

CEE DP 29

**The Growth of Psychometric Testing for Selection:
Why Has Test Use Increased, Will Growth Continue
and What Does this Mean for Education?**

Andrew Jenkins

Alison Wolf

**CENTRE FOR THE
ECONOMICS OF
EDUCATION**

June 2002

Published by
Centre for the Economics of Education
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

© Alison Wolf and Andrew Jenkins, submitted June 2001

ISBN 0 7530 1533 1

Individual copy price: £5

The Centre for the Economics of Education is an independent research centre funded by the Department for Education and Skills. The views expressed in this work are those of the authors and do not necessarily reflect the views of the DfES. All errors and omissions remain the authors.

Executive Summary

There is a growing body of evidence indicating that the use of psychometric tests for selection purposes has increased in recent years. However, the data are almost entirely cross-sectional. Thus, while the propensity to test, and levels of test use, can be related to various organisational characteristics such as size, costs of hiring and firing, or adoption of formal personnel policies, there is very little information on the underlying reasons for *changes* (and increases) in use. (See Jenkins, 2001a; Jenkins and Wolf, 2002.)

This paper reports findings from a series of in-depth interviews with senior human resources staff responsible for using selection testing. These formed part of a more general enquiry into patterns, and correlates, of test use, and drew on hypotheses which had been examined using national cross-sectional datasets. The interviews addressed directly the reasons for changes in test use, and so were deliberately focused on organisations (private and public) in which tests were already being used. As such, they obviously cannot provide national estimates of the general or relative importance of particular factors; but they do provide a first systematic exploration of the dynamics underlying increased test usage, and of likely future patterns. In all, data were collected for 53 organisations: 36 private sector, 12 public sector and 5 which were recently privatised or public-private collaborations.

To explain increasing test use, both of ability tests (*e.g.* numerical or verbal reasoning) and personality tests, we proposed six hypotheses – by no means mutually exclusive.

- *First*, that the increase in test use was a direct response to demand from senior managers, perhaps because of a loss of confidence in school-based formal qualifications and/or the standard of degrees.

This hypothesis was not supported by the data. There are certainly a few senior managers who are very enthusiastic about tests, but the use of tests for selection is almost always initiated by Human Resources departments. The strongly negative views about academic standards which (according to the CBI and Institute of Directors) senior managers now hold, did not appear to be having any major direct effect on recruitment practice. Conversely, it is unlikely that managers will react against test use. The data indicate that, while many are initially sceptics, there is a growing acceptance of and enthusiasm for

psychometric testing among senior staff. Many managers have come to accept tests as providing up-to-date information on general skills such as quantitative reasoning, thus complementing qualification-based evidence which may be decades old; and as providing data on a variety of skills and traits which are not suited to formal certification. (In part, no doubt, this is because growing numbers have themselves taken such tests prior to being appointed to their jobs.)

- *Second*, that the increase in test use was because of changes in the commercial or public environment in which the organisation found itself: so that while the tests themselves might be no different from a few years back – which is indeed the case – the information they offered was valued more than it had been in the past.

This hypothesis was supported by the data. There appear to be two major forces at work here. First, *any* reliable and valid information is more valuable than in the past both because the difficulty of dismissing staff means that decisions to hire are increasingly high-stakes, and because companies perceive the public and the government as increasingly unaccepting of visible failures. The hiring of low-paid staff in front-line jobs, or with health and safety responsibilities, is consequently treated with increasing seriousness. The second important factor is a trend away from the use of qualifications for selection, and, therefore, an increase in the value accorded to information from the alternative source of ability/attainment and other tests. The downgrading of qualifications is not a response to a perceived fall in their accuracy, but instead reflects concerns that using qualifications as a criterion for selection may create barriers to access and be viewed as contravening equal opportunity policies. A third possibility – that changes in skill needs had made test information more relevant to the selection process than in the past – was not supported by our interviews. The skills tested, whether academic-related (*e.g.* verbal and quantitative reasoning) or “soft” (*e.g.* teamwork, initiative) had been valued for many years, and were seen as important, but no more so than in the recent and moderately recent past.

- *Third*, that increased test use was a response to changes in the economy that have raised the cost of training and developing staff.

This hypothesis was not supported by the evidence. Cross-sectional data suggest that companies with heavy training expenditures also test more, and it has been suggested that this may be because the cost of staff development increases the value of all relevant selection information, and in particular justifies the additional cost of testing. Generalising to the effects of commonly described trends in the labour market – and especially the need of the ‘knowledge economy’ for continual learning and development – we hypothesised that test use might have increased in recent years alongside increased expenditures on, and attention to, staff training. We found no evidence to support this.

- *Fourth*, that increased test use was a response to the decreasing cost of testing relative to other methods of selection.

There was some support for this hypothesis in the data. The key change appeared to be less the up-front, ‘cover price’ of test purchase and training, and more the change to computer-based versions which can provide more or less instantaneous scoring and feedback. This greatly reduces one of the previous barriers to widespread usage: namely the need to have sufficient trained assessors to hand for results to feed into the selection process rapidly enough to be useful.

- *Fifth*, that increased test use was a largely defensive strategy, adopted in response to regulation and legislation.

This hypothesis was strongly supported by the data. The single most frequently volunteered reason for increases in testing was the need to have a selection process which would withstand challenges, including cases which actually reach tribunal stage. Testing is seen as playing an important role in this: the same motive underlies HR departments’ concern to have proper job and person specifications, and a clear link between job requirements and sources of evidence. Tests have two advantages in this context. They tend to align easily with the types of general skill or characteristic associated with the ‘competency frameworks’ which have become popular with organisations in recent years. (These frameworks are themselves seen as conducive to promoting access and equality of opportunity as well as reflecting a company’s substantive skill needs.) Tests also have been validated by established companies, who are able to provide assurances that the tests will not

have adverse impact on minority groups, and are therefore fully compatible with equal opportunity policies.

- *Sixth*, that increased test use was in large part a response to internal changes in organisations, and particularly the growth of ‘formalised’ human resources policies.

This hypothesis was supported by the data. The growth of human resource departments staffed by qualified professionals has been a response, in part, to the general regulatory climate. It also means that companies have, within them staff who, out of both professional conviction and departmental self-interest, tend to promote particular, formalised methods of selection. These staff (those who will typically be the only qualified test users in an organisation), are familiar with tests and tend to believe in their value. This further encourages test take-up.

Overall, these factors seem very likely (*ceteris paribus*) to encourage further growth in psychometric testing in the immediate future. No further changes in the external environment are needed for this to happen: simply the continuing influence of HR departments in a highly regulated labour market, and an environment where managers are increasingly comfortable with tests.

Overall, the surprising aspect of these findings is how little they have to do with developments in education or with changes in the demand for substantive skills. There are, however, clear policy implications for education (and training). The most widely used batteries of ability tests measure verbal and numerical reasoning, these are seen by companies as reflecting and measuring skills which they value and need; they are also, of course, skills central to the academic curriculum. At the same time, employers declining willingness to take formal note of qualifications calls into question the idea that education and training should be as qualification-driven as they have become over the last ten years. The unstated assumption behind much government policy in the UK seems to be that a course is worthless to the individual if no qualification is attached – hence the attaching of funding to qualifications rather than courses. This is clearly not the case if a course teaches skills which can be demonstrated during recruitment testing. Moreover, if it is the skills, rather than the diplomas, which are important, then it is important not to direct too much weight (and expenditure) towards assessment and certification at the expense of teaching and learning.

The Growth of Psychometric Testing for Selection

Why Has Test Use Increased, Will Growth Continue, and What Does This Mean for Education?

Alison Wolf and Andrew Jenkins

Introduction	1
1. The Policy Context: Governments' Skills Agenda	1
2. Employers' Use of Tests	5
3. Why Use Tests?	8
4. The Empirical Cross-Sectional Findings	17
5. Changes in Test Use	20
6. Conclusion	34
Appendix A: A Summary of Major Tests in Use	36
Appendix B: Sample Test Items	43
Appendix C:	46
References	54

The Centre for the Economics of Education is an independent research centre funded by the Department for Education and Skills. The views expressed in this work are those of the authors and do not necessarily reflect the views of the DfES. All errors and omissions remain the authors.

Acknowledgements

Alison Wolf and Andrew Jenkins are both members of the Centre for the Economics of Education, London School of Economics and the Institute of Education, University of London.

Introduction

This paper examines the current growth in psychometric testing for recruitment, and the degree to which it reflects dissatisfaction with formal qualifications in the general education and/or vocational sectors, changes in skill needs, or other non-education related factors. It is a companion piece to Jenkins (2001b) which reviews the evidence on levels of test use and confirms that there has been a significant increase and to Jenkins and Wolf (2002) which analyses the organisational and sectoral variables associated with higher or lower levels of test use. Its particular focus, however, is on the reasons for *change* in testing levels. Moreover, in focusing on reasons for change it draws heavily on evidence collected from case-studies of a wide range of public and private sector organisations – whereas the Jenkins and Wolf paper focuses largely on cross-sectional survey data.

1. The Policy Context: Governments' Skills Agenda

Contemporary education policy is highly preoccupied with how far the school curriculum, and to a lesser extent, university courses, provide employers with the precise skills they require of their workforce (see, for example, National Skills Task Force, 2000; Wolf, 1998). This activism co-exists with a rejection of any overt manpower planning. Consciously (if not unconsciously), modern Western governments believe this to be impractical and discredited in a dynamic economy. They do, however, believe not simply that high levels of education are important for a modern workforce, but also that the differences in the quality and quantity of educational provision that exist among developed countries are important to their individual economic success. (See, for example, Blunkett, 2000; 2001.) Moreover, in some countries (notably the UK and the USA), there is also a firm conviction, shared by politicians and the media, and by much of the business and professional class, that the country's current education system is performing at a highly inadequate level. This inadequacy is perceived to exist with respect to general education – notably command of written language and basic mathematics – and also in relation to the provision of particular specialist skills.

Politicians' preoccupation with the economic relevance of education, and its supposed inadequacy, is further strengthened by their belief that the economy itself has changed fundamentally in recent years, and that this must imply a demand for different (and probably additional) skills. In particular, the idea that we are entering a period of the 'knowledge

economy' and of accelerating technological change has passed beyond speculation to become received wisdom (Leadbetter, 1999; Keep and Mayhew, 1999).

Three influences thus come together to create acute political concern over the synergy between educational content and economic success. These are, first, the belief that detailed aspects of education policy can have a direct and substantial impact on economic success; second, in a good number of countries, a belief that the education system has, in recent years, been failing to meet the legitimate demands of employers for types and levels of skills among their employees, and third a belief that the nature of the links between curriculum and workplace needs has altered as technological change creates new demands from employers,. Most of the UK policy debate in this area is informed by government-funded surveys which ask employers directly both about their need/demand for skills and about the adequacy of new recruits' skills. Such surveys have the advantage of being direct and straightforward but they also have major limitations. As ample literature attests, responses to surveys are highly sensitive to the wording of the questions: since the latter are, in this case, likely to reflect policy-makers' concerns (such as the belief that there are skill shortages/educational inadequacies), there is a significant likelihood of self-fulfilling prophecy. (To give a couple of examples of how wording may affect responses, a UK study in the 1990s, designed to measure employer awareness of government initiatives, included in its list a completely fictional programme. Employers were asked to identify those programmes they had heard of, or were actively involved in and the fictional programme was chosen quite as often as most of the real ones. Again, respondents in a major UK longitudinal study were asked to identify which formal qualifications they had obtained while in school, and shown a list of possibilities. Many chose qualifications which, given their own age, and the date when the qualifications were first introduced, they could not possibly have obtained.)

Skill surveys of employers suffer from additional and more specific complications. Since they are typically completed (on paper or through a telephone survey) by a single individual in a personnel or human resources department, the respondent's knowledge of actual employee responsibilities and performance may be quite limited. Time series data might nonetheless in principle allow one to identify underlying trends with some confidence – since sources of error are the same on each occasion – but an additional complication is introduced by changes in the business cycle (Jenkins, 2001a; Robinson, 1996). Reported skill shortages rise, predictably and inevitably, as economies approach the peak of a cycle – indeed, if companies did not report substantial and increased skill shortages at such periods, it would suggest serious under-utilisation of resources. However, this makes it very difficult to

use employer reports of skill shortages to draw clear conclusions about underlying trends in the supply of skills by the education system.

An alternative source of evidence on skills requirements and their relationship to educational supply comes from longitudinal studies. US studies have included quite extensive testing of respondents, independently of and outside formal education and certification, and indicate (by analysing the determinants of income in successive cohorts) that reading, writing and, especially, maths skills are becoming increasingly important in the labour market (Pryor and Schaffer, 1999; Krueger, 2000). The UK data allow one to look more specifically at the labour market returns to *qualifications* which are of the same type, and at the same level, and therefore, presumably have much the same general 'signalling' value for employers; but which cover different material. From these one can make deductions about the relevance of curriculum content to individuals' labour market success. The work of Dolton and Vignoles on returns to A-levels is especially interesting here (Dolton and Vignoles, 1999; 2000) demonstrating significantly higher returns to A-level mathematics than to any other A-level subject.

