. M., & Johnston, L. (2000). Screening of job applicants: the impact of physical attractiveness and application quality. *International Journal of Selection and Assessment, 8*, 77–84.
- West, J., & Karas, M. (1999). Biodata: meeting clients' needs for a better way of recruiting entry-level staff. *International Journal of Selection and Assessment, 7*, 126–131.
- Westoby, J. B., & Smith, J. M. (2000). *The 16PI3 job spec*. Windsor, UK: Assessment and Selection in Employment (ASE).
- Wilkinson, L. J. (1997). Generalisable Biodata? An application to the vocational interests of managers. *Journal of Occupational & Organizational Psychology, 70*, 49–60.
- Wright, P. M., & McMahan, G. C. (1992). Theoretical perspectives for strategic human resource management. *Journal of Management, 18*, 295–320.