

Old and new selection methods

We've always done it this way

WHY SELECTION MATTERS

Clark Hull is better known, to psychologists at least, as an animal learning theorist, but very early in his career he wrote a book on aptitude testing (Hull, 1928), and described ratios of output of best to worst performers in a variety of occupations. Hull was the first psychologist to ask how much workers differ in productivity, and he discovered the principle that should be written in letters of fire on every HR manager's office wall: *the best is twice as good as the worst*.

Human resource managers sometimes find they have difficulty convincing colleagues that HR departments also make a major contribution to the organization's success. Because HR departments are neither making things, nor selling things, some colleagues think they do not add any value to the organization. This represents a very narrow approach to how organizations work, which overlooks the fact that an organization's most important asset is its staff. Psychologists have devised techniques for showing how finding and keeping the right staff adds value to the organization. Rational Estimate technique (described in detail in Chapter 14) estimates how much workers doing the same job vary in the value of their contribution. One 'rule of thumb' this research generated states that *The value of a good employee minus the value of a poor employee is roughly equal to the salary paid for the job*. If the salary for the job in question is £50,000, then a good employee, in the top 15%, is worth £50,000 more each year than one in the bottom 15%. Differences in value of the order of £50,000 per employee mount up across an organization. Hunter and Hunter (1984) generated a couple of examples, for the public sector in the USA.

- A small employer, the Philadelphia police force (5,000 employees), could save \$18 million a year by using psychological tests to select the best.
- A large employer, the US Federal Government (4 million employees), could save \$16 billion a year. Or, to reverse the perspective, the US Federal Government was losing \$16 billion a year, at 1980s prices, by not using tests.

Some critics see a flaw in such calculations. Every company in the country cannot employ the *best*, for example, computer programmers; someone has to employ *the rest*. Good selection cannot increase national productivity, only the productivity of employers that use good selection methods to grab more than their fair share of talent. At present, employers are – largely – free to do precisely that. The rest of this book explains *how*.

RECRUITMENT

Traditional methods

Figure 1.1 summarizes the successive stages of recruiting and selecting an academic for a British university. The *advertisement* attracts applicants, who complete and return an *application form*. Some applicants' *references* are taken up; the rest are excluded from further consideration. Applicants (As) with satisfactory references are shortlisted, and invited for *interview*, after which the post is filled. The employer tries to attract as many As as possible, then pass them through a series of filters, until the number of surviving As equals the number of vacancies.

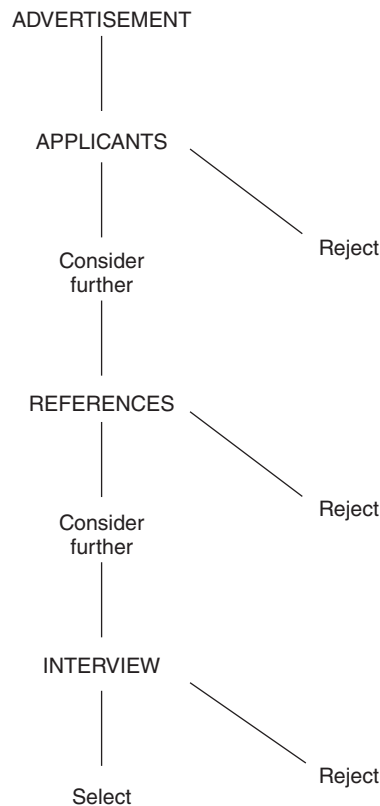


Figure 1.1 Successive stages in selecting academic staff in a British university.

Recruitment sources

There are many ways employers can try to attract applicants: advertisements, agencies – public or private, word of mouth, ‘walk-ins’ (people who come in and ask if there are any vacancies), job fairs, and the Internet. Employers should analyse recruiting sources carefully to determine which find good employees who stay with them. Employers also need to check whether their recruitment methods are finding a representative applicant pool, in terms of gender, ethnicity, and disability. Newman and Lyon (2009) investigate targeted recruiting, through the wording of advertisements for job. They suggest that saying the organization is ‘results oriented’ will tend to attract more As high in conscientiousness, and saying the organization is ‘innovative’ will attract more As high in mental ability. Later chapters will describe research showing As high in conscientiousness and mental ability tend to make better employees. Newman and Lyon suggest the right advertisement can attract such applicants, both overall and from minorities, so meeting the twin aims of many employers: good employees *and* a representative workforce.

Realistic job previews

Any organization can paint a rosy picture of what is really a boring and unpleasant job because they fear no one would apply otherwise. In the USA realistic job previews (RJPs) are widely used to tell applicants what being, for example, a call-centre worker is really like: fast-paced, closely supervised, routine to the point of being boring. Earnest, Allen and Landis’s (2011) analysis confirms the results of several earlier reviews that there is a very modest link with reduced turnover, suggesting RJPs may be worth using, given that RJPs cost employers very little whereas turnover costs them a lot. Earnest *et al.* suggest RJPs work by making As see the employer as more honest.

Informal recruitment

Applicants are sometimes recruited by word of mouth, usually through existing employees. Besides being cheaper, the grapevine finds employees who stay longer (low *turnover*) possibly because they have a clearer idea what the job really involves. Zottoli and Wanous (2000) report informal recruits on average do slightly better work; the difference is small ($d = 0.08$; page 31) but is achieved very cheaply. However, fair employment agencies, for example the (British) Equality and Human Rights Commission, generally frown on informal recruitment; they argue recruiting an all-white workforce’s friends is unfair because it tends to perpetuate an all-white workforce. Weller *et al.* (2009) report data from the German Socio-Economic Panel Study, nearly 3,000 people, representative of the whole German working population, tracked over five years. Weller *et al.* confirm that informal recruitment results in lower turnover: more employees recruited through agencies or advertisement leave in the first two years, and

leave sooner, departures peaking at 9 months compared with 17 months for employees recruited through informal contacts.