Longitudinal data also have their limitations, however. Even the largest samples often provide rather few cases of a particular type (*e.g.* people who took a particular type of qualification, or entered a particular high-growth industry); particular cohorts will be unusual or atypical in certain respects (*e.g.* many of the BS70 cohort entered the labour market at an unusually difficult period for school leavers); returns to qualifications often reflect a curriculum which has long since changed. Analyses generally emphasise income/wage returns to different skills or qualifications. Generalising from this to skill needs in the economy implies that there is a direct and uniform connection between productivity and pay. However, this assumption may be more or less justified across different countries, sectors and time periods. (Leaving aside factors such as labour market regulation which affect the nature of the connection, it will simply be harder to detect links between curriculum content or skills on the one hand, and labour market requirements or productivity on the other, in countries with very compressed wage structures, such as much of Scandinavia, than it will in those with high dispersion, such as the US.)

Work on the demand for skills therefore needs to supplement large-scale survey data, however rich, with more detailed and qualitative work at enterprise level. In both this country and elsewhere, the greatest amount of research has addressed mathematics requirements, and their implications for the curriculum (see *e.g.* Sutherland and Wolf, 1995; Wolf and Silver, 1995; Noss and Hoyles, 1996; Wolf, 1997; Hoyles, Noss and Pozzi, 1999

and Hoyles *et al*, 2002). These studies have been detailed and ethnographic, and concerned with the quite complex ways in which school-taught skills are adapted and transmogrified in the workplace context.

Finally, firms' recruitment practices, and, in particular, changes in these over time, offer an alternative and valuable form of evidence which has been surprisingly neglected.¹ Indeed, the whole area of recruitment is remarkably under-researched and under-conceptualised, even though it absorbs a very high proportion of senior managers' time, and involves what, in most companies, is the single largest item of expenditure (salaries and wages). Moreover it is becoming increasingly critical as regulation makes it more and more difficult to reverse ill-judged hiring decisions by firing people.

In these circumstances it seems likely, *a priori*, that recruitment procedures will be informed by up-to-date skill demands. We do not wish to suggest that these practices are always optimal, in terms of meeting companies' objectives as effectively as possible, or that they reflect an accurate and complete analysis of skills needs. However, because they are very high-stakes for the firm, and because of the way they tend to evolve over time, they reflect input and decisions by a large number of experienced and senior personnel – including many line managers, and not just human resources/personnel departments² – with a strong incentive to develop appropriate mechanisms. They also (unlike complaints to government over the products of the school system) involve the commitment of substantial amounts of the companies' own funds, in the form of personnel and line management time, and also the use of specialist expertise, and specialist instruments – notably tests.

Recruitment practices thus offer a particularly promising arena for investigating changes in skill demands. If a large number of organisations spend significant sums of money identifying particular skills that were not a priority in the past, this may indicate

¹ The standard textbook on Personnel Economics (Lazear, 1995) contains one short chapter on recruitment, most of which is concerned with a particular (and interesting) comparison of the incentives offered by piecework rather than flat-rate payments in terms of the type of applicants attracted to a company as well as the effort they expend on the job. One of the effects/arguments for adopting piecework payment scheme is that it is likely also to sort applicants by ability. It does this by deterring those who suspect or know they are unlikely to be able to perform at a level which, on piecework terms, provides them with good returns while attracting the more able workers who have higher expected earnings under piecework than under salaried conditions in a comparable firm. However, the number of industries, and the number of jobs, in which piecework payment is applicable is small (and almost certainly shrinking): so that the relevance of this analysis to most recruitment processes, especially for higher-salary openings, is limited.

² The 2001 Recruitment Survey carried out by the Chartered Institute of Personnel and Development notes a general trend towards increasing line managers' control over the recruitment process, and general decentralisation of decision making, especially in non-public sector organisations (CIPD, 2001).

genuine and important changes in the skill mix which the economy demands. Similarly, if companies alter the way in which they evaluate and use formal qualifications, this may indicate either that the labour market relevance of formal educational curricula has changed, or that perceptions of educational standards have altered (for more or less well founded reasons).

Either way, there are clear policy implications for government agencies. In the case of growing use of formal (mostly psychometric) tests by the employer, the two possible explanations which have received most attention are first, that employers are increasingly interested in a set of skills which are not well catered for by formal education (and/or not attested to by formal qualifications) and second that employers have decreasing confidence in formal qualifications as evidence of required levels of skill. Both these positions have been advanced with confidence by employers' organisations and government agencies. The first has been argued most actively by the CBI, when advocating the formal accreditation of such 'key skills' or 'core skills' as 'problem solving' or 'working with people' (CBI, 1989 and 1995): a policy of developing such accreditation was pursued very actively by the former National Council for Vocational Qualifications in the 1990s, and, far less actively, by the DfES and the Qualifications and Curriculum Authority, successor authority to NCVQ. The latter charge has been brought by both the CBI and the Institute of Directors (IoD, 1994). Our research investigated how far either of these factors actually explains rising test use, as well as the importance of a number of other possible explanations which are, *a priori*, equally plausible (but without the same implications for curriculum content and education policy).

2. Employers' Use of Tests

The major psychological testing agencies in the UK and their products are summarised in Appendix A, which also explains the use of 'level A' and 'level B' certificates for test users: the former basically covering ability and the latter personality tests. Tests currently in widespread use can be categorised as:

- general ability tests (usually critical reasoning, verbal reasoning, numerical reasoning, plus some spatial, fine dexterity and mechanical ability tests)

- personality questionnaires and profiles, generally built around the ‘big five’,³ although these domains are often broken down further, typically into 16 or 32 factors (*e.g.* 16PF, OPQ32)
- literacy and numeracy tests
- specific skill and competency tests (*e.g.* clerical)

A large number of skill/competency tests are also generally available, from a range of sources other than the large testing companies – notably the ubiquitous typing test, but also a range of literacy and numeracy tests for which no special licence is necessary, and which can be copied freely. A good number of companies create their own skill tests for internal use: others commission highly specialised tests, developed by external consultants for particular, specialised situations and jobs. In addition, specific tests of integrity and of motivation (rather than broader personality profiles) are available. These appear to be widely used in North America, but not in the UK, and are not discussed specifically in this paper.

There has been a considerable amount of attention paid to levels of test use by companies but the data are generally drawn from rather small and non-representative samples. Overall, however, the sources (summarised in Jenkins, 2001a) are consistent in indicating that use has been growing steadily and that test use for recruitment is now very common and supports a substantial commercial sector of test creation and processing. Data from the CIPD (Chartered Institute of Personnel and Development) also show increasing use of tests *within* companies (CIPD, 2001). However, very little of this work has attempted to explain differences in test use in theoretical terms. The largest recent survey of test use is the Workplace Employee Relations Study, according to which, in 1998, 19.4% of all establishments with 10+ employees were using some form of personality test, and 47.9% were using some form of competency test for recruitment.⁴ Among establishments with more than 100 employees, these figures rose to 39.2% and 63.2% respectively. As part of our research, we analysed the WERS data in some depth, in order to establish how far levels of

³ Personality traits are most commonly examined in terms of the ‘big five’ dimensions of *agreeableness* (how far someone is genuinely consensual and concerned with other people’s feelings as compared to results-driven); *adjustment/neuroticism* (how far people value order and control, are highly structured and prone to plan to the last detail); *extroversion/introversion*; *conscientiousness*; and *openness to experience* (including the degree to which people are conservative and like implementing plans rather than creative, independent and resentful of structure).

⁴ See Appendix C for the precise wording of the WERS questions.

test use could be explained by other organisational and sectoral variables (see Jenkins and Wolf, 2002).

Another data source is the annual recruitment survey conducted by the Chartered Institute of Personnel and Development, which, in 2001 (CIPD 2001), enabled us to insert some very specific questions about test use.⁵ The CIPD survey covered a much smaller number and range of companies (253 organisations all employing 50+ workers) and also fewer organisational variables than did WERS. It did, however, ask much more detailed questions about test use and also included a number of questions about the organisation of recruitment practices not covered by WERS: analysis of these is also reported in Jenkins and Wolf. The results of the survey indicated similar levels of use to those reported in WERS – 60.1% for tests of specific skills (equivalent to competency tests); and 40.7% for personality. In addition, 54.5% reported using general ability tests and 44.6% literacy and numeracy tests. The CIPD was also the main source of contacts for qualitative work on test use, and, specifically, on changes in test use which we undertook alongside the survey data analysis, and which is the main focus of this paper. Through their Recruitment Forum (enrolling CIPD members who are professionally involved in recruitment) we were able to identify individuals in a range of organisations who were experienced test users and willing to be interviewed in depth about current and changing patterns of test usage. Through our initial CIPD contacts, a number of further such case studies were also made possible. In all, we conducted 29 interviews (with one or more respondents), some of which were with specialist companies or independent consultants who provided recruitment services, including testing, to a considerable number of different clients (almost all sizeable private sector companies). In total, the case study data covered current testing practice, and its recent evolution, in 36 private sector companies; 12 public sector or non-profit organisations, and 5 organisations which were either public-private collaborations, or recently privatised companies in which practices still reflected (in part) their public sector inheritance. All of these were active users of tests; most, but not all, had increased test use in recent years.

⁵ See Appendix C.

3. Why Use Tests?

What might account not simply for a high level of test use, but, more specifically, for the increase? To understand this, it is useful to start with the way in which the decision to use (or not use) tests is actually made within a company with respect to a particular hire.

First of all, as our case studies make clear, the decision is in most cases dependent on other decisions which were made well in advance of the particular recruitment/hiring case in point. With the exception of large graduate recruitment programmes and a few major recruiting and redeployment efforts (associated with new companies, major restructuring or with public sector programmes opening up non-traditional avenues for promotion) managers and personnel staff alike are largely concerned with hiring either single individuals or small groups, usually within a tight time-span.

If a test is to be used in such circumstances it will, in almost every case, be one which is already available in the company or through a 'retained' outside consultant: in other words, the decision to test is conditional on other decisions, which have resulted in tests being available as and when required. Moreover, these decisions will always involve the Human Resources or Personnel department, even if they are not the only ones involved.

Because of the costs of tests and, for in-house administration, of training staff (initially and when trained staff leave), the decision to purchase is described by respondents as one which has been the subject of considerable deliberation: this is not something which, in cost terms, can be seen as equivalent to a minor item in the stationery bill.⁶ A few tests, notably typing/word processing tests, can be bought very simply in an over-the-counter fashion, but many commercially produced tests are only available under restricted circumstances (see Appendix A). SHL, the company which dominates the UK market for both personality and general ability tests (e.g. verbal reasoning, mathematical reasoning) requires test users to have undergone dedicated training and will only sell to companies with registered trained users. Where 'restricted' and commercially developed tests were

⁶ Many tests are administered as part of an assessment centre, or by in-house personnel departments who do not price their time overtly. However, psychometric tests, where licenses are required, are also the ones which are most likely to be out-sourced to recruitment consultants and where we can therefore obtain some quite good estimates of the costs of testing. The minimum quoted to us was between £60 and £80 per test administration (*i.e.* including feedback/interpretative report). In many cases, more than one test is used, and an average figure of £500 per candidate was cited by one consultant who works for a range of large companies. He emphasised, however, that this had to be put in the context of what he calculated to be a

administered by company staff, these were, in our study, invariably located in central HR departments. In cases where testing was out-sourced to an independent specialist, it was again personnel officers who organised this, and they (if anyone) who monitored the extent to which there were adequate trained staff available to meet demand for test use.⁷

Some companies use tests written in-house by previous or current staff (*e.g.* in-tray exercises) and others set group exercises as part of an assessment centre. Approximately half the organisations in the sample report some activity of this type, but this was, in each case, a small-scale activity compared to the use of purchased, commercial tests: nor were any of the in-house tests which we encountered validated in a formal way. In a few sectors there are also commercially available, highly sector-specific tests, such as scenario-based exercises for the hotel industry. (These are used by large chains and developed and marketed by a specialist US company.) Some of our case study companies had paid for the development of ‘bespoke’ tests which were designed specifically for them by external consultants. All these involved cases where very large numbers of people were hired, either regularly, or because of a start-up or restructuring. They were also all cases where there were serious concerns about the danger of hiring inadequate staff (especially for safety reasons), or where redeployment decisions were being made which were high-stakes for large numbers of current employees (and where the unions were often actively engaged). For all these types of tests, the rule was again that they were organised and administered by HR specialists and not by line managers.

To summarise: psychometric test use in recruitment presupposes either that the central personnel/HR department has equipped itself with tests and trained staff, or that the organisation has an on-going contract with an outside agency to provide testing services, and can call on these at short notice – something that will also be administered through central HRD. What might lead to this happening? And what, in particular might lead to it happening more often in recent years than ever before? Six possible scenarios or hypotheses present themselves – by no means mutually exclusive.

- **First**, the expectation that (senior) managers will in the near future ask for tests. In other words, the ‘prior’ decision to purchase tests may itself be a function of expected

total of, on average, £15,000 per short-listed candidate in cases where national advertisements, major commitments of senior staff time, *etc* are involved.

⁷ In fact, a surprising number of organisations did not seem to plan ahead in any way to ensure the future availability of trained staff.

demand from outside the HR function. HR staff respond to this expectation because they are in a service relationship to senior managers. The latter directly or indirectly control their budgets, positions *etc* and they therefore wish to be in a position to accede to future requests. In this scenario, HR staff may also be assumed to be either neutral or positive regarding the effectiveness of tests (in relation to managers' beliefs about what tests can offer in making 'good' recruitment decisions).

If managers start to demand tests more often, this in itself may reflect a number of factors: for example, decreasing faith in the standards of school leavers, or in the qualification system (vocational as well as academic); growing awareness that interviews are fallible, and that there is a substantial body of evidence to this effect; or, more simply, a belief in tests (which more and more managers will themselves have experienced – and, since they got the job, will tend to see as valid!) In any of these cases, the result will be that **increased test use is a direct response to increased demand from managers** (Hypothesis A).

