Best Practice in the Management of Psychometric Tests

GUIDELINES FOR DEVELOPING POLICY

SHL

Best Practice in the Management of Psychometric Tests

GUIDELINES FOR DEVELOPING POLICY

Contents

SHL®

1.Introduction

Psychometric tests are powerful tools used by organisations for the selection, development and management of people. Tests can be used:

- to enhance the decision making process in assessment for selection and promotion,

- as an aid to management in areas such as motivation and team building

- to identify development needs, as a basis for employee counselling as well as in organisational areas such as management of change or succession planning.

Psychometric tests provide additional relevant information over and above that obtained from more traditional assessment methods. For instance, tests aid recruitment processes by ensuring that all candidates are treated fairly and measured against a common yardstick. Tests can be used to challenge stereotyped judgements made by interviewers and often enable a more objective analysis to take place than is possible by interviewing alone. Occupational tests have consistently been shown to be better predictors of job success than interviews.

Using tests can lead to substantial gains for an organisation in terms of increased output and efficiency, better quality staff, higher morale, more effective performance, lower training costs and reduced turnover. When properly used, psychometric testing helps to ensure a common language for assessment standards throughout an organisation, match people to jobs, identify individual capabilities and predict on-the-job performance.

2. Why is a policy needed?

Use of tests must be properly managed in order to reap the potential benefits. If inappropriate tests are used, or the information from tests is misunderstood, or even ignored, there will be no advantage in using tests - indeed there will be unproductive costs. The management of tests should include regular reviews and monitoring to ensure the organisational aims are being met.

In addition, as with any powerful tool, the potential for misuse is ever present. Bad practice can reduce the effectiveness of test use, cause considerable suffering to individuals, damage the company's image with employees, trade unions and clients, and even lead to contravention of the law and the costs and bad publicity associated with litigation.

A well thought out policy on test use will help to ensure that the organisation gains maximum benefit from test use and that potential misuse is avoided. It will also demonstrate commitment to good practice. The aims of every organisation in using tests are different and policy must be geared to the needs and structures of each organisation. For one company, a central unit supervising all test use may be appropriate. For another, individual test users may be required to take final responsibility for their own test use. For this reason each organisation must develop its own way of managing test use and formulate a local policy document.

The purpose of this booklet is to highlight the areas that ought to be covered in such a formal policy document. It also provides examples of what policy statements may look like. Readers are free to use or modify these as appropriate but the statements do not in themselves constitute a policy. It is beyond the scope of this booklet to provide detailed discussion of all the issues that need to be considered in developing a policy for managing tests. SHL is happy to supply expert consultancy services to help organisations create their own policies for test use if further guidance is desired.

3. What should a policy look like?

Policies work best when they are brief and focused. Long manuals that lie unread in a drawer have little effect on practice. However, general statements of principle tend to be vague, easily ignored and provide little guidance for the user. An effective strategy is to have brief policy statements supplemented with more detailed guidelines. The statements can encompass general principles and are likely to need few changes even in changing circumstances. The procedural guidelines put the policy into practice, and then can, and should, be amended whenever they cease to meet the aims of the organisation or the policy. The detail with which such procedures are specified will depend on the nature, style and needs of the organisation.

4. What should a policy include?

4.1 Mission statement

It is helpful to start with a short general statement of aims in using tests. This sets the tone and purpose of the document. It ensures that readers are aware of the positive objectives of the use of tests and the reasons for the existence of the policy. Otherwise, a policy can be seen as restrictive and bureaucratic, rather than as a helpful guide to doing things well.

Example

We use psychometric tests to enhance the quality and quantity of information available for selection, development and training decisions and as an aid to organisational change. We are committed to the highest standards of practice in the use of all psychometric tests, in order to maximise the benefit of testing to the organisation and the individual, and to promote fairness and equality of opportunity for all.

4.2 Overall responsibility for testing standards

Test users are always responsible for the way they personally use tests. However, in some organisations there is a central unit or department that decides policy, oversees practices and provides support for individual users. Often such a unit contains one or more professional occupational psychologists. Sometimes such units have an advisory role; at other times they dictate procedures.

In other organisations each test user makes his or her own autonomous decisions regarding the use of tests; small organisations may only have a few trained users; large organisations may have many users in a single personnel department or test use may be decentralised.

Whatever the structure, the responsibilities and accountabilities of each individual must be clearly defined. It is highly recommended that in large organisations with high volume test use, one or more chartered psychologists are available to support the design, implementation, validation and review of test procedures - either on a consultancy basis or as permanent members of staff.

Examples

Each test user must ensure that he/she uses tests to the highest professional standards and only in accordance with the guidelines set out in this policy.

The Central Testing Unit is responsible for ensuring that all test use in this organisation is in accordance with this policy. Any procedure that will involve the use of psychometric testing must be referred to the Unit for approval before implementation. The Unit will provide guidelines for the use of tests that must be followed by all users.