New technology and recruitment

Advertising, making applications, sifting applications and even assessment can now be carried out electronically, which can make the whole process far quicker. People even talk of making 'same-day offers'. More and more jobs are advertised on the Internet, through the employer's own website or through numerous recruitment sites. People seeking jobs can post their details on websites for potential employers to evaluate, which gives the job seeker an opportunity that did not exist before. Most employers now use electronic application systems, eliminating the conventional paper application form. Internet recruitment can greatly increase the number of As, which is good for the employer if it broadens the field of high-calibre As, but it does also create work sorting through a mountain of applications.

APPLICATION SIFTING

The role of the application form (AF), or its new technology equivalent, is to act as first filter, choosing a relatively small number of applications to process further, called *sifting*. Sifting can take up a lot of time in HR departments so any way of speeding it up will be very valuable, so long as it is fair and accurate. Research suggests sifting is not always done very effectively. Machwirth, Schuler and Moser (1996) used *policy capturing* analysis, which works back from the decisions HR make about a set of applications to infer how HR decides. Machwirth *et al.* showed what HR *do* often differs from what they *say*. Managers said they sifted on the basis of proven ability and previously achieved position, but in practice rejected applicants because the application looked untidy or badly written. McKinney *et al.* (2003) analysed how US campus recruiters used grade point average (GPA; college course marks) to select for interview. Some chose students with high marks, which is the logical use of the information, given that GPA does predict work performance to some extent, and that it is linked to mental ability, which also predicts work performance. A second large group ignored GPA altogether. A third group selected for lower GPA, screening out any As with high grades, which does not seem a good way to sift, given the link between work performance and mental ability. The choice of strategy seemed essentially idiosyncratic, and not linked to type of job or employer.

Accuracy and honesty

Numerous surveys report that alarming percentages of AFs, résumés and CVs contain information that is inaccurate, or even false. These surveys often seem to have a 'self-serving' element, being reported by organizations that offer to verify information supplied by As; not much independent research has been

reported. Goldstein (1971) found many applicants for nursing vacancies exaggerated both previous experience and salary. More seriously, a quarter gave a reason for leaving that their previous employer did not agree with, and 17% listed as their last employer someone who denied ever having employed them. McDaniel, Douglas, and Snell (1997) surveyed marketing, accounting, management and computing professionals, and found that 25% to 33% admitted misrepresenting their experience or skills, inflating their salary, or suppressing damaging information, such as being sacked. Keenan (1997) asked British graduates which answers on their application forms they had 'made up ... to please the recruiter'. Hardly any admitted to giving false information about their degree, but most (73%) admitted they were not honest about their reasons for choosing that employer, and 40% felt no obligation to be honest about their hobbies and interests. Electronic media, such as the Internet, do not bypass these problems. It is just as easy to lie through a keyboard as it is on paper or in person.

RESEARCH AGENDA

- The accuracy of CV and application form information.
- What sort of information is wrongly reported.
- What sort of people report false information.
- Why people report wrong information.
- Whether the amount of incorrect information is increasing.
- The role of careers advice, coaching, self-help books and websites.

Fairness and sifting

Equal opportunities (EO) agencies in the USA have produced long lists of questions that application forms should not ask for one reason or another. Some are obvious: ethnicity, gender, and disability (because the law forbids discrimination in all three). Others are less obvious: for example AFs should not ask about driving offences, arrests or military discharge, because some minorities have higher rates of these, so the question may create indirect discrimination. Questions about availability over holidays or weekends may discourage some religious minorities. A succession of surveys (reviewed by Kethley & Terpstra, 2005) have consistently shown that most US employers seem unaware of, or unconcerned by, this guidance and continue to ask questions the agencies say they should avoid. Kethley and Terpstra review 312 US Federal cases involving AFs, and find complaints centred on sex (28%), age (25%), and race (12%). Some questions listed as 'inadvisable' – military discharge, marital status, arrest – have never been the subject of a court case. Internet recruitment and selection could raise another set of 'fairness' issues, because not everyone has access to the Internet. In 2014 a UK government-run

recruitment system caused some embarrassment – to itself and the government – by advertising vacancies as suitable for ‘recent graduates’, alleged to be a code word for ‘young’.

Bias in sifting

Many studies have used the *paper applicant* method, which prepares sets of equally suitable As who differ in one key feature – gender, age, having a beard, etc. – then has HR staff rate their suitability. This is an easy type of research to do, and one that usually ‘gets results’, by finding evidence of bias:

- Davison and Burke (2000) reviewed 49 studies of gender bias, and found both male and female sifters biased against female As. The less job information is given, the greater the bias.
- Ding and Stillman (2005) reported New Zealand data showing overweight female As tend to be sifted out.
- Correll, Benard and Paik (2007) find women with children tend to be sifted out, but men with children are not, and may even be favoured.

Paper applicant research has a flaw, however: the sifters know they are being scrutinized by psychologists, so may be on their best behaviour. Also they are not really hiring As, and will not have to work with the people they ‘select’.