- **Second** the belief that tests can indeed provide valuable and reliable information about candidates and add value both in identifying good/the best recruits and inadequate/the poorest ones and that the company should be encouraged to use them. This set of beliefs may be translated into action either because the staff concerned identify their own interests with the company's longer-term success (and wish it to select the best recruits possible); and/or because they believe that the perceived efficacy of tests will rebound to the HR department's credit and their own immediate benefit.

The nature of tests, the information they provide, and the validation techniques used by companies (and the resulting validity coefficients) have changed very little in the last two decades, so increased use is unlikely to reflect changes in the 'absolute' perceived accuracy of the information. However, the perceived relative value of that information may alter. For example, HR staff may come to regard qualifications as less accurate sources of information than in the past; or as providing evidence of skills which is less useful than it used to be. Or they may feel that references have become less detailed or more anodyne, and therefore that test information has greater value compared to information from testimonials than was the case in the past. Or shifts in

the skill mix demanded may make the domains addressed by personality, ability or aptitude testing more important than they were in the past. (Some recent discussions (*e.g.* CBI 1995, SCANS 1991; see also Financial Times 11.5.2002) imply that ‘soft’ skills such as team-working are more important in the modern workplace than they were in the past, and that employers will therefore place added value on evidence that applicants possess them.)⁸

Alternatively, test outcomes may be more valued because *any* information that helps people make the ‘right’ decision has become more valuable. The most likely reason for this would be an increase in the potential cost/difficulty to the company of getting rid of unsatisfactory workers, and, therefore, the generally increased importance of getting initial selection decisions right. Constraints on employee dismissal may vary both between companies (with one obvious proxy for this being union penetration) and between employees (depending on the nature of the contract awarded) but may also change over time within a whole country as a consequence of *e.g.* employment legislation. In that situation one might expect HR departments to advocate test use more often because the cost of incorrect hiring decisions has increased. In both these cases, we would find that **increased test use is a response to increases in the net, or relative, value of the information provided** (Hypothesis B).

- **Third**, employees who will receive costly training in the early period of their employment will, *ceteris paribus*, justify higher levels of recruitment expenditure. This, too, may have an impact on levels of test use. If changes in the structure of employment (‘the knowledge society’) or in firms’ HR practices, mean that more recruits are involved in costly training, this too may translate into more frequent decisions to use tests during recruitment.. In this scenario, **increased test use is a response to increased training and development costs for staff** (Hypothesis C).

⁸ In the 2002 survey of graduate employers carried out by the Association of Graduate Recruiters and the Financial Times, interpersonal skills were consistently given the highest rating among factors determining how successful a graduate will be in their long term career. It should be noted, however, that the respondents were focusing on graduate’ success once in employment: *i.e.* after they had cleared the selection hurdle. FT ‘Universities’ supplement, published 11.5.02.

- *Fourth*, tests may be purchased in the belief that their use will reduce the time demands on managers and the costs associated with selection and decision making. Provided that the tests are also positive, neutral or insignificantly negative in their impact on the quality of the recruits selected, this will again rebound to the HR Department's credit. (NB How much money is saved obviously depends on the cost of managers' time as well as the cost of HR staff time during test administration but the balance is likely to be positive. See footnote 6.) If this is a major underlying factor in levels of test use, increases in purchase and provision may reflect changes in the relative costs of test use and manager time: test use may be cheaper or managers' time more expensive. In other words, in this scenario, **increasing test use is the result of a changing balance between test administration costs and costs of other recruitment methods** (Hypothesis D).
- *Fifth*, HR departments (and senior managers) may believe that test use is a helpful defensive strategy for a company to adopt, in anticipation of possible future legal and regulatory challenges to its recruitment procedures. In recent decades, companies' recruitment procedures, in both the USA and Europe, have become increasingly subject both to regulatory constraints and (especially in the US) to legal challenge (see Wing and Gifford 1993 for a discussion of landmark US court decisions). A major objective of regulation, in each case, is to safeguard equal opportunities for all applicants, and particularly those from recognised minorities (ethnic groups, the handicapped, women). For example, one high-profile case in the UK (one with which most personnel professionals are well acquainted) involved recruitment/selection tests used by British Rail. These were judged by a tribunal to discriminate against non-native English speakers in a way that could not be justified by the requirements of the jobs in question.

If a company's procedures and decisions are challenged, the potential value of having used tests at recruitment will vary with the context. In a number of highly-publicised US test cases, company procedures were struck down because the recruitment tests used were judged inappropriate to the vacancies concerned, and thus actually or potentially discriminatory. However, in other cases, companies have been criticised for not using tests. The general position is now that a company which uses tests is

thereby more likely to have its procedures judged acceptable *provided* the tests used are demonstrably related to job requirements, and have been developed in line with professional demands and standards.⁹ In the absence of such conditions, test use will not provide a strong defence of the appointments made, and may actually be held against the employer.

Changes in the regulatory environment are thus likely to lead to changes in levels of test use, but in which direction - higher or lower use - is not clear *a priori*. If human resources departments in a given country perceive test use to be a useful defence in a potential tribunal case involving charges of unfair recruitment and/or promotion decisions, then test use will increase when new laws and regulations are brought in relating to these activities, and also when employees' access to tribunals and courts is increased. However, if use of any given test is seen as something which it can be extremely difficult to defend and justify, then regulatory and statutory changes may actually tend to decrease levels of use. The relevant hypothesis here (Hypothesis E) is that **increased test use is a defensive strategy adopted in response to regulatory and legislative changes.**

- *Sixth*, and finally, it may be that a major reason for HR Departments' purchase of psychometric and other tests is the belief that test use will strengthen their position and influence within the organisation. This is unlikely to be articulated in quite this way, publicly, or even among HR staff themselves. More probably, it will be phrased in terms of how best to achieve high quality results. It may nonetheless be quite important in explaining test take-up. There is a large literature on the way in which professional and semi-professional groups consistently attempt to use their (genuine) professional knowledge to strengthen their overall influence in organisations and in society, and increase their financial and other rewards (see *e.g.* Cohen and Pfeffer, 1986; Dobbin *et al*, 1988; Baron *et al*, 1986; 1988). Group members (and, indeed, individual humans generally) find it natural to believe that the good of the collective and their own-self-interest are intimately entwined. It is not just a case of believing

⁹ Embodied in the Codes of Fair Testing Practices issued by organisations such as the National Council for Measurement in Education.

that what is good for General Motors is good for America. Employees tend to be equally convinced that what is good for them is good for General Motors.

From an organisational perspective, it is clear that high levels of test use will tend to benefit HR departments. As discussed earlier, psychometric tests can be used and interpreted only by trained staff, who will typically be within an HR or personnel department. Any recruitment or promotion process which involves testing will therefore have to involve that department closely, and in a way which gives its members an important substantive input into the decision-making process. Moreover, to the degree that tests are accepted, used and valued, the status accorded to professional HR expertise will tend to rise: since HR staff not only can use tests currently held by the organisation, but also know about alternatives, and about the relative strengths, weaknesses and defensibility of each.

It follows that, over time, HR departments will tend to promote test use, and that, *for a given degree of formalisation of the "HR function"*, one would predict a (slowly) rising level of test use without any additional factors coming into play. More rapid increases would follow from an increase in the number of firms with formal HR departments, and from the number of formally trained and qualified HR staff found in such roles. In other words, we would find that **increased test use is the result of increased formalisation of Human Resources activities and departments** (Hypothesis F).

All of these factors may apply in a general way to the decision whether or not to purchase tests (and, for many tests, training which permits staff to administer them). In the following sections we examine how far the evidence supports any, or all, of the hypotheses advanced. In addition, it is important to highlight a number of factors that may affect which types of test organisations do or do not adopt, and when, within an organisation, they are used.

First, tests which are cheap, easy to administer, and seen as potentially informative *may* be purchased more often than tests that are potentially at least as valuable, but a lot more expensive. It may seem self-evident that companies will always go for the cheaper option but actually this is not the case. Cheaper tests that can be used without special training may not be as attractive to HR departments as those which require trained users. This may be partly

an issue of professional self-interest but may also reflect well-founded professional concerns that cheaper tests will not be as well-developed or as valid. They will not, in consequence, stand up as well to hostile scrutiny; and they are also likely to be of substantially lower quality, and the classic false economy.

Second, tests are likely to be valued more, and easier to ‘sell’ to managers and budget holders, if they promise additional information in cases where (a) appointments are very high-stakes, and (b) when appointing to jobs for which it is very hard to obtain objective and reliable evidence about candidates’ suitability. From an employer’s point of view, recruitment will merit more attention and expenditure the greater the potential benefit/cost to the organisation of the employee’s performance or *the greater the potential for large losses* (directly or through adverse publicity) associated with poor performance. This former situation will generally be at least partly reflected in an employee’s salary – so, other things being equal, more attention will be paid to recruiting senior managers than low-grade clerical workers (and the organisation will be willing to pay more for the recruitment process.). However, workers such as train drivers and signalmen do not earn a great deal. Recruitment to these jobs is nonetheless taken extremely seriously, and involves higher levels of expenditure by the relevant companies than does recruitment to many higher-paying positions.¹⁰

Third, the nature of the tests used by companies is likely to be very different depending on the selection context. “High stakes” appointments to relatively low-level jobs requiring very specific skills are likely to involve quite specific competency-type tests to ensure that recruits either already have the particular skills, or have the potential to acquire that same highly specialised sub-set. Highly paid professionals, who are expected to carry out a wide range of less well specified activities, are likely to be tested using a combination of tests tied to professional competencies and skills and more general tests (*e.g.* verbal and quantitative reasoning. See Appendix B for sample questions). Highly paid generalists, such as managers, are less likely to be given competency-type tests, and very likely to be given a range of general tests, including personality tests (see Appendix B). This wide range of testing is related not just to the fact that generalist skills will be less well attested by formal

¹⁰ Companies’ level of risk aversion, and the probabilities they subjectively assign to possible worst-case scenarios, are likely to change with the political and economic climate, and will also affect the way they recruit for positions such as these.

qualifications than is the case with professional (and craft) skills, but also to what economists call the ‘principal agent’ problem.

Recent work in economics has elaborated the dilemma faced by almost any employer who wishes to maximise the effort expended by employees. The employee, however, is first and foremost interested in his or her own private ends, which may or may not be the same as the employer’s. This is a more general statement of the problem alluded to earlier in the context of HR departments where we pointed out that HR professionals’ activities need to be seen in terms of their own interests and objectives and not just those of ‘the organisation’.

In situations where there is such a problem (see *e.g.* Dixit, 2000; Prendergast, 1999), the employer may respond by instituting high levels of oversight, and by measuring precisely what the employee does and produces. But measurement is itself generally very costly and in many cases it is effectively impossible. Measurement is also very problematic when workers are expected to perform a number of different tasks, and when some behaviours or outputs are more easily measured than others. Attempts to measure employee behaviour tend, naturally, to focus on the easily/cheaply measured (because measurable) outcomes. In so doing they distort employees’ behaviour, as they, in turn, focus on the measured tasks to the detriment of others which may be more important. Much of the economic literature on incentives focuses on bureaucracies, and, specifically, the greater likelihood that public-sector bureaucrats will be multi-taskers whom it is difficult to monitor and measure than will be the case in the private sector. However, many private sector employees are also multi-tasking and/or difficult to monitor.

As a general rule, this is especially true of managers, compared not only to skilled manual and clerical workers, but also professional and technical workers. We can predict that this will manifest itself in the nature of the recruitment process. Employers will want to know not only about individuals’ levels of actual and potential skill, but also will value any information that indicates whether they are better or worse suited to a particular job in terms of personality traits (*e.g.* extroversion/introversion); and are likely to make a long-term commitment to the company, rather than concentrating on activities with high visibility outside the company and short-term individual pay-offs. Psychometric testing of attitudes, personality *etc.*, are thus much more likely to be attractive in the context of high-stakes hiring decisions to jobs where close supervision is difficult; and this will hold true irrespective of other factors which may tend to raise or lower testing levels overall.

4. The Empirical Cross-Sectional Findings

In this section, before turning to the evidence from the case studies, and our hypotheses about why test use overall has increased, we summarise the evidence from the cross-sectional analyses carried out with WERS and CIPD recruitment survey data and reported in detail in Jenkins and Wolf (2002). For both surveys we had available both data on *current* test use (see Appendix C for the survey questions) and information on a considerable number of other variables relating to the organisation, some of which were highly relevant to our concerns. Equally, we were without information on a number of other key variables, notably anything relating to the use of qualifications or references in the recruitment process, or any questions relating directly to changes in skill requirements.

The following main findings are especially relevant here, because of their direct bearing on one or other of the hypotheses advanced above. WERS findings can generally be taken to be more robust, since the sample size is far larger but the CIPD data include information on a number of additional factors as well as being broadly consistent with those from WERS.

Key findings from the WERS analyses are as follows (the econometric results are summarised in Appendix C):

- a. Companies who engage in high levels of formal off-the-job training and workplace development policies are significantly more likely to use personality tests for recruitment. However, there is no such relationship for use of competency tests.
- b. Workplaces reporting managerial and professional vacancies were also more likely to use personality tests for recruitment than others, and those reporting vacancies for routine unskilled manual occupations less so.
- c. Levels of personality and competency test use are both clearly associated with the extent to which workplaces practise a range of formal policies to promote/ensure equal opportunities in the workplace. In addition, use of competency tests is directly associated with the presence of formal grievance procedures and whether there have been any recent employee complaints to a tribunal.
- d. Public sector organisations are significantly less likely to use personality tests than are private sector ones – there is no such relationship for competency tests.