4.3 Who should use tests?

Knowledge and experience are required to use psychometric tests effectively. It is recognised throughout the world that psychometric instruments are potentially dangerous in the wrong hands. Indeed, in many countries only qualified psychologists are allowed to use them.

In the UK, all reputable test publishers maintain a register of qualified people to whom test materials may be supplied. To access materials requires both general training in test use and often familiarisation training for the specific instruments in question. Publishers usually give some recognition to training from other reputable providers. There are several recognised levels of training.

Test Administration

Typically based on a short (two day) training course, this qualifies the individual to carry out the standard procedures for administering and scoring tests. It does not cover the choice of tests or the interpretation of results. Neither does it allow the individual access to materials. Someone trained to this level may only work under the close supervision of a fully trained test user within the organisation. Some organisations rely heavily on test administrators to relieve more highly trained staff. SHL suggests a ratio of one fully trained test user to three test administrators to ensure proper supervision at all times.

Level 1, Occupational Testing

The British Psychological Society (BPS) has defined a set of essential competencies constituting a minimum standard for any test user (BPS Level A Certificate of Competence). These competencies cover the basic principles of psychometric testing and the techniques of selection, administration, scoring, interpretation and validation of ability tests. Typically based on a five day training course, this qualifies the individual to use ability tests and interest inventories.

Level 2, Personality Measurement

This covers the theoretical and practical issues involved in the use of measures of personality, including scoring, interpretation, feedback and applications. Typically a five plus one day training course is required, in addition to Level 1 training, in order to qualify an individual in the use of a personality instrument, such as the OPQ. The BPS Intermediate Level B certificate is available for those who have completed this level of training. Further familiarisation training is then required for each additional instrument to be used.

Level 3

The BPS full level B certificate is available to those with a deeper and broader perspective on personality and general test use, as well as full training in at least two different measures of personality. It is recommended for those who have responsibility for other test users and policy decisions. SHL can help those interested to gain the relevant competence.

Chartered Psychologists

Professional occupational psychologists who specialise in testing and assessment have a greater knowledge and experience of the use of tests than can be gained from short training courses. Chartered psychologists have reached a standard of training and experience required by the BPS and have agreed to abide by the BPS code of conduct.

Qualified users should ensure that materials are only used appropriately and are not used by untrained people or for a purpose for which they were not intended. It is also their responsibility to work within the confines of their own expertise and to recognise when refresher training, skills updates or expert advice is needed.

Example

Only trained test users who hold the relevant qualifications may use and interpret psychometric instruments. Trained test users may delegate test administration to a person trained in this area.

4.4 When should tests be used?

Tests may be used for selection, (either for shortlisting or final decisions), placement or promotion decisions, development, team building, counselling, out-placement and organisational development. It is not possible to discuss the considerations for using tests in each context in this short booklet. In each case, the situation must be evaluated to see whether test use would be appropriate to help achieve the desired objectives.

There are some occasions where it is not usually appropriate to use tests. For instance, it is unlikely that test results would be suitable for making **redundancy decisions**, since direct information on job performance should be available. (Tests can be valuable in making redeployment decisions or in outplacement counselling, however.) Similarly, an organisation may want to restrict the use of some tests to counselling or development applications.

Tests are best used in decision making, in conjunction with other relevant information. In a promotion decision, test results may be integrated with interview performance, track record and managers' recommendations to provide the best information about individual suitability. Use of a single test result alone should be avoided whenever possible.

Consideration must be given to where test results fit into a procedure. Tests may be used for shortlisting from a large pool of applicants, to suggest areas to be explored during an interview, or as a final check on the suitability of a chosen candidate.

Example

Tests may be used for selection, development and counselling purposes. Any additional uses should be referred to the Central Testing Unit for approval.

4.5 Test Choice

Whenever tests are used, it is vital that there is a match between the skills and characteristics measured and the job or organisational demands. This is particularly important when selection or promotion decisions are based on test results. Objective job analysis is the best way to determine the skills required for a particular job. These skills are then matched to appropriate tests. In large scale testing procedures, it may be appropriate to perform criterion related validation studies before test implementation.

Trained test users should have the skills to evaluate the quality and relevance to job requirements of an instrument from the information provided by the test publisher in the test manual. Where insufficient information is provided, caution should be exercised in the use of the test.

Some organisations may allow individual users to select the test they want to use. Others may require the involvement of internal or external experts in test choice. A middle way would be the provision of a general list of 'approved' tests from which individual users could choose those appropriate to their needs. This allows an organisation to prevent the use of instruments of uncertain quality. It will not control the relevance of measures - in other words the job analysis stage will still be required.

It is important that whenever tests are chosen there is written documentation of the reasons behind the choice. This may include copies of job analysis reports, job descriptions, person specifications, validation studies, etc. If the relevance of a particular measure is challenged, such evidence supports the test choice, shows the care taken and helps ensure users do not take dangerous shortcuts.

Example

All psychometric tests used must be clearly relevant to the given purpose. Detailed job descriptions and person specifications based on objective job analysis must be prepared prior to the choice of tests for any selection or promotion procedures. All decisions to use tests should be clearly documented with a copy sent to the Central Testing Unit.