Research on sifting in the USA had reached the reassuring conclusion that it seemed free of racial bias, but a study by Bertrand and Mullainathan (2004) suggested there may be a problem after all. They used a different approach, called the *audit* technique. They sent their ‘paper applicants’ to real employers, applying for real jobs, and counted how many were shortlisted for interview. Choice of first name identified A as white or African American. (Americans will assume ‘Brad’ and ‘Carrie’ are white, while ‘Aisha’ and ‘Leroy’ are African American). For every 10 ‘white’ As called for interview, there were only 6.7 ‘African Americans’; African Americans were being sifted out by ethnicity. Bertrand and Mullainathan could argue their data show what is really happening in the real US job market, which justifies the slightly unethical practice of sending employers fake job applications. The International Labour Organization seems to approve, for it publishes a manual on how to conduct ‘natural experiments’ to test for discrimination. Some research, described in Chapter 4, takes this method a step further, by accepting invitations to interview, and noting how the interview proceeds. The audit method is in one respect even easier than the paper person method, because HR do not have to agree to participate. Hoque and Noon (1999) wrote to British employers enquiring about possible vacancies, not applying for a specific job, calling themselves ‘Evans’, implying a white person, or ‘Patel’, implying a South Asian person. ‘Evans’ got on average slightly longer and more helpful replies. McGinnity and Lunn (2011) find Irish applicants in Ireland twice as likely as African, Asian or German applicants to be interviewed. McGinnity and Lunn note the effect seemed stronger than found elsewhere and suggest this may

reflect the low number of non-Irish people in Ireland, and a strong feeling of national identity. The data were collected between March and September 2008, just before the financial crisis that started the recession; a replication today would be interesting. In the Netherlands, applicants with Arab-sounding names are four times less likely to be called back by the employer (Derous, Ryan & Nguyen, 2012). Agerström *et al.* (2012) get the same result in Sweden. It is sometimes argued that providing more information about people will avoid 'snap' judgements based on apparent race or nationality, but Agerström *et al.* find that providing information about Erik's coldness and lack of commitment or Hassan's warmth and high commitment did not prevent discrimination. Research has widened to include other possibly discriminated-against classes of applicant.

Social class

Jackson (2009) confirms the continuing importance of social class in Britain. Applicants with high-status names (Charles Bartle-Jones vs Gary Rodgers) or with high-status pastimes (polo vs darts), or who have been to public (i.e. private) rather than state schools get (slightly) more favourable responses from employers.

Pregnancy

Morgan *et al.* (2013) list four elements of some US employers' perception of pregnant women as job applicants: lower competence, lack of commitment, inflexibility, and as needing 'accommodation' (changes in working conditions or hours etc.). Morgan *et al.* employ a variant of the 'audit' approach, in which women go into a department store and ask if it has any jobs, and to complete an application form. Sometimes they wear a pregnancy prosthesis that makes them look about five months pregnant, and they are provided with four different scripts to counter one of the four stereotypes listed above: for example, 'I have the help I need so I can work whenever you need me'. The scripts had some effect in reducing discrimination.

Age

Ng and Feldman (2012) list six common stereotypes of older workers: that they are less motivated, less trusting, less healthy, more resistant to change, more vulnerable to work-family imbalance, and less willing to participate in training and development. All would tend to result in older As being sifted out at the shortlisting stage. Ng and Feldman also review evidence on actual age differences and conclude the first five stereotypes are false, but there is evidence that older workers are less willing to participate in training and development. In an earlier paper Ng and Feldman (2008) examined the actual correlation of age with a range of work performance measures, and found no age differences in most: core task performance, creativity, training programme performance, safety performance, and counterproductive work behaviour

(ranging from taking too long a break to wrecking machinery). In some respects older workers perform better: less often late, less often absent, more likely to go beyond the letter of their job description. The correlation of age with (avoiding) absence and lateness is 0.26–0.28, which is as good as many selection tests achieve. Ng and Feldman's results suggest the stereotyped view of older workers is incorrect, and that in some ways they are better employees.

Weight

Agerström and Rooth (2011) report an audit study from Sweden. Pairs of parallel real applications were sent to employers advertising a range of jobs through the Swedish Employment Agency, one with a photo (face only) of an obese person, one with a photo of a similar but not obese person. The obese person was less than half as likely to be invited for interview. Two months later some of the HR managers who had done the sifting agreed to do an implicit association test (described in detail in Chapter 8) which showed they tended to think of obese persons as ineffective, slow, lazy, incompetent, and lacking in initiative. The implicit association test does not ask sifters to rate, for example, obese people on, for example, energy, but aims to uncover automatic associations sifters have, but may be unaware of.

Gender and 'backlash'

'Agentic' people are ambitious, task-oriented and assertive, whereas 'communal' people are caring, helpful and collaborative. In Western countries gender stereotypes expect men to be agentic, and women to be communal. Carlsson *et al.* (2014) report an audit study on CVs that conform to this stereotype, or do not: caring, helpful men, and ambitious, assertive women. They test the hypothesis that there will be a 'backlash' against applicants who do not fit the stereotype. However over 5,000 real applicants to 3,000 real jobs in Sweden showed no evidence of 'backlash'.

IMPROVING APPLICATION SIFTING

Behavioural competences

Applicants are asked to describe things they have done which relate to key competences for the job. *Ability to influence others* is assessed by A describing an occasion when A had to persuade others to accept an unpopular course of action. This method might improve the AF as a selection assessment, but there is no research on whether it does.

Weighted application blanks and biodata

Application forms can be converted into weighted application blanks (WABs) by analysing past and present employees for predictors of success (Chapter 9). One study found American female bank clerks who did not stay long tended

to be under 25, single, to live at home, and to have had several jobs (Robinson, 1972), so banks could reduce turnover by screening out As with these characteristics. (Robinson's list probably would not be legal today, however, because it specifies female bank clerks.) Most WABs are conventional paper format, but the technique would work equally well for electronic applications. Biodata also uses biographical items to select, but collects them through a separate questionnaire, not from the AF.

Training and experience (T&E) ratings

In the USA application sifting used to be assisted by T&E ratings which sought to quantify applicants' training and experience by various rating systems, instead of relying on arbitrary judgements. T&E ratings seem to have been overtaken in the USA by various Internet application coding systems. Note, however, that T&E ratings had extensive research (McDaniel, Schmidt and Hunter, 1988) showing they do actually predict work performance; similar research for Internet application coding has yet to be reported.