- e. Companies with clerical and technical vacancies (not managerial and professional, or sales) are appreciably more likely to report competency test use.
- f. Use of tests for recruitment is positively associated with an emphasis on the importance of motivation being shown at the time of recruitment.¹¹

The much smaller CIPD survey distinguished four categories of test. The econometric results are again summarised in Appendix C. Overall, it appears that:

- a. Use of *personality tests* is more common when companies have formal equal opportunities policies (notably, policies intended to promote diversity in the workforce); when they involve external consultants in recruitment; and when they are recruiting with an eye to the long term rather than for a specific current job. The extent of use (in terms of number of occupational categories involved, and whether or not use has changed in the last two years) was positively associated with the same factors, plus organisation size.
- b. *General ability tests* are more likely to be used when local personnel staff (rather than staff from central office or external consultants) are involved; when there is an emphasis on the long term; and in enterprises which devote a particularly high level of effort to recruitment and evaluate their recruitment practices. The extent of use (measured by how many different categories of staff were tested) was significantly related to being in the public sector and to evaluation of recruitment practices.
- c. *Literacy and numeracy tests* are again found more often when there is a commitment to the long term when recruiting and also when there are formal measures to promote equal opportunities and diversity. Extent of use was related to these factors, and to organisation size and whether or not recruitment practices are evaluated.
- d. Few variables were, in this sample, clearly associated with use of '*tests of specific skills*' at recruitment. The only one that was clearly and strongly related to use was the company's economic sector, with public sector organisations more likely to use them. *Extent* of use was also significantly higher in the public sector: as was taking a long-term perspective.¹²

¹¹ This finding is not discussed in detail in the companion Jenkins and Wolf paper, but is consistent with the argument that tests may be used by companies seeking to identify employees who will identify with the organisation and share its objectives.

¹² It might seem obvious that test use would be associated with whether or not organisations had formal methods of recruitment rather than relying on informal (word of mouth) recruitment. This is true, but in the

None of these findings, whether from WERS or CIPD, rules out the possibility that qualification and skill factors are also important in explaining test uptake. They do, nonetheless, suggest that they are unlikely to be the *sole* influence. The pattern of association between test use and variables relating to HR practices is strong and consistent enough to indicate that these are almost certainly a major, independent influence.

The cross-sectional data are also interesting for what they do *not* show. Information on union recognition was available in WERS but was not significantly related to test uptake. The degree to which line managers are directly associated with recruitment is covered by the CIPD data and is not related to use of any sort of test. This again suggests that the proximate causes of change in test use are to be found within HR departments, rather than in, for example, changes in levels of union recognition or activity, or changing managerial attitudes.

Finally, there is also a cluster of labour market variables which, *a priori*, seemed likely to affect whether or not companies are inclined to spend money on intensive testing: notably the degree to which it is hard to fire people who turn out to be unsatisfactory, which means that ‘wrong’ decisions are more expensive (and test data more valuable), and the degree to which recruitment is solely or largely internal. Unfortunately, the available data did not enable us to examine these directly or satisfactorily. The only proxies for these factors which we could construct in the WERS data (see Jenkins and Wolf, 2002) turned out to be insignificant, but they were also very approximate measures indeed¹³ and this may well account for the lack of relationship.

To summarise the evidence from these cross-sectional data sets, in relation to the major hypotheses advanced earlier about increases in testing, we find:

- no evidence in support of Hypothesis A: that high levels of use, and increases in use, reflect direct line manager demand for test evidence as against other sources.
- no opportunity, in these data, to examine Hypothesis B: that there have been changes in the relative value of (valid) evidence from testing, in relation to its cost, because of

CIPD data the relationship appears (statistically) significant only for the use of general ability and literacy/numeracy tests.

¹³ For ease of firing, we used a dummy variable taking the value 1 if guaranteed job security or non-compulsory redundancy policies exist for at least one occupational group within the organisation (otherwise 0). Union recognition also provides a proxy for ease or difficulty of firing but, as already noted, was not significantly related to test use.

either the decreasing value of other evidence, or the increased value of any valid evidence which reduces the likelihood of increasingly costly incorrect decisions.

- some evidence for Hypothesis C: that test use, and increased levels of use, are associated with the cost to companies of (competent) staff, and that, to the degree that training and staff development costs rise, so test use tends to increase.
- no opportunity, in these data, to examine Hypothesis D: that test use levels, and changes in these, are a direct function of the relative cost of managers' time compared to test administration costs.
- some evidence, in these data, for Hypothesis E: that test levels and increases in these are a direct response to perceived risks associated with equal opportunity and anti-discrimination legislation, and with the way tribunals operate.
- considerable evidence in favour of Hypothesis F: that levels of use, and increases in these, are a result of increases in the formalisation of the "HR function" in organisations.

The next section of the paper looks at the evidence from the case studies in relation to these same hypotheses. As noted earlier, the great advantage of the case study evidence is that it allowed us to investigate in depth the way test use had developed over time in respondents' organisations. However, these studies were not of a representative (let alone a random) sample of UK organisations: in every case, responses relate to organisations which used, or had used tests, and in which relatively senior staff agreed – indeed, mostly volunteered - to talk at some length about their practices. While the case studies thus provide invaluable insights into trends, and the reasons for test adoption, or rejection, they cannot be used as a basis for quantitative estimates of practice across the workplace.

5. Changes in Test Use

This section summarises findings from the 53 organisations on which test use information was collected. We requested informants to describe changes over as much of the last decade as they could; but in over half the cases, they had personally arrived less than ten years ago, and no organisations appeared to have kept records or documentation of policies stretching that far back. We therefore had to rely on individual knowledge, which, in 85% of cases was

felt, by the respondents, to be reliable for at least the last five years. While repeated visits to the same organisations over a five or ten year period would obviously provide a more robust evidence base, the data collected are nonetheless more detailed, and more explicitly focused on changes in test use, than any other collected up to now.

All of the long-standing private sector companies had increased test use within the last five years. Just under a quarter reported increased personality testing but not increases in use of other types; just over 10% had increased other forms of testing but not personality testing (including one company which had done so for a while, but recently reversed this, dropping all use of OPQ); and the remainder (*i.e.* two thirds) had increased their use of both personality and other forms of testing. Among public-sector organisations, some had increased use and some had maintained a fairly constant (though substantial) level over the period. Privatised organisations (successors to public sector organisations) showed the most mixed pattern: some had increased use, the others reported lower levels than during (parts of) their public sector existence. However, the number of organisations in this latter category was small ($n=5$).¹⁴ All firms running graduate recruitment schemes made heavy use of tests but always as part of a larger assessment centre/selection centre.

No-one used tests as the sole basis for short-listing, though some instances of this practice were described to us (disapprovingly) by respondents. Formal tests were usually used after short-listing, but in a few cases were part of the short-listing procedure.

Since the focus of our enquiry was increased use of commercially produced tests, our case studies examined only organisations that used or had used these. The vast majority of users (over 90%) used tests that require users to be trained and licensed, with SHL dominating the market in both personality and aptitude/ability testing. A few used other commercially produced tests (*e.g.* Psytech) which are available for immediate use without such training. Two thirds also made use of other tests that were designed in-house, or had been obtained/bought at some previous period and were not restricted in their use. The latter would typically be literacy and numeracy-related; the former were typically group exercises, in-tray exercises or other work-sample exercises for use as part of an assessment centre. Six organisations (just over 10%) had paid for the development, by expert outside companies, of bespoke tests for use with particular groups of recruits.

¹⁴ Strictly, four privatised and one newish public-private collaboration.

As already noted, licensed users within organisations were found in Personnel or HR departments: this meant that these tests were kept there (usually now, in computer-based form). The number of individuals who had trained to administer and interpret tests was generally very small, though more might be allowed to administer. Only 11 organisations had an explicit policy of maintaining numbers of licensed practitioners over time; and in almost half the cases where testing (using licensed tests) was an in-house concern, only one person had the relevant full level A and B training. However, a substantial number of private-sector companies, including some very large ones, out-source a large part of recruitment activity, including testing. Nine of our respondents were independent consultants who worked for a range of private companies, usually as part of a company rather than as self-employed sole practitioners. From the company's point of view, this ensures stable access to licensed professional staff.

A. Is increased test use a response to increased managerial demand?

One simple explanation for increased test use in organisations would be increased management/employer demand (Hypothesis A). For example, if there is indeed a widespread perception that school standards, and qualification standards, have declined, as employer organisations report, this might be translated into growing demand from managers for HR departments to test candidates' language and mathematics skills directly. Or changing, or rising, skill requirements might result in managers' perception that current recruitment procedures were inadequate, and needed to address particular skills more directly. A response to this might be to demand more qualifications, or to demand direct skill testing.

In fact, we found very little evidence of test use increasing in response to explicit management demand. In two organisations we were told that senior managers had, independently, encouraged the use of tests: in one of these, the policy came directly from the chief executive and reflected his conviction that mathematical skills were critically important for all employees from mid-level supervisory upwards, and in the other (where tests were already in use), managers were aware of research which indicated that interviews were a highly fallible basis for recruitment, and requested more testing.¹⁵ This is not to deny the critical role senior managers' views play, however large the organisation: one respondent

spoke for many when describing how ‘We’ve used methods for which there was no good valid evidence at all if a really senior manager was set on it’; while others emphasised the need for top management to be supportive, or at least accepting, from the start. (‘Our new Chief Executive came out top on the tests, so of course he approves of them’, remarked one personnel officer in a large public sector organisation which is highly committed to test use.) But the initiative rarely came from above.

Instead, in the vast majority of cases, test adoption and increases in test use, were initiated by the HRD staff. In the latter case, the responses of most managers initially often ranged from mild to very strong scepticism:¹⁶ but in every case study site where our respondents had been directly involved in introducing or increasing testing, they reported a rapid move towards approval and enthusiasm by at least a sizeable proportion of managers. Of course, these responses do not preclude there being significant changes in managers’ skill needs, or perceptions of recruits’ capabilities: indeed, there may well be, and this may predispose them to look favourably on recruitment innovations which promise to increase the quality and appropriateness of hiring decisions. However, the increase in test use does not appear to be a result of direct requests by managers, or direct criticism by them of the value of qualifications as a basis for recruitment.

B. Is increased test use a reflection of perceived increases in its net value?

As discussed above (Section 3) the nature and content of tests has altered very little in recent years. So if increased use reflects greater value being placed on the substantive results of testing – *i.e.* what they tell you about the actual capacities and skills of an individual – this implies that their net value has changed relative to either other sources or to the likely costs of making a wrong decision. The former might be the result of other sources of evidence becoming less valuable (*e.g.* if qualifications are seen as less informative than in the past), or of the particular information provided by tests becoming more relevant because of changes in the skill mix of occupations. The second would reflect changes in the organisation’s environment.

¹⁵ In a third case, the senior manager concerned had headed the HR department previously, so the situation was rather different.

¹⁶ Survey evidence from the late 1970s and early 1980s suggests that managerial scepticism was, at that period, a serious barrier to uptake. See Jenkins (2000a).

All recruitment processes are approximate and imperfect: they carry an x% chance of not choosing the best candidate(s) or of hiring people who are actually deficient in key requirements. However, if the costs of making the wrong decision increase, a level of risk which was acceptable in the past may become significantly less acceptable, and companies will wish to pursue additional and costly procedures which promise a reduction in that risk level.

The precondition for such responses is, of course, the belief that tests are indeed a valid source of evidence and certainly case study interviewees believed them to be such. A typical comment on personality profiling was ‘The new OPQ ...is remarkably accurate. I’ve given feedback ...after only 40 minutes, people have said “That’s me!” No-one has ever complained about the opposite.’ Others noted that ‘I really need to know if someone can read a piece of information and draw a logical inference, and the test tells me’, and that ‘aptitude tests really can pick up someone who is very bright and has potential’. In one organisation, running large assessment centres involving extensive test use, one senior interviewee emphasised the need for high quality tests and expert assessors but argued that ‘the success of tests as predictors has been proven — they are at least 60% more accurate than other methods as indicators of potential — they can save you as much as two years of on-job training’. For the most part, respondents also emphasised the importance of treating them as just one piece of evidence among many; of interpretation by a trained and experienced person; and of using high quality tests. But with the exception of two sceptics who questioned the value of personality testing,¹⁷ the validity of the tests was accepted or, more often, enthusiastically endorsed on the basis of personal experience, by every one of our interviewees.

Our evidence suggests that a rise in the perceived relative value of test-based evidence is indeed one of the reasons for the growth in testing (though probably not the most important.) It is associated with a markedly reduced emphasis on qualifications as a basis of recruitment, and also, in some cases, with shifts in the perceived risks of making wrong appointment decisions: changes in the skills required by organisations, conversely, did not appear to be very important.

¹⁷ The sceptics did not consider that the results of such tests were necessarily wrong, but did query their usefulness, noting that the vast majority of people scored down the middle, so that their usefulness for discriminating among people was actually very limited; moreover, they might be used as a trigger for questioning and ultimately rejecting the application of someone with particular, ‘extreme’ scores without there being good evidence to indicate that this was appropriate.