Minimum qualifications (MQs)

The advertisement says As need a civil engineering qualification plus a minimum of five years' experience, the intended implication being that people who lack these will not be considered, so should not apply. MQs are generally based on education and experience. However, educational MQs may exclude some minorities, while length of experience may exclude women, who tend to take more career breaks. Hence in the USA MQs may be challenged legally, and so need careful justification. Buster, Roth and Roth (2005) described elaborate systems of panels of experts, discussions and rating schedules for setting MQs. (As opposed to setting an arbitrary MQ, or using the 'one we've always used', or the 'one everyone uses'.) For example the experts might be asked to 'bracket' the MQ; if it is suggested that three years' experience is needed, then ask the experts to consider two and four years as well, just to make sure three years really is the right amount.

Background investigation, aka positive vetting

Application forms contain the information applicants choose to provide about themselves. Some employers make their own checks on As, covering criminal history, driving record, financial and credit history, education and employment history, possibly even reputation and lifestyle. Background checking is rapidly growing in popularity in the USA, from 51% of employers in 1996 to 85% in 2007 (Isaacson *et al.*, 2008), possibly driven by several high-profile cases where CEOs have been caught falsifying their CVs. In Britain, background investigations are recommended for childcare workers, and used for government employees with access to confidential information (known as *positive vetting*). Presently there is little or no research on whether background checks succeed in selecting 'good' employees and rejecting unsuitable ones.

Isaacson *et al.* compare As who failed a background check with those who passed, and find those who failed score slightly higher on a test of risk-taking. The closest they could get to work performance was a realistic computer simulation of manufacturing work, where the failed group worked slightly faster but slightly less well. Roberts *et al.* (2007) report a long-term follow up of a New Zealand cohort of 930 26-year-olds, which found no link between criminal convictions before age 18 and self-reported counterproductive work behaviour.

Internet tests

Many employers are replacing their conventional paper application forms by short tests completed over the Internet. Some assess job knowledge; it can be useful to screen out people who know little or nothing about subjects, for example, Microsoft Excel, they claim expertise in. Testing can improve the whole selection process by screening out, early on, As who lack the mental ability necessary for the job. (Chapter 6 will show that mental ability is generally a good predictor of work performance.) In conventional selection systems, tests are not normally used until the shortlist stage, by which time many able As may have been screened out. It is theoretically preferable to put the most accurate selection tests early in the selection process, but the cost of conventional paper-and-pencil testing tends to prevent this. Some Internet tests assess personality or fit. Formerly HR inferred, for example, leadership potential from what As said they did at school or university. Some Internet sifting systems assess it more directly by a set of standard questions. No research has been published on how well such systems work.

Application scanning software

Numerous software systems exist that can scan applications and CVs to check whether they match the job's requirements. This is much quicker than conventional sifting of paper applications by HR. Automated sifting systems can eliminate bias directly based on ethnicity, age, disability or gender, because they are programmed to ignore these factors. However they will not necessarily ignore factors *linked* to ethnicity, disability, age or gender, such as sports and pastimes. Sifting software will do the job consistently and thoroughly, whereas the human sifter may get tired or bored and not read every application carefully. Sifting electronically is not necessarily any more accurate. Accuracy depends on the decision rules used in sifting, which in turn depend on the quality of the research they are based on. Reports (Bartram, 2000) have suggested some scanning systems do nothing more sophisticated than search for key words; once applicants realize this, they will take care to include as many as possible. There is an urgent need to know what application sifting programs actually do. Psychologists tend to be rather sceptical, for one fairly simple reason. If these systems are doing something tremendously subtle and complex, where did the people who wrote them acquire this wisdom? There is

no evidence that human application sifters are doing anything highly complex that software can model, nor is there any body of research on application sifting that has described any complex, subtle relationships to put into software. Stone *et al.*'s (2013) survey concludes that little or no research has been reported on several fairly important questions, including keyword searching in applications.

RESEARCH AGENDA

- The link between various application sifting systems and later work performance, for competence based applications, background investigations, Internet testing, application scanning and sorting software systems.
- Policy capturing research on application scanning and sorting software systems.
- Investigation of how application sifting software operates, and what it can achieve.

OVERVIEW OF SELECTION METHODS

The first column in Table 1.1 lists the main techniques used to select staff in North America, Europe, and other industrialized countries. The list is divided into traditional and 'new', although most 'new' methods have been in use for some time. Table 1.1 also indicates which chapter contains the main coverage of each method.

WHAT IS ASSESSED IN PERSONNEL SELECTION?

The short answer to this question is: ability to do the job. A much more detailed answer is provided by job analysis, which lists the main attributes successful employees need (see Chapter 3). Table 1.2 lists the main headings for assessing staff.

Mental ability

divides into general mental ability (or 'intelligence'), and more specific abilities such as problem solving, clerical ability, or mechanical comprehension. Some jobs also need sensory abilities: keen hearing, good balance, or good eye-hand co-ordination

Physical characteristics

Some jobs need specific physical abilities: strength, endurance, dexterity. Others have more implicit requirements for height or appearance.

Table 1.1 Traditional and new(er) selection assessment methods.

Traditional methods	Chapter	Alternative names
Application form/CV/résumé	1	
Traditional interview	4	
References	5	
New(er) methods		
Electronic application	1	
Structured interview	4	
Peer rating	5	
Mental ability test	6	Aptitude test
Personality questionnaire	7	Personality inventory
Honesty test	7	Integrity test
Projective test	8	
Graphology	8	Handwriting analysis
Biodata	9	Weighted application blank
Assessment centre	10	Extended interview
Emotional intelligence	11	Situational judgement Social intelligence
Work sample test	11	Trainability test, in tray/basket
Physical ability test	11	
Drug use testing	11	

Table 1.2 Seven main aspects of applicants assessed in selection.

Mental ability
Personality
Physical characteristics
Interests and values
Knowledge
Work skills
Social skills

Personality

Psychologists list from five to 30 underlying dispositions, or personality traits, to think, feel and behave in particular ways. An extravert person likes meeting people, feels at ease meeting strangers, etc. The employer may find it easier to select someone who is very outgoing to sell insurance, rather than trying to train someone who is presently rather shy.

Interests, values and fit

Someone who wants to help others may find charity work more rewarding than selling doughnuts; someone who believes that people should obey all the rules all the time may enjoy being a traffic warden. People cannot always

find work that matches their ideals and values, but work that does may prove more rewarding. 'Fit' means the applicant's outlook or behaviour matches the organization's requirements. These can be explicit: soldiers expect to obey orders instantly and without question. Fit may be implicit: the applicant may not sound or look 'right for us', but there is no written list of requirements, or even a list that selectors can explain to you.

Knowledge

Every job requires some knowledge: current employment law, statistical analysis, or something much simpler, how to use a telephone, how to give change. Knowledge can be acquired by training, so need not necessarily be a selection requirement. Mastery of higher-level knowledge may require higher levels of mental ability.

Work skills

The ability to do something quickly and efficiently: bricklaying, driving a bus, valuing a property, diagnosing an illness. Employers sometimes select for skills, and sometimes train for them. Mastery of some skills may require levels of mental or physical ability not everyone has.

Social skills

These are important for many jobs; and essential for some. They include communication, persuasion, negotiation, influence and leadership, teamwork, etc. Hogan, Chamorro-Premuzic and Kaiser (2013) argue that the changing nature of work in the developed world means social skills are now much more important, especially in the service sector, and in teamwork. Hogan *et al.* also suggest that being able to get on with colleagues in general has always been overlooked in selection.

Construct and method

Arthur and Villado (2008) note that many reviews fail to distinguish clearly between *construct* and *method*. Construct refers to *what* is being assessed, for example, personality; method refers to *how* it is assessed, for example, personality questionnaire.

NATURE OF THE INFORMATION COLLECTED

Discussions of selection methods tend to focus on the merits of personality questionnaires (PQs), or structured interviews, or work samples; they do not usually address the issue of what sort of information the method generates. Table 1.3 sorts selection methods by five qualitatively different types of information.

Table 1.3 Five categories of qualitatively different information, obtained by selection tests.

Self-report	Information provided by the applicant <i>Application form, including online application, T&E rating, biodata, personality questionnaire, honesty test, projective test, interest questionnaire, interview.</i>
Reported	Information provided by other people about the applicant <i>References, peer rating.</i>
Demonstrated	The applicant performs a task or demonstrates a skill
a) Test	<i>Work sample, mental ability test, job knowledge test, physical ability test.</i>
b) Behavioural	<i>Group exercise, behavioural test.</i>
Recorded	Information that is recorded somewhere <i>School marks, degree class, professional qualification, published work</i>
Involuntary	Information the applicant may not be able to control consciously <i>Graphology, drug use testing, polygraph, psychophysiology, voice stress analysis.</i>

Self-report evidence

Information that is provided by the applicant, in written or spoken form, on the application form, in the interview, and when answering personality questionnaires, attitude measures, and biographical inventories. Some self-reports are free-form or unstructured, for example some interviews or application forms. Other are more structured, such as personality questionnaires, biodata, or structured interviews. Some self-reports are fairly transparent, notably interviews and personality questionnaires. (Transparent in the sense that As will have little difficulty working out what inference will be drawn from what they say.) Other assessments may be less transparent, such as biodata, or projective tests; As may find it less easy to decide what answer will be seen as 'good' or 'poor'.

Self-report data have some compelling advantages in selection. The system is generally very cheap and very convenient; applicants are present, and eager to please, so collecting information is easy. Self-report can also be justified as showing respect and trust for applicants. It can be argued that some questions can only be answered by the person him/herself. It is not usually very sensible to start an argument with someone about whether they are satisfied with their job. If they say they are, then they should know. Unless, however, there is any reason to think they are not telling the truth, which may be a possibility in the context of personnel selection. In fact self-report has a fundamental disadvantage in selection: As provide the information, and the employer generally has no way of verifying it. Self-report has two other limitations: coaching, and lack of insight. There are many books on how to complete job applications; career counselling services advise students what to say at interviews. The second

problem is lack of self-insight. Some As may genuinely think they are good leaders, or popular, or creative, and incorporate this view of themselves into their application, PQ, or interview. However by any other criterion – test, others’ opinion, or achievement – they lack the quality in question. This issue has not been researched much, if at all, in the selection context. These problems make it important to confirm what applicants say about themselves by checking information from other sources. Kruger and Dunning (1999) note that people who score badly on tests often greatly overestimate their performance, suggesting that people may lack the ability to recognize their own shortcomings. This has been dubbed the Dunning Kruger effect.

Self-report can also create *common method variance* problems. If all the information about Jones’s personality, and Jones’s job performance, comes from Jones him/herself, there is a risk that a correlation might reflect something other than a true link between the two, such as a general optimism, or general pessimism, or simply not telling the truth. Research should try to avoid common method variance problems by getting information about aspects of Jones from different sources, or by having some way of checking what he or she says.

Other report evidence

Information about the applicant is provided by other people, through references or ratings. Other reports vary in the degree of expertise involved. Some require no special expertise, such as peer ratings and the letter of reference. Others use experts, generally psychologists.

Demonstrated evidence

The applicant performs a task or demonstrates a skill. Tests include general mental ability (GMA)/intelligence tests, as well as tests of aptitudes, and specific knowledge (trade or job knowledge or achievement tests). These are real tests, with right and wrong answers. Demonstrated evidence also includes work samples, group exercises, simulations and other behavioural exercises typically included in assessment centres. Demonstration evidence has fewer limitations than self- or other reports. Ability tests cannot generally be faked. On the down side, demonstrated evidence tends to be more difficult and expensive to collect.