The downgrading of qualifications. One of the most striking findings from our case studies was how many organisations explicitly reject the use of academic or vocational qualifications as a recruitment criterion and how many others are trying to reduce their importance. Leaving aside graduate recruitment schemes, and recruitment to specific professional jobs, we found that, in over a third of cases, qualifications were not viewed as a legitimate criterion for use in making recruitment decisions; and a further eight respondents (including some with multiple clients) noted that they were encouraging a move away from use of qualifications in their recruitment practices. Overall, therefore, well over half the organisations surveyed either excluded qualifications from their formal criteria, or were reducing their role. The corollary, in these cases, is increased use of tests.

It is important, however, to note that this is not, for the most part, a reaction to perceived or experienced changes in the nature of current qualifications. The two most frequently mentioned reasons for downgrading qualifications were concerns about equal opportunities (discussed further below), or a move towards use of company-specific competency frameworks, in which qualifications may be accepted as a possible source of evidence relevant to a particular competency but are not a necessary or even a privileged source.

The adoption of competency frameworks for HR purposes (recruitment, appraisal, promotion, restructuring) became widespread among UK companies during the 1990s. One respondent, a psychologist with an unusually long period of involvement with his (large) company's selection policies, noted that 'From a psychology perspective, there was a big change in the early '90s — a move from dimensions to competencies.' Only some of the organisations we visited had moved to wholesale use of competencies. Most of these were public sector, and in them, 'Competencies have a huge impact on recruitment — as they do on all the processes.' But in many others, competence frameworks were being introduced, and used, in particular, when new positions were established or for senior appointments. 'We do a job and person specifications, and agree the competence framework indicators ...tests fit really well with the skills and qualities identified there: critical reasoning, analysis...'

Ability, aptitude or literacy/numeracy tests, are thus preferred to qualifications as a source of evidence because they link quite directly with one or more competencies (e.g. analytical reasoning); the same is true of personality batteries. In some cases where competencies are linked to ability tests (most often verbal, numerical or analytical reasoning) qualifications may be taken as a proxy, and candidates tested only if they do not have

particular qualifications. However, it is far more common for all candidates to be given the same tests regardless of their formal qualifications.

Respondents who commented directly on the evidence provided by qualifications most often underlined the fact that, in most recruitment situations, qualifications would be quite old and hence no guarantee of current skills. Related to this were anxieties about the truthfulness of CVs: 'We've had people where we inferred ...that they had the written skills, and then it turned out not to be true'. However, four respondents did comment directly on the current standard of qualifications; always in relation to academic awards and degrees. (None of our respondents mentioned vocational ones.) All of them did so negatively. One (in local government) noted that, in his experience, there was 'an amazing lack of correlation' between GCSE and ability test results, and he therefore now relied on the latter; and another that he and others felt an 'increased mistrust' of exam results.¹⁸ However, it was far more common for respondents to emphasise either equal opportunity issues or HR's commitment to a competence framework, when explaining the exclusion from or decreasing role played by formal qualifications in selection for their organisations.

It should not be concluded, however, that formal qualifications have become totally irrelevant. (And certainly not that their content has. The skills that produce high performance levels in qualifications are also likely to produce high test performance on, in particular, verbal and numerical reasoning tests.)¹⁹ Qualifications are frequently taken into account, formally or informally. For example, in many cases where large numbers of applications are received, preliminary short-listing is carried out by personnel specialists in-house or externally. Qualifications are often one possible source of evidence used in assessing eligibility against job and person specifications, and in drawing up shortlists, even though they are not the only acceptable such source. One recruitment specialist notes that 'If I'm sifting, and have two candidates with the same experience, but one has GCSEs at C+ and the other doesn't, I would go for the one with GCSEs'. And another noted that 'some clients do still insist on high quality degrees from prestigious universities ...though opinion is becoming more enlightened'.

¹⁸ One of the respondents who expressed negative views regarding the standard/reliability of current school-level qualifications also hoped that key skill qualifications would establish themselves as an accurate indicator. His was the only mention made in any interview of key skills.

¹⁹ It would be perfectly possible, for this reason alone, to find high returns to qualifications in macro studies alongside a complete ban on using formal qualifications for recruitment.

In a number of occupations, moreover, specific professional and vocational qualifications are of major importance or even mandatory, often for legal reasons: social work, school teaching, medical and medically-related occupations, legal and engineering positions all fall into this category. And a considerable number of large organisations also run graduate recruitment schemes, entry into which, by definition, is qualification-dependent.

Nonetheless, the growing reluctance to use qualifications (and the commensurate lack of interest in the details of qualification change and reform) is a striking feature of the case study evidence. We return to the implications (if any) for the education system in the Conclusion.

Changes in the skill mix. Another possible reason for an increase in test use because of changes in its net value would be that the skills required for jobs had changed. (The underlying scenario here would be a change in the nature of jobs which made test-based evidence more relevant, because of a better match between test outcomes and job requirements than in the past or because of greater productivity differences between workers with different test scores.) Test use is controlled and to a large degree recommended and driven by HR professionals: so increases in test use for this reason would occur only if they were aware of such underlying shifts.

We found little evidence of this. In most cases, it is unlikely that such an adjustment would happen quickly, since, except when large numbers of people are being hired, most organisations report that managers do not have clear, well-founded job and person specifications to hand when they decide to recruit. One of the main tasks of the HR team is to create these in consultation with the recruiting department. While changes in what employers want may thus feed through into test use gradually, there was no clear and discernible link in the bulk of case studies.

The exception is when major restructuring or expansion occurs, and large numbers of new positions must be filled. The use of formal testing programmes, including the commissioning of bespoke tests, is then extremely common: a considerable number of such recruitment campaigns were described to us, and in every one, testing played a central part. It seems likely that tests are more common in such situations now than they were 10, let alone 20, years ago but the link has, on the basis of the reports, more to do with other factors (HR formalisation, regulatory environment *etc*) than it has to do with changes in the importance of skills or their relationship to test coverage. Skill needs changed in the past, and they change today, but we could not find anything to link *increases* in test use to changes in the *speed* or *direction* in which skill demands alter.

Shifts in perceived risk (or in levels of risk aversion). In contrast, we did find considerable evidence that the output of tests was valued more than in the past as a result of organisations perceiving the risks attendant on a ‘bad’ hire to be greater. No-one mentioned test use as increasing because the returns to a ‘good’ hiring decision had risen. In every case where external changes were referred to as the source of altered risks, or higher risk aversion, they were changes in the regulatory or political environment, rather than changes in the nature of the job, or increases in the variability of worker productivity. (Again, this may reflect the HR-driven nature of selection practices.) However, there was also some indication that testing was more likely than in the past to be a response to these situations because of increased confidence that the information from tests would, in these situations, be genuinely valid, as well as evidence of ‘good’ recruitment practice.

Often, the growth in company ‘risk aversion’ was in response to situations which were not particularly novel or specific to the late 1990s or early 2000s. For example, ‘our company grew fast with virtually blind recruitment and got some very incompetent staff. There was panic and they brought in new people, including me, with a remit for recruitment in key areas.’ Or, ‘we’re constantly struggling for new ways to get the right person for these jobs. They’re our public face — and we found we’d made some bad decisions. We were asking ourselves, “How on earth did we get you on board?”’ In these situations, the higher risk aversion is company-specific, not sector-wide but respondents turned automatically to testing as a possible way forward, where twenty years ago, that would not, they feel, have been the case.

In other cases, there is a feeling that the general level of concern has risen, because the costs of failure have increased, and/or public tolerance has fallen. This is particularly the case for low-paid ‘front line’ employees, where tests are seen as a critical defence against incorrect hires. ‘One respondent noted that ‘We’ve hired people in the past who’ve turned out to be illiterate or innumerate. In a hospital, this can have potentially serious health and safety implications.’ According to another, ‘There are no specialist skills we can’t cover in six or seven weeks’ training, but there’s a very strong safety element. We *have* to identify those who aren’t going to succeed – maybe they don’t have the ability to learn at the rate we expect, or have an inadequate concentration span. The tests are very important for that and we have very strong validation of the link between these, success on the job and system-wide safety.’

There is obvious overlap here with concerns about equal opportunity (see section E below) but the objective is less to cover oneself in a formal way, and more to do with

substantive decision making. The changes in the external environment which create greater risk aversion involve, as already noted, changes in the costs of failure — and those, in turn, include alterations in employment law, which make it increasingly difficult to dismiss people. That these changes translate into increased test use, however, once again seems to reflect the increasing formalisation and professionalisation of HRDs, and growing knowledge of tests and testing procedures, to which we also return below.

The belief of HR professionals in test results is worth some further discussion here because it encourages the development of bespoke tests, users of which invariably expressed satisfaction. Bespoke tests have been in existence for a long time, as have general ones: hence the point that increased test use is unlikely to be explained by changes in the nature of tests and the evidence they produce. It was a consistent perception among our respondents that, in the last decade, more large private-sector companies than in the past have commissioned specialised tests, either because they were expanding into new fields and had to hire or promote/re-appoint a large number of people at once, or because of major reorganisations involving redeployment of staff. These tests are typically developed by specialised external companies, and reflect a further growth in the level of activity in this area, and in knowledge about testing and test opportunities. It is also striking that bespoke tests are very often concerned with quite specialised skills among employees who will occupy relatively low-level but front-line jobs, involving either safety-related activities or customer contact.²⁰

C. Is increased test use a response to increased training and development costs for staff?

Many commentators believe that changes in the world economy mean that staff skills need to be updated more regularly than in the past; that ever-higher levels of skills are required in general; and that one can therefore expect that effective companies will increase their training and development budgets. If so, then one might expect that, as the cost of a new recruit increases (in terms of more or less fixed training costs per hire), this would translate into higher levels of test use at recruitment. The argument here is analogous to the one made

²⁰ In the public sector there may well have been a decrease in *in-house* bespoke test development with the disappearance of some clusters of psychologists. This is certainly true for privatised industries.

earlier, in relation to changes in risk/risk aversion at hiring but the posited change is in costs per employee rather than risks to the company from a wrong decision.

As we saw earlier, the cross-sectional evidence finds a strong relationship between companies' levels of training activity and their levels of test use. This suggests that it may indeed be the case that, the higher the fixed up-front costs of hiring someone, the more test evidence is valued as a source of improved recruitment decisions. However, the case study evidence did not suggest that this is a major reason for changes in test use levels. On the contrary: the issue and possibility were never raised in explaining which tests were used, why, and why, in many cases, test use was more frequent than in the past.

D. Is increased test use the result of a changing balance between test administration costs and costs of other recruitment methods?

Recruitment is expensive in terms of advertisements, headhunters, internal HRD staff or external recruitment specialists to process, sift, test, *etc.* – and, as importantly, the time of senior members of staff, many of who devote substantial proportions of their time to this key activity with associated direct and opportunity costs. If the cost of testing, for given levels of perceived validity, shifts in relation to other costs, then one might expect test use to increase or decrease accordingly.

The case study interviews suggest that, in many companies, increased test use has been 'sold' successfully to senior managers as an effective way of providing information that it might be much more difficult and expensive to obtain from interviews, and as a way of keeping the length of interviews down. (No organisation dispenses with interviews which remain critical to recruitment and the context for final selection decisions.) We did not encounter any cases where this was actually presented in explicitly costed terms. However, just under half our respondents mentioned that they evaluate the entire recruitment process in terms of the relative cost-effectiveness of particular components – something which has led on occasion to the replacement of one test by another, or even, in one case, the decision to cancel personality testing. Five mentioned explicitly the need to keep interviews quite brief and efficient because of the cost of managers' time, and the contribution that test data can make to this process.

Equally important is the increasing ease with which large amounts of testing can be done. More and more of the large well-known tests are available in computer-based form, so that, instead of lengthy report writing, the basic grind of scoring and interpretation is replaced

by a computer-generated report. Comments included: ‘For the personality test, we made our decision on cost grounds – it’s really easy to download the manuals and the training is quite quick...’; ‘Within the assessment centre, we can only afford to make time for one test, and it needs to be something which is quick to use’; and ‘These clients are massively over-subscribed – they can easily afford to pay huge amounts for the selection, but what they need is to process lots of people in a day’. While computerisation may not have a major effect on the ‘face’ cost of a test it means that a single individual or department can produce far more test reports in a given amount of time; and do so with the speed required if the data are to be fed into the interview and recruitment decision. Respondents note that, in the past and even today, it was far from uncommon for test data to be effectively irrelevant since the reports reached the interview panels too late. Also important in this context is the growth of independent recruitment specialists who provide an out-sourced facility for testing and reporting back: this significantly reduces the in-house investment required to train, and maintain a number of trained personnel. It was noticeable in our sample how many private sector companies used this route – over half the private sector organisations surveyed, as compared to only two of the public sector organisations, relied largely or entirely on external consultants for testing as well as for much of the handling of recruitment.

E. Is increased test use a defensive strategy adopted in response to regulatory and other external changes?

The answer to this is unquestionably yes: it was the most frequently cited factor, and seen as important by many respondents. In the vast majority of cases, respondents referred explicitly either to the increased risks now attendant on making hiring decisions in an inappropriate way, or (even more frequently) to the importance of being able to defend hiring practices as fair and objective should there be a challenge. In all these cases, tests were seen as an important safeguard and potential defence. As mentioned earlier, in the United States, a number of court cases in the 1980s, which ruled against companies on the grounds of inappropriate test use, are thought to have reduced test use by companies. There was no evidence of such a reaction here. On the contrary, whether through direct experience with tribunals (in two cases), knowledge of cases affecting others, or simply a generalised awareness of the situation, there was unanimity that using tests in recruitment was powerful evidence in favour of properly thought out and objective recruitment procedures.

The following five quotations are typical.