Recorded evidence

Some information used in selection can be characterized as recorded fact. The applicant has a good degree in psychology from a good university. The information is recorded, and verifiable. (Although some employers make the mistake of relying on self-report data, and failing to check applicants’ qualifications at source.) Work history can also provide a record of achievement, for example the applicant was CEO/MD of organization XYZ when XYZ’s

profits increased. Published work, grants obtained, inventions patented, prizes and medals also constitute recorded evidence.

Demonstrated and recorded information tends to have an asymmetric relationship with self- or other reported information. Evidence that someone cannot do something disproves the statement by the applicant or others that he/she can. However the converse is not true: being told that someone cannot do something does not disprove demonstrated or recorded evidence that he or she can. To this extent, demonstrated and recorded evidence is superior to self- and other reported evidence.

Involuntary evidence

Some evidence is provided by applicants, but not from what they tell the assessors, nor from things they do intentionally. The classic example is the polygraph, intended to assess A's truthfulness from respiration, heart rate and electrodermal activity, not from the answers A gives. In fact the polygraph is used to decide which of A's self-reports to believe, and which to classify as untrue. Two other involuntary assessments are graphology and drug use testing. The former seeks to infer As' characteristics from the form of their handwriting, not from its content. Drug use testing assumes that drug use can be more accurately detected by chemical analysis of blood or urine than by self-report. DNA tests of mental ability or personality might – or might not – be devised at some future time. Some types of demonstrated evidence seem to have an almost magical quality about them which appeals to some.

WORK PERFORMANCE

Selection research compares a *predictor*, meaning a selection test, with a *criterion*, meaning an index of the person's work performance. This can be very simple when work generates something that can be counted: widgets manufactured per day, or sales per week. It appears very simple if the organization has an appraisal system whose ratings can be used. The supervisor rating criterion is widely used, because it is almost always available (in the USA), because it is unitary, and because it is hard to argue with.

On the other hand finding a good criterion can soon get very complex, if one wants to dig a bit deeper into what constitutes effective performance. Questions about the real nature of work, or the true purpose of organizations soon arise. Is success better measured objectively by counting units produced, or better measured subjectively by informed opinion? Is success at work unidimensional or multi-dimensional? Who decides whether work is successful? Different supervisors may not agree. Management and workers may not agree. The organization and its customers may not agree.

Objective criteria are many and various. Some are more objective than others; *training grades* often involve some subjective judgement in rating written work. *Personnel criteria* – advancement/promotion, length of service,

turnover, punctuality, absence, disciplinary action, accidents, sickness – are easy to collect (but can be hard to interpret). Analyses of selection research (Lent, Aurbach & Levin, 1971) have shown that a subjective criterion – the global supervisor rating – was clearly the favourite, used in 60% of studies. Criteria of work performance are discussed in greater detail in Chapter 12.

FAIR EMPLOYMENT LAW

Most people know it is against the law to discriminate against certain classes of people when filling vacancies. These protected classes include women, ethnic minorities, and disabled people. Most people think discrimination means deciding not to employ Mr Jones because he is black or Ms Smith because she is female. Direct discrimination is illegal, but is not the main concern in personnel selection. The key issue is indirect discrimination, or *adverse impact*. Adverse impact means the selection system results in more majority persons getting through than minority persons. For example some UK employers sift out applicants who have been unemployed for more than six months, on the argument that they will have lost the habit of working. The former Commission for Racial Equality (CRE) argued this creates adverse impact on some ethnic minorities, because their unemployment rates are higher. Adverse impact assesses the *effect* of the selection method, not the *intentions* of the people who devised it. Adverse impact means an employer can be proved guilty of discrimination, by setting standards that make no reference to ethnicity or gender. Adverse impact is a very serious matter for employers. It creates a presumption of discrimination, which the employer must disprove, possibly in court. This will cost a lot of time and money and may generate damaging publicity. Selection methods that do not create adverse impact are therefore highly desirable, but unfortunately are not always easy to find. Fair employment issues are discussed in detail in Chapter 13.

CURRENT SELECTION PRACTICE

Surveys of employers' selection methods appear quite frequently, but should be viewed with some caution. Return rates are often very low: Piotrowski and Armstrong (2006) say 20% is normal. There is also the grey (and black) side of selection. Some methods are not entirely legal or ethical, so employers are unlikely to admit to using them. Rumours suggest that some employers gain unauthorized access to criminal records by employing former police officers, or use credit information to assess applicants (forbidden by credit agencies in the UK). There are even rumours of secret databases of people to avoid employing, because they are union activists or have complained about unsafe working conditions. Many organizations forbid the use of telephone references, but Andler and Herbst (2002) suggest many managers nevertheless both ask for them, and provide them.

Selection in Britain

Table 1.4 presents three recent UK surveys, by IRS (Murphy, 2006), CIPD (2006), and Zibarras and Woods (2010). Table 1.4 confirms earlier UK surveys, showing that most UK employers are still using interviews of various types, that most still use references, that most use tests at least some of the time, but less frequently online. Only half use assessment centres or group exercises, while few use biodata. Two surveys gave no information about return rate. Zibarras and Woods survey organizations of various sizes from a handful of

Table 1.4 Three surveys of UK selection, by CIPD (2006), IRS (Murphy, 2006) and Zibarras & Woods (2010).

	CIPD	IRS	Z&W
% Return rate ^a			10
Sample size	804	100	579
Application form		85	60
CV		20	85
Criminal check			27
<i>Interview</i>			
Unstructured IV			42
Face to face IV		98	
Panel IV		28	
Structured IV			69
Structured panel	88		
Structured one to one	81		
Competency based	85		
Telephone	56	32	
<i>References</i>			
References		85	72
Employment ref (pre-interview)	49		
Academic ref (pre-interview)	36		
<i>Tests</i>			
Tests for specific skills	82		
General ability tests	75		39
Literacy/numeracy	72		28
Personality/aptitude Qs	60		26
Psychometric tests (mostly PQs)		64	
Online test	25		
Biodata		4	27
Work sample			19
<i>Behavioural assessments</i>			
Assessment centre	48	35	17
Group exercise	48		15

^aCIPD % are employers who use that method (rarely/occasionally/frequently). IRS and Zibarras & Wood % are employers who use that method (extent/frequency unspecified).

employees to more than 1,000, and find no significant differences in choice of selection methods. They find differences between different sectors, which they characterize as the public and voluntary sectors more being likely to use formalized methods, for example application forms rather than CVs, and structured interviews rather than unstructured ones.