‘The design of the selection process was driven very much by the need to show fairness.’

‘Psychometric tests are certainly seen partly as a safeguard against accusations of unfair practice.’

‘Fairness is always an issue. We’ve given up using role plays because some roles are easier/more sympathetic to play than others.’

‘Using personality tests has pre-empted the trouble experienced by other local authorities. On the contrary, we’ve been praised for objectivity and the rigour of our testing procedures. People very seldom opt to go before tribunals because it is so easy to refer back to our systems. ...If they did the same with political appointments they would have avoided a lot of trouble.’

‘If the tests used don’t correlate closely with hard data on job requirements, and with the job and person specifications, it could lead to confusion and even to tribunals.’

This does not mean that respondents were not genuinely committed to equal opportunities policies. On the contrary, we are sure that many felt strongly about their importance and saw themselves as having a mission to educate their organisations. Nonetheless, they were also keenly aware of the potential risks to their employers - and, in consequence, to their own reputations and jobs. It is also the case that, while the issue was raised directly in every interview, only four respondents felt that there was any risk of adverse impact as a result of test use itself. All of them felt that, in the case of ‘restricted’ tests (requiring licenses) the validation statistics, and the guarantees of lack of bias which the companies concerned provide, were central to the potential value of the tests in cases of legal challenge and all but the same four respondents accepted these statistics and statements without reservation. As one of our interviewees put it, ‘High calibre test publishers ensure that tests cannot reflect either gender or ethnicity. In any case, any good assessor would know if the test was impartial or not’.

This same set of forces is also critically important in explaining the declining use of qualifications. As noted above, there is some evidence of declining faith in the quality of academic qualifications (vocational ones were never mentioned). There is considerably more evidence of concern that using qualifications as a basis for recruitment decisions could lay an employer open to challenge on equal opportunity grounds. As one respondent explained, in describing the decision of a major government department to bar any use of qualifications for

selection, ‘We were really concerned about adverse impact, for older women and minorities especially. The Civil Service Commission wants recruitment to be fair and open: we thought we were far better off using tests and competencies’.

F. Is increased test use the result of increased formalisation of Human Resources activities and departments?

Traditionally, personnel managers have tended to occupy a weak and subordinate role in organisations (Legge, 1995; Lupton, 2000). There is evidence that HR professionals have grown in number and in status as a self-conscious and qualified profession over the last twenty years. The management of employee relations has increasingly become the province of specialists rather than generalists, and practitioners are much better qualified than in the past (Cully *et al*, 1999; Millward *et al*, 2000). These changes have occurred partly in response to the increased volume of regulation and partly as a result of more general management trends. This means there are more people in organisations who know about and value tests, and who, as part of the more general (and partly defensive) formalisation of recruitment practice, are in a position to promote testing in their organisations. In doing so, they may also ‘educate’ other staff in their departments, and familiarise managers with their use, so that the latter start to expect and demand them. How far do these developments help to explain increases in test use?

As the previous sections will have suggested, HR formalisation is important as an ‘enabler’, making it possible for tests to be used, at short notice, either through in-house expertise or through outsourcing to recruitment specialists. There is also a symbiosis between the growth of potentially threatening regulation and the importance of specialists who can protect organisations against threats and costly litigation. In addition, interviewees suggested another way in which HR formalisation has increased test use – namely, through more general reforms of the whole recruitment process.

Many of our respondents described how tests had been introduced into organisations: by themselves or their predecessors, in conjunction with a general attempt to tighten up the recruitment process, and to make procedures and criteria fit clear job and person specifications. (This was seen as highly desirable on substantive grounds, in order to improve the quality of decisions, and not simply as a way of showing that the process was fair.) From the perspective of an HR professional, managers are, most of the time, seen as remarkably unclear about the precise nature of the skills and characteristics they want, and

slow to analyse the requirements of the posts they are trying to fill. ‘Managers aren’t actually sure what they are looking for — there’s no matrix there of information wanted against method to be used. So if they do say they want a test, it will be on the lines of “Oh, I’ve used them before — I’ll have one of them.” When HR ask “Why?” they haven’t a clue.’

The companies included in this study tended not only to be increasing test use, but also to have rather ‘formalised’ selection procedures generally. Moreover, in the cases where formalisation had increased markedly in recent years, this had very clearly been HR-driven. (It seems likely that this was also the scenario for others, with *e.g.* long-established assessment centres, but we cannot be certain.) As one explained, ‘The HR department was the motivating force for change. Test use was just a part of the drive to get the right person for the job — we’ve had to move to proper job and person specs, and are working towards competencies.’ Change may be stimulated by appeals to ‘defensive’ self-interest: another respondent explained ‘What I found here was potential legal dynamite — no clear job description, no thought about selection criteria’. But the justification is also the quality of appointments, and their central importance to the organisation. Since there is every reason (given the regulatory climate) to expect at the very least that formal centralised ‘HR functions’ will maintain their current importance (and the strong likelihood that this will increase), we can expect further increases in test use from this source alone.

6. Conclusion

On balance, it seems that the major forces driving increased test use are to be found among the regulations and employment law that constrain companies’ behaviour rather than in the content of academic and vocational curricula, or in the degree to which ‘standards’ are being maintained. It also seems extremely likely that test use levels will increase, simply from the continuing effects of factors already in place, including the professionalisation and formalisation of ‘Human Resources’ activities. Changes in underlying skill needs, or concerns over qualification standards, appear, by contrast, to have little direct effect on decisions about when, and which, tests to use.

This does not, however, mean that firms’ increasing reliance on tests has no implications for education policy: on the contrary. This is particularly the case with the most widely used batteries of ability tests which measure verbal and numerical reasoning, and

which are seen by companies as reflecting and measuring skills which they value and need. In other words, they are looking for some of the core skills taught in the academic curriculum, and it is important for young people and those involved in policy to be aware that, while the formal qualification will not secure a job, the skills that are learned in studying for it may.

At the same time, employers' declining willingness to take formal note of qualifications, which is strengthened by a definite (albeit moderate) level of scepticism about their meaning, does call into question the idea that education and training should be as qualification-driven as they have become over the last ten years. The unstated assumption behind much government policy seems to be that a course is worthless to the individual if no qualification is attached. This is clearly not the case if it teaches skills which can be demonstrated during recruitment testing and it is simply untrue, on the basis of these studies, to imply that all formal qualifications have general (or even specific) labour market value.

These remarks apply predominantly to tests of ability and aptitude, and to some specific skill tests. The implications for education policy of increasing use of personality batteries is less clear. Some of our more sceptical respondents doubted if they really had very much effect on selection, except for a few individuals with outlying scores (because most profiles cluster around the middle range). However, it does seem important that people entering or re-entering the workplace should be made aware of the nature of these tests, and of how they are scored. It would also seem important that further study is conducted on the validity of such tests and of the dimensions underlying them.

Appendix A: A Summary of Major Tests in Use

Manufacturers generally require test users to be formally trained. This is particularly true for personality tests, but also for many ability tests. There is a distinction between *users* and *test administrators*, who brief candidates, hand out and collect test papers and sometimes score the results (under supervision).

British users and personnel/HRD employees often refer to “level A” or “level B” training. This refers to standards and certificates originally developed and introduced by the British Psychological Society (BPS) and the Institute of Personnel and Development (IPD: now the CIPD).

- *Level A* covers the general foundations of testing and the administration and interpretation of ability tests.
- *Level B* covers personality assessment, and the interpretation and use of personality tests.

The two certificates complement each other, and an individual has to pass Level A before taking Level B (although some trainers offer a course combining the two). The competence of the potential test user is assessed by a chartered psychologist whose assessment practices will in turn have been verified to BPS standards. A number of publishers offer part of the Level A training as an affordable way to train as a test administrator (as opposed to being a test user). Although this does not make those who undertake the training eligible to purchase test materials, they can, however, run test sessions.

As a general rule, Level A enables holders to use ability tests but most (not all) personality tests are available only to those certified at Level B.

Level A

The Level A course with a registered chartered occupational psychologist covers seven units of core competence, each of which must be met by trainees before they are given certification:

- *defining assessment needs* (understanding the differences between psychometric measures)
- *basic principles and validity* (e.g. calculating standard deviations of test scores)
- *reliability and validity* (understanding correlations and why test scores may be unreliable)
- *deciding when and when not to use tests as part of an assessment process* (equal opportunities; practical considerations; weighting tests in selection decisions)
- *administering and scoring tests* (carrying out test sessions; using scoring keys)
- *using test results appropriately and providing accurate written and oral feedback* (presenting reports in lay terms; encouraging candidates to comment on the information from their test)
- *security and confidentiality of test materials and data* (ensuring access is not given to unauthorised personnel).

Level B

The Level B qualification covers tests where interpretation requires knowledge of psychological theories or models (such as trait theory). There are actually two Level B qualifications: intermediate and full. The full qualification is intended for practitioners and is a certificate of competence in more than one ‘type’ of personality test. In the majority of cases, would-be test users take the intermediate Level B qualification. This qualifies the user in a specific type of personality assessment. Some publishers accept that the BPS Level B (intermediate qualification) gives users enough of a broad grounding in personality assessment so as not to require additional training if buying a test other than that in which they are trained. Other suppliers require users to take some additional training.

Some examples of tests in use are given below, distinguishing between ability/aptitude tests and personality tests. We derived much of this information from Incomes Data Services (2000).

Ability and aptitude tests

Examples of this include the following from the major companies operating in the UK:

ASE

General Ability Tests 2

Modern Occupational Test Series

Graduate and Managerial Assessment

First Graduate Assessment

Skillscape

Critical Reasoning Tests

ACER tests

New Technology Tests

Computer Programmer Ability Battery

The AH Series of Tests

Oxford Psychologists Press

The Able Series

This was first published in 1996, and consists of tests combining work simulation exercises and psychometric testing. They aim to relate candidates' skills and abilities in a working environment and assess the potential to learn tasks, to quickly become successful in a job, and to adapt to changes in the working environment. The eleven tests in the series comprise business decision analysis, commercial reasoning, fault identification, critical business planning, critical information analysis, legal interpretation, financial appraisal, product inspection, performance programming, vetting applications and helpline. The tests take between 30 and 45 minutes to complete.

Raven's Progressive Matrices

This is a very old test, first developed in 1938. It is a test of general cognitive ability, and consists of a series of progressively more difficult problems. It is a measure of general ability.

Destiny Series

Critical Reasoning Skills Series

Psytech International

This company's products include

Graduate Reasoning Tests

General Reasoning Tests

Critical Reasoning Test Battery

Technical Test Battery

Clerical Test Battery

SHL Group

Among SHL's tests for measuring ability are the following.

Advanced Management Tests (AMT)

Four tests for middle/senior managers, professionals and graduates. These tests are at a higher level of difficulty than the MGIB (listed below).

Management and Graduate Item Bank (MGIB)

MGIB consists of tests which assess critical reasoning abilities at graduate or middle to senior management level. Eight versions are available, four verbal and four numerical tests. The verbal tests take 25 minutes and the numerical tests 35 minutes to complete.

Critical Reasoning Test Battery (CRTB)

CRTB comprises tests of reasoning skills at administrative, supervisory and junior management level.

Information Technology Test Series

Customer Contact Aptitude Series (CCAS)

CCAS consists of aptitude tests aimed at sales and customer service staff and assessing verbal and numerical reasoning skills.

Personnel Test Battery

Automated Office Battery

The Automated Office Battery (AOB) includes a numerical estimation test which assesses the ability to estimate the correct answer to a calculation; computer checking test which measures the ability to check machine input against the resulting output; another test assesses the ability to comprehend written instructions when a form of coded language is used. The battery aims to indicate whether a candidate has the skills necessary to work in an automated office environment.

Technical Test Battery

The Technical Test Battery is designed to select for a range of technical occupations. Specific tests include a test of verbal comprehension of vocabulary from a technical environment; numerical computation; numerical reasoning; spatial recognition of shapes in two dimensions; mechanical comprehension, covering basic mechanical principles and application to levers, pulleys *etc*; technical understanding, testing based on written passages from technical literature; and fault diagnosis.

Applied Technology Series

Work Skills Series Transport

Work Skills Series Manual Dexterity

Personality tests

Popular examples include:

ASE

16PF – the Sixteen Personality Factor Questionnaire

The fifth edition was launched in 1994. The questionnaire is designed for professional and technical staff, as well as sales staff and graduates. The 16 personality factors are: warmth, reasoning, emotional stability, dominance, social boldness, liveliness, role consciousness, sensitivity, vigilance, abstractedness, privateness, apprehension, openness to change, self-reliance, perfectionism and tension. Norms are available for the British population in

general, males and females, manual and non-manual occupations. The questionnaire has 185 items and takes 45 minutes to complete.

OPP

Myers-Briggs Type Indicator

This test was first released in the UK market in 1991/2 and revised in 1998. It is based on Jung's theory of personality with an individual's preferences categorised on four separate dimensions allowing the identification of 16 different 'types'. The test is not timed but usually takes 20 to 30 minutes to complete.

California Psychological Inventory

Fundamental Interpersonal Relations Orientation - Behaviour (FIRO- B)

Innovation Potential Indicator

Psytech

15FQ

Occupational Personality Profile

Jung Type Indicator

16PF industrial

Values and Motives Inventory

SHL

OPQ 32

OPQ32 is the latest version of the Occupational Personality Questionnaire, launched in April 1999. The OPQ32 assesses personality using 32 characteristics which are grouped under three main headings: relationships with people; thinking styles; feelings and emotions. 'Relationships with people' is sub-divided into influence, sociability and empathy. Thinking styles is broken down into analysis, creativity and change, and structure. Feelings and emotions is split into emotion and dynamism. There are further sub-divisions within each of the categories.

OPQ32 is available in ipsative format and in normative format. The ipsative format (OPQ32I) consists of 100 blocks of 4 statements, and the respondent is asked to state which of each set of statements is most and least true of them. This takes about 45 minutes to complete. The normative version (OPQ32N) gives a list of statements and asks respondents.

OPQ32 is part of a family of tests: others in the series include the customer contact styles questionnaire, Work Styles Questionnaire, Images and Factor Models.

OPQ Factor 4.2 and 5.2

Customer Contact Styles Questionnaire

Work Styles Questionnaire

Images 1

Appendix B: Sample Test Items

1. Verbal reasoning

A principal objective of the Health and Safety at Work Act 1974 is to involve everybody in the workplace. The promotion of health and safety is an essential function of good management in all areas. The employer, however, also has a duty to protect other persons on the premises. The term 'premises' includes all workplaces, and in particular, includes any vehicle, vessel, aircraft, or any other installation on land. Regulations may be made exempting certain employers from the necessity to provide all employees with written copies of the Act.

1. The Health and Safety Act is the sole responsibility of the employee.
2. All installations on land, and vehicles are subsumed under the heading 'premises'.
3. The Act states that all employees must receive written copies.
4. There were no health and safety rules before 1974.

Mark Circle A if the statement is TRUE, or follows logically, given the information contained in the passage.

Mark Circle B if the statement is FALSE, or the opposite follows logically, given the information contained in the passage.

Mark Circle C if you CANNOT say without further information.

1. A B C
2. A B C
3. A B C
4. A B C

2. Numerical reasoning

Newspaper readership

Daily Newspapers	Readership (millions)		Percentage of copies sold in Year 2, by outlet:	
	Year 1	Year 2	Subscription	Newsagents
The Daily Chronicle	3.6	2.9	47%	37%
The Daily News	14.8	10.0	50%	33%
The Tribune	1.1	1.4	43%	43%
The Herald	8.4	10.9	29%	55%
The Daily Echo	4.8	4.9	30%	60%

1. What was the combined readership of **the Daily Chronicle, The Tribune and The Daily Echo** in Year 2 (in millions).

- A B C D E
- 9.2 9.5 18.7 31.8 None of these

2. Which daily newspaper showed the largest percentage change in readership from Year 1 to Year 2?

- A B C D E
- The Daily Chronicle The Daily News The Tribune The Herald The Daily Echo

3. If 4.5 million copies of **The Daily News** were sold by subscription, approximately how many copies (in millions) were sold by newsagents in Year 2?

- A B C D E
- 1.5 2.3 3 6 Cannot say

3. Personality Tests

(a) Mark the [1] if you strongly agree with the statement; the [2] if you agree; the [3] if you are unsure; the [4] if you disagree; and the [5] if you strongly disagree.

- | | | | | | | |
|----|-----------------------------------|-----|-----|-----|-----|-----|
| 1. | I enjoy social gatherings | [1] | [2] | [3] | [4] | [5] |
| 2. | I am always late for appointments | [1] | [2] | [3] | [4] | [5] |
| 3. | I avoid taking risks | [1] | [2] | [3] | [4] | [5] |
| 4. | I like playing team games | [1] | [2] | [3] | [4] | [5] |

(b) For each block of four statements indicate which one is most like you [M], and which one is least like you [L]. leave the other two statements blank.

I am...

- | | | |
|----------------------------------|-----|-----|
| A highly organised person | [M] | [L] |
| Anxious when meeting people | [M] | [L] |
| The 'life and soul' of the party | [M] | [L] |
| An energetic person | [M] | [L] |

(c) For each statement indicate whether you think it is true [T] or false [F]:

- | | | | |
|----|--|-----|-----|
| 1. | I like to try new ways of doing things | [T] | [F] |
| 2. | I find it difficult to relax after a hard day's work | [T] | [F] |
| 3. | I am an optimistic person | [T] | [F] |
| 4. | I like organising meetings and gatherings | [T] | [F] |

(d) For each statement indicate whether you think the statement is true [T]; false [F]; or if you are not sure [?]:

- | | | | | |
|----|---|-----|-----|-----|
| 1. | People should sort out their own problems | [T] | [?] | [F] |
| 2. | I can get more done when I can work alone | [T] | [?] | [F] |
| 3. | I like telling people funny stories | [T] | [?] | [F] |
| 4. | I show my emotions easily | [T] | [?] | [F] |

Appendix C

We had available to us two primary sets of data relating to current use of testing for recruitment by UK organisations which enable us to examine some (not all) of the hypotheses outlined above. One is the 1998 Workplace Employee Relations Survey, a large government-sponsored survey of workplace practices which draws on a stratified sample of public and private sector UK workplaces with 10 or more employees.²¹ It achieved a very high (80%) response rate which did not vary to any substantial degree by either workforce size or industrial classification. 2191 workplaces took part: in each there was an interview with the manager most responsible for personnel matters, an interview with a worker representative and an employee survey.

The WERS manager questionnaire included the following two questions:

‘When filling vacancies at this workplace, do you ever conduct any type of personality or attitude test?’

and

‘When filling vacancies at this workplace, do you ever conduct any type of performance or competency test?’

The survey also collected information on a large number of workplace variables, including size, union penetration, equal opportunities policies, personnel function and training practices.

The second source of up-to-date data on current test use was the 2001 Recruitment Survey conducted by the UK’s Chartered Institute of Personnel and Development, the main professional body for personnel and training staff. The survey has been carried out for the last five years. For 2001, in collaboration with the CIPD, we inserted some questions which collected more detailed information on test use than in previous years. Respondents were asked the following questions:

²¹ The sample is stratified by workforce size and industrial activity. Large workplaces are over-sampled relative to their number as a proportion of all workplaces but results can be weighted to give estimates for the entire population of workplaces employing 10+ people.

14. I'd now like to ask you about the ways that you select applicants for your vacancies. I'm going to read out a list, and ask you to indicate whether you use each method for managers; professional staff such as IT specialists, engineers and accountants; or skilled manual workers. First of all, I'd like to ask you about your use of testing as part of selection. Do you use:

Personality questionnaires?

General ability tests?

Literacy/numeracy tests?

Tests of specific skills?

15. Still thinking about your use of testing in selection, has the amount of your usage changed over the past two years, 1999 to 2000? Is there:

Increased use? Less use? About the same?

Personality questionnaires

General ability tests

Literacy/numeracy tests

Tests of specific skills

16. Here are some reasons for using selection tests. For each reason, please tell me how important it is to you:

Is it 'important'? 'Fairly important'? or 'Not important'?

The shortlisting of candidates

Assessing the fit of candidates to specific jobs

Providing further information about candidates

Measuring specific skills or abilities

The objectivity of the information from the tests

In analysing the WERS data, we examined the effect on (a) personality test use and (b) competency test use of the following variables:

Workplace size

Levels of workforce training

Investor in People status

Existence of short term contracts (<1 year; or 1 year+)

Whether any groups with guaranteed job security
Whether employees can expect long term employment
Proportion with high salaries
Composition of workforce (% professional; managerial; technical; clerical; sales; operative; routine unskilled)
Type of vacancies in past 12 months
Unemployment rate in the travel to work area
Whether there are workers on performance related pay

Operation of internal labour market
Union recognition
Sector: public/private
 industrial sector
Workplace age

Number of diversity promoting practices
Whether a tribunal complaint in the last year
Whether an equal opportunities policy
Whether a formal grievance procedure

The following tables summarise those variables for which a statistically significant relation was found. (For details, see Jenkins and Wolf 2002.)

Table C1**Combined effect of economic variables, organisational characteristics, establishment size and industrial sector on the use of personality tests (WERS)**

SURVEY LOGIT REGRESSION	Coef.	Std. Err.	P> t
Constant	-3.763	0.548	0.000 ***
<i>Workplace Size (base is 10 to 24 employees)</i>			
25 to 49 Employees	-0.162	0.281	0.564
50 to 99 Employees	-0.025	0.274	0.929
100 to 199 Employees	0.118	0.298	0.692
200 to 499 Employees	0.271	0.325	0.405
More Than 500 Employees	0.659	0.370	0.075 *
<i>Training (% in LOG received training)</i>			
High (80 to 100 %)	0.972	0.383	0.011 **
Medium (40 to 79 %)	1.398	0.424	0.001 ***
Low (1 to 39 %)	0.285	0.360	0.429
Investor in People	0.500	0.214	0.020 **
Short term contracts (< 1 year)	0.458	0.246	0.062 *
% Professional in Workplace	-0.029	0.008	0.001 ***
% Operatives in Workplace	0.021	0.007	0.002 ***
<i>Type of Vacancies</i>			
Managerial	0.554	0.218	0.011 **
Professional	0.754	0.295	0.011 **
Technical	-0.325	0.290	0.262
Clerical	0.036	0.240	0.882
Craft/Skilled	-0.018	0.284	0.949
Personal Service	-0.547	0.359	0.128
Sales	0.216	0.287	0.453
Operative	-0.451	0.364	0.214
Routine unskilled	-0.501	0.219	0.022 **
<i>Unemployment rate in twwa</i>			
2 to 3 %	-0.101	0.420	0.810
3 to 5 %	0.068	0.358	0.850
5 to 7 %	-0.352	0.356	0.323
7 to 7.75 %	-0.479	0.464	0.302
more than 7.75 %	-1.818	0.598	0.002 ***
No of Diversity Promoting Practices (0 to 6)	0.228	0.062	0.000 ***
<i>Workplace age (time at current address, base 0 to 4 yrs)</i>			
5 to 9 years	0.927	0.341	0.007 ***
10 to 24 years	0.387	0.325	0.234
25 or more years	0.993	0.313	0.002 ***
<i>Industrial Sector (base, manufacturing)</i>			
Electricity, Gas, water	2.746	0.608	0.000 ***
Construction	0.083	0.431	0.847
Wholesale & Retail	0.880	0.370	0.017 **
Hotels & Restaurants	0.563	0.468	0.229
Transport & Communication	0.832	0.584	0.154
Financial Services	1.651	0.762	0.030 **
Other Business Services	0.827	0.413	0.045 **
Public Administration	-0.156	0.571	0.785
Education	-0.186	0.629	0.768
Health	1.041	0.539	0.053 *
Other Community Services	0.231	0.719	0.748
Number of Observations			1984
% Correctly Predicted			73.085
% Test Users Correctly Predicted			45.257
% Non-Test Users Correctly Predicted			86.428

* Significant at 10% level, ** 5% level, *** 1% level

Table C2
Combined effect of economic variables, organisational characteristics, establishment size and industrial sector on the use of competency tests (WERS)

SURVEY LOGIT REGRESSION	Coef.	Std. Err.	P> t
Constant	-0.903	0.474	0.057 *
Fixed/Short term contracts (> 1 year)	0.483	0.255	0.058 *
% Routine Unskilled in Workplace	-0.011	0.005	0.019 **
Workers on PRP	0.492	0.244	0.044 **
<i>Type of Vacancies</i>			
Managerial	0.029	0.185	0.875
Professional	-0.290	0.225	0.198
Technical	0.486	0.248	0.050 **
Clerical	0.302	0.196	0.124
Craft/Skilled	0.562	0.205	0.006 ***
Personal Service	-0.166	0.272	0.542
Sales	-0.144	0.232	0.534
Operative	-0.716	0.297	0.016 **
Routine unskilled	0.091	0.229	0.691
No of Diversity Promoting Practices (0-6)	0.213	0.054	0.000 ***
Grievance Procedures	0.691	0.339	0.041 **
Complaints to Tribunal in last Year	0.455	0.224	0.042 **
<i>Industrial Sector (base, manufacturing)</i>			
Electricity, Gas, water	-0.116	0.528	0.826
Construction	-0.735	0.501	0.142
Wholesale & Retail	-0.748	0.381	0.050 **
Hotels & Restaurants	-0.954	0.452	0.035 **
Transport & Communication	0.497	0.517	0.337
Financial Services	0.344	0.519	0.507
Other Business Services	-0.094	0.411	0.819
Public Administration	0.625	0.482	0.195
Education	-0.183	0.455	0.687
Health	-0.392	0.447	0.381
Other Community Services	-0.010	0.474	0.983
Number of Observations			2101
% "Correctly" Predicted			65.731
% Test Users "Correctly" Predicted			74.798
% Non-Test Users "Correctly" Predicted			52.723

* Significant at 10% level, ** 5% level, *** 1% level

In analysing the CIPD data, the following variables were used:

- Organisation size
- Broad sector (manufacturing/services/public)
- Difficulty in filling vacancies
- Any steps taken to promote diverse workforce (scored 0,1)
- Number of different practices for promoting diversity (maximum 6)
- Number of different recruitment methods
- Use of informal recruitment methods
- Use of personnel department for recruitment
- Involvement of local personnel staff
- Involvement of central office personnel staff
- Involvement of external consultants
- Involvement of line managers
- Use of recruitment agencies
- Degree of involvement of line managers
- Time perspective when hiring
- Insistence on exact skills when recruiting
- Whether evaluate recruitment and selection methods

Tables C3 to C7 summarise the main relationships found.