Selection in Europe

European countries favour a social negotiation perspective on selection, which emphasizes employee rights, applicant privacy, and expectation of fair and equitable treatment. Salgado and Anderson (2002) concluded that mental ability (MA) tests are now more widely used in Europe than in the USA. The most recent comprehensive survey of European practice remains the Price Waterhouse Cranfield survey from the early 1990s (Dany & Torchy, 1994) which covered 12 western European countries and nine methods. Table 1.5 reveals a number of interesting national differences:

- The French favoured graphology but no other country did.
- Application forms are widely used everywhere except in the Netherlands.
- References were widely used everywhere but less popular in Spain, Portugal and the Netherlands.

Table 1.5 Summary of surveys of selection test use in Europe (Dany & Torchy, 1994; Schuler *et al.*, 2007; König *et al.*, 2010).

		AF	IV	Psy	Gph	Ref	Apt	AC	Grp
UK	1994	97	71	46	1	92	45	18	13
Ireland	1994	91	87	28	1	91	41	7	8
France	1994	95	92	22	57	73	28	9	10
Portugal	1994	83	97	58	2	55	17	2	18
Spain	1994	87	85	60	8	54	72	18	22
Germany	1994	96	86	6	6	66	8	13	4
	1993 ^a	98	60/63	21	9	71	34	39	51
	2007 ^a	99	73/42	20	2	57	30	58	42
Switzerland	2010 ^b	100	99	32 ^c	16	89	19	26	nr
Netherlands	1994	94	69	31	2	47	53	27	2
Denmark	1994	48	99	38	2	79	17	4	8
Finland	1994	82	99	74	2	63	42	16	8
Norway	1994	59	78	11	0	92	19	5	1
Sweden	1994	na	69	24	0	96	14	5	3
Turkey	1994	95	64	8	0	69	33	4	23

Methods: AF – application form; IV – interview panel; Psy – psychometric testing; Gph – graphology; Ref – reference; Apt – aptitude test; AC – assessment centre; Grp – group selection methods.

^aData from Schuler *et al.* (2007).

^bData from König *et al.* (2010).

^cPersonality test.

- Psychometric testing was most popular in Spain and Portugal and least popular in West Germany and Turkey.
- Aptitude testing was most popular in Spain and the Netherlands and least popular in West Germany and Turkey.
- Assessment centres were not used much but are most popular in Spain and the Netherlands.
- Group selection methods were not used much but were most popular in Spain and Portugal.

Schuler *et al.* (2007) report a survey of 125 German HR managers, which they compare with a similar survey from 1993. Methods that have become less popular include references, unstructured interviews, biographical questionnaires, and medical examinations. Methods that have become more popular include structured interviews and assessment interviews. The 2007 survey also collected data on perceived validity, and perceived acceptability to applicants. Methods seen as having high validity include structured interview, assessment centre, group discussion and work sample; methods perceived as having low validity include graphology and online PQ. This indicates the German HR people seem generally well informed about selection test validity, with the possible exception of tests of mental ability. The 2007 survey includes some information linking selection methods to type of vacancy. Use of references seems to be linked very closely to status; references are most likely to be used for executives, are rarely used for clerical and retail jobs, and are not used at all for unskilled or skilled workers. Assessment centres are used for apprentices, trainees and management. MA tests are only used for selecting apprentices.

Selection in the USA

Piotrowski and Armstrong (2006) report the most recent US survey, of 151 companies in the Fortune 1000. US employers used application form, résumé and reference check virtually without exception. Half used 'skills testing', and a substantial minority used personality tests and biodata. A few employed drug use testing. Piotrowski and Armstrong did not enquire about use of interviews.

Selection further afield

Less is known about selection in other parts of the world. Surveys of New Zealand (Taylor, Keelty & McDonnell, 2002) and Australia (Di Milia, 2004) found a very similar picture to Britain; interview, references and application were virtually universal, with personality tests, ability tests and assessment centres used by a minority, but gaining in popularity. Arthur *et al.* (1995) described selection in Nigeria and Ghana; interviews were nearly universal (90%), references widely used (46%); paper-and-pencil tests were less frequently used, as were work samples (19%) and work simulations (11%). Ryan *et al.*'s (1999) survey covered no fewer than 20 countries, although some

samples were rather small. Mental ability tests were used most in Belgium, the Netherlands, and Spain, least in Italy and USA. Personality tests were used most in Spain, least in Germany and USA. Projective tests were used most in Portugal, Spain, and South Africa, and least in Germany, Greece, Hong Kong, Ireland, Italy, and Singapore. Drug tests were used most in Portugal, Sweden, and USA, and least in Italy, Singapore, and Spain. Ryan suggested the data confirmed a prediction from Hofstede's (2001) discussion of national differences in attitudes to work: countries high in *uncertainty avoidance* (see Box 1.1) used more selection methods, used them more extensively, and used more interviews. Huo, Huang and Napier (2002) surveyed 13 countries including Australia, Canada, China, Indonesia, Taiwan, Japan, South Korea, Mexico, USA, and Latin America. They found interviews were very widely used, but less so in China and South Korea. Some countries, including Mexico, Taiwan, and China, based selection partly on connections (school, family, friends, region, or government). Selection in Japan emphasized ability to get on with others, possibly because Japanese employers traditionally offer people lifelong employment.