Table C3
Logit regression for personality test use, CIPD dataset final specification

	Coef.	Std. Err.	z	P> z
Constant	-2.151	0.570	-3.770	0.000 **
<i>Organisation Size</i>				
100 to 199 Employees	1.318	0.527	2.500	0.012 **
200 to 499 Employees	0.931	0.488	1.910	0.057 *
500+ Employees	1.312	0.465	2.820	0.005 **
*				
Steps Taken to Promote Diverse Workforce	0.897	0.412	2.180	0.030 **
External Consultants Used for Recruitment	0.647	0.285	2.270	0.023 **
Filling Vacancies: Focus Only on Current Requirements	-0.645	0.288	-2.240	0.025 **
N	242			
pseudo-R2	0.086			
Log Likelihood	-150.207			

* Significant at 10%, ** Significant at 5%, *** Significant at 1%

Table C4
Logit regression for general ability test use, CIPD dataset final specification

	Coef.	Std. Err.	z	P> z
Constant	-0.772	0.535	-1.440	0.149
Local Personnel Function involved in Recruitment	0.700	0.291	2.410	0.016 **
Filling Vacancies: Consider More demanding jobs in future	0.503	0.304	1.660	0.097 *
Informal Recruitment Methods Used	-1.194	0.391	-3.050	0.002 **
				*
Recruitment Agencies Used	-0.559	0.322	-1.730	0.083 *
Number of Recruitment Methods	0.209	0.079	2.630	0.008 **
				*
Evaluation of Recruitment & Selection	0.539	0.327	1.650	0.099 *
N	239			
pseudo-R2	0.083			
Log Likelihood	-150.256			

* Significant at 10%, ** Significant at 5%, *** Significant at 1%

Table C5
Logit regression for literacy/numeracy test use, CIPD dataset final specification

	Coef.	Std. Err.	z	P> z
Constant	-0.109	0.394	-0.280	0.781
No of Practices for Promoting Diverse Workforce	0.249	0.085	2.920	0.004 **
				*
Filling Vacancies: Focus on Current Requirements Only	-0.570	0.282	-2.020	0.044 **
Informal Recruitment Methods Used	-0.608	0.323	-1.880	0.060 *
N	235			
pseudo-R2	0.0614			
Log Likelihood	-152.01404			

* Significant at 10%, ** Significant at 5%, *** Significant at 1%

Table C6
Logit regression for specific skills test use, CIPD dataset final specification

	Coef.	Std. Err.	z	P> z
Constant	0.449	0.233	1.930	0.054 *
<i>Business Sector</i>				
Services	-0.416	0.327	-1.270	0.204
Public Sector	0.685	0.326	2.100	0.035 **
Insist on Exact Skills when Recruiting	-0.596	0.336	-1.770	0.076 *
N	249			
pseudo-R2	0.041			
Log Likelihood	-159.699			

* Significant at 10%, ** Significant at 5%, *** Significant at 1%

Table C7
Logit regression for any test use, final specification

	Coef.	Std. Err.	z	P> z 	
1. CONSTANT	0.087	0.441	0.200	0.844	
2. Size 2: 100-199 employees	1.068	0.579	1.840	0.065	*
3. Size 3: 200-499 employees	0.157	0.480	0.330	0.744	
4. Size 4: 500+ employees	0.996	0.522	1.910	0.057	*
5. No. of diversity-promoting practices	0.281	0.119	2.360	0.018	**
6. Local personnel staff involved	0.856	0.387	2.210	0.027	**
N	240				
pseudo-R ²	0.092				
Log L	-99.687				

* Significant at 10%, ** Significant at 5%, *** Significant at 1%
 (Note: Size 1 = 50 to 99 employees)

Table C8
OLS regression for overall test use, final specification

	Coef.	Std. Err.	z	P> z 	
1. CONSTANT	3.405	0.488	6.980	0.000	**
2. MANUFACTURING SECTOR	-1.148	0.505	-2.270	0.024	**
3. SERVICES SECTOR	0.921	0.619	1.490	0.138	
4. NO OF DIVERSITY-PROMOTING PRACTICES	0.354	0.151	2.350	0.020	**
5. EXTERNAL CONSULTANTS INVOLVED	0.982	0.441	2.220	0.027	**
6. FOCUS ON CURRENT VACANCY ONLY	-1.020	0.428	-2.380	0.018	**
N	232				
R ²	0.159				

* Significant at 10%, ** Significant at 5%, *** Significant at 1%
 (Note: The residual sector category — only three were used — is Public Sector)

References

- Baron, J. N., Dobbin, F. R. and Jennings, P. D. (1986), 'War and Peace: the Evolution of Modern Personnel Administration in US Industry', American Journal of Sociology, 92, pp. 248-273.
- Baron, J. N., Jennings, P. D. and Dobbin, F. R. (1988), 'Mission Control? The Development of Personnel Systems in US Industry', American Sociological Review, 53, pp. 497-514.
- Blunkett, D. Rt. Hon. MP. (2000), *Modernising Higher Education – Facing the Global Challenge*, Department for Education and Employment.
- Blunkett, D. Rt. Hon. MP. (2001), *Education into Employability: the Role of the DfEE in the Economy*, Department for Education and Employment.
- CBI (1989), *Towards a Skills Revolution - a Youth Charter*, Interim Report of the Vocational Education and Training Task force, CBI: London.
- CBI (1995), *Realising the Vision: a Skills Passport*, CBI: London.
- Chartered Institute of Personnel and Development (2001), *Recruitment*, CIPD: London.
- Cohen, Y. and Pfeffer, J. (1986), 'Organizational Hiring Standards', Administrative Science Quarterly, 31, pp. 1-24.
- Cully, M., Woodland, S., O'Reilly, A., Dix, G. (1999), *Britain at Work: as Depicted by the 1998 Workplace Employee Relations Survey*, Routledge: London and New York.
- Dixit, A. (2000), *Incentives and Organizations in the Public Sector: An Interpretive Review*. unpublished paper, University of Princeton.
- Dobbin, F. R., Edelman, L., Meyer, J. W., Scott, W. R. and Swidler, A. (1988), 'The Expansion of Due Process in Organizations', in L. G. Zucker, (Ed.), *Institutional Patterns and Organisation*, Ballinger: Cambridge, MA.
- Dolton, P. and Vignoles, A. (1999), 'The Economic Case for A-Level Reform', Discussion Paper No. 422, Centre for Economic Performance, London School of Economics.
- Dolton, P. and Vignoles, A. (2000), 'The Pay-Off to Mathematics A-Level', in C. Tikly and A. Wolf (Eds.), *The Maths We Need Now*, Institute of Education: London, pp. 52-73.
- Hoyles, C., Noss, R. and Pozzi, S. (1999), 'Mathematizing in Practice', in C. Hoyles, C. Morgan and G. Woodhouse (Eds), *Rethinking the Mathematics Curriculum*, Falmer Press: London, pp. 48-62.
- Hoyles, C., Molyneux-Hodgson, S. and Wolf, A. (2002), *Mathematics Skills in the Workplace*, Final Report to the Science, Technology and Mathematics Council.
- Incomes Data Services (2000), *Psychological Tests*, IDS Studies in Personnel Policy and Practice: Study Plus.

- Institute of Directors (1994), *Performance and Potential: Education and Training for a Market Economy*, IoD: London.
- Jenkins, A. (2001a), *The Occupational Use of Psychometric Tests: Results of the CIPD Recruitment Forum*, unpublished report, Centre for the Economics of Education, London School of Economics.
- Jenkins, A. (2001b), *Companies' Use of Psychometric Testing and the Changing Demand for Skills: A Review of the Literature*, Discussion Paper No. 12, Centre for the Economics of Education, London School of Economics.
- Jenkins, A. and Wolf, A. (2002), *Why Do Employers Use Selection Tests? Evidence from British Workplaces*, CEE Discussion Paper No. 27: Centre for the Economics of Education, London School of Economics.
- Keep, E. and Mayhew, K. (1999), Knowledge, Skills and Competitiveness, Oxford Review of Economic Policy.
- Krueger, A. (2001), *Education Matters: Selected Essays*. Edward Elgar, Cheltenham.
- Lazear, E. (1995), *Personnel Economics*. MIT Press, Cambridge, Mass.
- Leadbeater, C. (1999), *Living on Thin Air: the New Economy*, Penguin: London.
- Legge, K. (1995), *Human Resource Management – Rhetorics and Realities*, Macmillan: Basingstoke.
- Lupton, B. (2000), 'Pouring the Coffee at Interviews? Personnel's Role in the Selection of Doctors' Personnel Review, 29/1, pp. 48-64.
- Millward, N., Bryson, A. and Forth, J. (2000), *All Change at Work?* Routledge.
- National Skills Task Force (2000), *Skills for All: Proposals for a National Skills Agenda*, Final report of the National Skills Task Force Department for Education and Employment: London.
- Noss, R. and Hoyles, C. (1996), 'The Visibility of Meanings: Modelling the Mathematics of Banking', International Journal of Computers for Mathematical Learning, 1, pp. 3-31.
- Prendergast, C. (1999), 'The Provision of Incentives in Firms', Journal of Economic Literature, 37.1, pp. 7-63.
- Pryor, F. L. and Schaffer, D. L. (1999), *Who's Not Working and Why: Employment, Cognitive Skills, Wages and the Changing US Labor Market*, CUP: Cambridge.
- Robinson, P. (1996), 'Skill Shortages and Full Employment: How Serious a Constraint? in J. Michie and J. Grieve-Smith (eds), *Creating Industrial Capacity: Towards Full Employment*, Oxford University Press: Oxford.
- Secretary's Commission on Achieving Necessary Skills (SCANS) (1991), *What Work Requires of Schools*, US Department of Labor: Washington DC.
- Sutherland, R. and Wolf, A. (1995), *A Spreadsheet Approach to Mathematical Modelling for Engineering Students*, Final Report to The Nuffield Foundation.

- Wing, L. and Gifford, B. (eds) (1993), *Policy Issues in Employment Testing*, Kluwer: Boston.
- Wolf, A. (1997), 'Mathematics For All: the Teaching of Post-Compulsory Students in International Context', Teaching Mathematics and its Applications, Special Issue 'Mathematics 16-19', 16, 4.
- Wolf, A. (1998), 'Politicians and Economic Panic', History of Education, 27, 3, pp. 219-234.
- Wolf, A. and Silver, R. (1995), *Measuring 'Broad' Skills: the Prediction of Skill Retention and Transfer over Time*, Research Series No. 61, Department for Education and Employment, Sheffield.

CENTRE FOR THE ECONOMICS OF EDUCATION
Recent Discussion Papers

- | | | |
|----|--|--|
| 28 | Arnaud Chevalier
Gauthier Lanot | Monotonicity and the Roy Model |
| 27 | A. Jenkins
A. Wolf | Why Do Employers Use Selection Tests? Evidence from British Workplaces |
| 26 | J. Blanden
A. Goodman
P. Gregg
S. Machin | Changes in Intergenerational Mobility in Britain |
| 25 | S. Gibbons | Geography, Resources and Primary School Performance |
| 24 | D. Clark | Participation in Post-Compulsory Education in England: What Explains the Boom and Bust? |
| 23 | R. Layard
S. McIntosh
A. Vignoles | Britain's Records on Skills |
| 22 | D. Bonjour
L. Cherkas
J. Haskel
D. Hawkes
T. Spector | Returns to Education: Evidence from UK Twins |
| 21 | A. Chevalier
P. Dolton
S. McIntosh | Recruiting and Retaining Teachers in the UK: an Analysis of Graduate Occupation Choice from the 1960s to the 1990s |
| 20 | G. Conlon | The Determinants of Undertaking Academic and Vocational Qualifications in the UK |
| 19 | A. Jenkins
A. Vignoles
A. Wolf
F. Galindo-Rueda | The Determinants and Effects of Lifelong Learning |
| 18 | S. Gibbons | Neighbourhood Effects on Educational Achievement |

- | | | |
|----|---|---|
| 17 | S. Gibbons | Paying for Good Neighbours? Neighbourhood Deprivation and the Community Benefits of Education |
| 16 | R. Blundell
L. Dearden
B. Sianesi | Estimating the Returns to Education: Models, Methods and Results |
| 15 | S. Gibbons
S. Machin | Valuing Primary Schools |
| 14 | C. Megir
M. Palme | The Effect of a Social Experiment in Education |
| 13 | G. Conlon | One in Three? The Incidence and Outcomes of Lifelong Learners in the United Kingdom |
| 12 | A. Jenkins | Companies Use of Psychometric Testing and the Changing Demand for Skills: A Review of the Literature |
| 11 | G. Conlon | The Differential in Earnings Premia Between Academically and Vocationally Trained Males in the United Kingdom |
| 10 | P. Dolton
O. D. Marcenaro
L. Navarro | The Effective Use of Student Time: A Stochastic Frontier Production Function Case Study |
| 9 | P. Dolton
M. Silles | Over-Education in the Graduate Labour Market: Some Evidence |
| 8 | A. Chevalier
G. Lanot | The Relative Effect of Family and Financial Characteristics on Educational Achievement |
| 7 | A. Chevalier | Graduate Over-Education in the UK |
| 6 | B. Sianesi
J. Van Reenen | The Returns to Education: A Review of the Macro-Economic Literature |
| 5 | C. Harmon
H. Osterbeek
I. Walker | The Returns to Education: A Review of Evidence, Issues and Deficiencies in the Literature |
| 4 | L. Dearden
S. McIntosh
M. Myck
A. Vignoles | The Returns to Academic and Vocational Qualifications in Britain |

To order a discussion paper, please contact the
Publications Unit
Centre for Economic Performance
Tel 020 7955 7673 Fax 020 7955 7595