Economic climate

A very plausible hypothesis states that high unemployment will mean that employers need worry less about what applicants think of selection methods, so some might feel free to use less popular methods, such as mental ability (intelligence) tests, or personality questionnaires. The years from 2007 to the present (2015) would have been a good time to test this, an opportunity, however, that seems to have been missed.

Box 1.1 Uncertainty avoidance

Uncertainty avoidance means organizations do not like unpredictable situations, and maintain predictability by adhering to formal procedures and rules. Countries that tend to be high in uncertainty avoidance include Greece and Portugal, while countries low in uncertainty avoidance include Singapore.

REASONS FOR CHOICE OF SELECTION METHOD

One survey (Harris, Dworkin & Park, 1990) delved a little deeper and asked why personnel managers choose different selection methods. Factors of middling importance were fakability, offensiveness to the applicant, and how many other companies use the method. Interviews, although very widely used, were recognized to be not very accurate, as well as easy to fake; Harris *et al.* suggest personnel managers are aware of the interview's shortcomings, but continue using it because it serves other purposes besides assessment. Terpstra and Rozell (1997) by contrast asked personnel managers why they did *not* use particular methods. Some they did not think

useful: structured interviews, mental ability tests. Some they had not heard of: biodata. They also avoided mental ability tests because of legal worries. Muchinsky (2004) noted that the commonest questions American managers ask about selection tests are *How long will this take?* and *How much will it cost?* not *How accurate is it?*

Box 1.2 Role Repertory Grid

The Role Repertory Grid is method of finding out what people think about things, such as selection procedures, or the nature of work (Chapter 3), that seeks to dig a little deeper than simply asking *Do you use interviews to select, if so why?* HR people are given a list of selection methods, and asked to consider them in sets of three at a time, for example interview, reference and personality questionnaire. They are then asked to say which two differ from the third, for example PQ differs from interview and reference. They then have to say *how* they differ, to which one answer might be that interviews and obtaining references are traditional ways of selecting staff while PQs are a recent invention. This is repeated, with different sets of three selection methods up to 20 times. Respondents are usually forbidden to give the same reason twice, and are usually asked to apply the distinction they have drawn, for example *traditional/recent invention*, to all the other selection methods in the list. The Repertory Grid method is intended to discover the person's own conceptual framework for thinking about selection, rather than to impose the researcher's.

König, Jöri and Knüsel (2011) use the Repertory Grid method (Box 1.2) to uncover ideas about 11 selection methods in 40 Swiss HR practitioners. Some of their distinctions are purely descriptive, even practical: *spoken vs written*; *need to bring in outside people to do it vs can do it ourselves*. Other distinctions show a concern with the nature of the information: *candidate's view of him/herself vs other people's view of candidate*. Things the work psychologist might think very important were mentioned by only a few HR people: five of the 40 mentioned validity, while only four mentioned fakability. Most surprising perhaps are the things not one of the 40 mentioned: law, applicant reaction, or cost. Koenig *et al.* speculate (their word) that Swiss HR may be unconcerned about cost because of a 'generally good economic situation', and less worried about legal problems because Switzerland is 'relatively free of legal pressure'. Eleven of the 40 viewed selection methods in terms of *the present vs long-term axis into the past*. Koenig *et al.* suggest some may be influenced by psychoanalysis. As Koenig *et al.* note, there is much scope for more research here.

Asking applicants

All the surveys discussed so far ask HR how they select. Billsberry (2007) presents 52 UK accounts of selection procedures by those on the receiving end. The accounts amply confirm the hypothesis that some of the 80% of employers who do not reply to surveys have something to hide. Applicants describe rudeness, unprofessional behaviour, blatant lying, obvious bias, and sexual

harassment. The most generally favoured form of assessment seems to be the interview, often conducted very incompetently. Billsberry's data suggest that a large survey of job applicants is an urgent necessity to find how many employers are behaving badly towards applicants.

RESEARCH AGENDA

- Employers' reasons for choosing selection methods, and more use of repertory grid methods.
- Information from applicants about use of selection methods.

KEY POINTS

In Chapter 1 you have learned the following:

- Employees vary greatly in value so selection matters.
- How employees are recruited may be linked to turnover.
- Deciding which application to proceed with and which to reject is called sifting and is often done inefficiently or unfairly.
- Sifting can be improved by T&E ratings, and careful setting of minimum qualifications.
- Conventional paper application methods can be improved.
- The Internet may greatly change the application process.
- Sifting software is something of an unknown quantity.
- Selection uses a range of tests to assess a range of attributes.
- Information used in selection divides into five main types.
- Selection methods must conform with fair employment legislation.
- The problem with fair employment is not deliberate or direct discrimination, but adverse impact meaning the method results in fewer women or minority persons being successful. Adverse impact will create problems for the employer so should be avoided if possible.
- Selection in developed countries follows broadly similar patterns, with some local variations.

KEY REFERENCES

Agerström *et al.* (2012) describe possible unconscious bias against overweight applicants in Swedish HR staff.

Bertrand & Mullainathan (2004) describe discrimination in selection in the USA.

Billsberry (2007) presents 52 accounts of how applicants experienced selection.

Buster *et al.* (2005) describe a system for setting minimum qualifications.

Davison & Burke (2000) review research on gender bias in application sifting.

König *et al.* (2011) analyse thinking about selection methods in Swiss HR personnel.

McKinney *et al.* (2003) describe how information on college grades is used in sifting.

Ng & Feldman (2012) analyse six common stereotypes of older employees.

Stone *et al.* (2013) review research on electronic selection and recruitment systems.

USEFUL WEBSITES

checkpast.com. A (US) background checking agency.

zerochaos.com. Another (US) background checking agency.

hrzone.com. Offers advice on range of HR issues in the USA.

incomesdata.co.uk. Income Data Services, a UK company that reports interesting research on HR issues, including survey of selection tests.

siop.org. (US) Society for Industrial and Organizational Psychology, includes details of conferences, and The Industrial/Organizational Psychologist.