4.6 Equal Opportunities

Research has shown that well-constructed psychometric tests are the single most effective predictor of job performance. Tests give objective information about a candidate and have been shown in general to lead to better and fairer employment decisions. Tests of aptitude or ability have sometimes been found to have disparate impact on ethnic or gender groups; that is proportionately fewer members of one ethnic or gender group do well in the test in some cases. In that these patterns exist, it is particularly important that appropriate guidelines are followed to avoid improper use of tests and potential contravention of anti-discrimination laws. Considerations of fairness are important in themselves. In addition, there are legal implications of unfair practices in the selection and promotion of employees.

The main issues concern choosing relevant measures and monitoring outcomes. A full discussion of all the elements of good practice can be found in *Equal Opportunities Guidelines for Best Test Practice in the Use of Personnel Selection Tests,* available from SHL. Users should take particular care when a candidate is not a native speaker of the language in which the tests are presented. Actual ability may be confounded with language proficiency and results may be difficult to interpret. Similar issues arise in the testing of candidates who are disabled. In addition, in these cases, it may be necessary to adjust standardised administration procedures to allow for the disability, as required by the Disability Discrimination Act of 1995. Test users can call the SHL Helpline for advice. *Guidelines for Testing People with Disabilities* is also available from SHL to help in this process.

Examples

The organisation is committed to selection on merit and only measures which are clearly relevant to job demands and free of extraneous bias should be used. All assessments for selection and promotion must be monitored to ensure they do not unfairly exclude or disadvantage any section of the population.

Whenever a disability prevents a suitably qualified individual from undergoing standard selection procedures, appropriate alternative arrangements for assessment must be found. Always contact the test publisher for advice before making any changes to test administration procedures.

4.7 Use of test scores

Interpretation of results should be accurate and not exceed the limits of the information. Users should beware, for example, of failing to take account of the standard error of measurement in interpreting the difference between any two scores, or of drawing far-fetched inferences from personality data.

Appropriate norms should always be used in interpreting scores. Where no suitable norm groups are provided with a test, further guidance should be sought from the publisher. Large scale test users should produce their own local norms for ability and aptitude tests. SHL will provide clients with their own norm groups, if data is submitted for analysis[1].

Tests should always be interpreted by properly trained individuals in the context of clearly defined criteria. Both quantitative and qualitative interpretation can be used. The former should be restricted to cases where there is sufficient supporting evidence. The use of fixed cut-offs with personality measures can be particularly misleading, without relevant evidence of validity. Qualitative interpretation may be more appropriate in these cases. Decision rules and their rationale should be properly documented.

Thought should be given to the integration of test results with other relevant information. This is particularly important where many different assessment techniques are used. The booklet *Guidelines for Best Practice in the Use of Assessment and Development Centres* (available from SHL) discusses the issues that arise here. However, even where only test results and an interview are combined, decisions need to be made about, for instance, whether good performance in one sphere can compensate for poor performance in the other, or whether some minimum standard needs to be reached on all elements. Similarly, in the use of personality questionnaires, users need to decide whether a criterion referenced, "danger zone" or integrative approach is to be used.

Examples

Test scores must be interpreted on the basis of relevant norm groups. Fixed cut-offs may only be imposed where specific evidence of test relevance is available (e.g. job analysis, validation study).

The interpretation of test scores is in accordance with the formula, which will be amended from time to time on the basis of ongoing validation results.

[1]This service is currently being offered free of charge provided that permission is given to publish the results anonymously.

4.8 Confidentiality and storage of results

Test results, like all personal information, should be stored with due regard to confidentiality. Access should be restricted to those with a need to know and in accordance with what has been agreed with the respondent during administration and feedback. Persons who are untrained should not be allowed access to raw data from tests, but only to clearly described interpretations.

Individuals do change and develop and so psychometric data can become less accurate over time. Test scores should therefore not be kept on file indefinitely. The time period for which scores are valid will differ depending on the nature of the measures and the particular use made of them. Care should be taken with results over 6-12 months old for selection purposes. Little reliance should be put on results over three years old for any purpose.

The Data Protection Act (1998) covers the storage of any information about an individual, whether on computer or in another form. Test results should not be stored or used without candidates' permission and they have a right to see information stored or used, or a meaningful interpretation of it.

Example

Test results should be kept by test users in locked files. A written interpretation of results should be kept in personnel files and provided to relevant individuals. Results over 12 months old are invalid for selection or promotion decisions. All results are to be destroyed after three years or when the respondent ceases to be employed, whichever is the sooner.

4.9 Responsibility to test takers

Testers should be honest and open with candidates about why the tests are being used and what will happen to the results. Candidates should be offered feedback of their results.

Before the test session

Whenever tests or questionnaires are used it is important that respondents are given clear information about the nature of the instruments and the reason for using them. Candidates should be given examples of the types of tests to be used as well as general information about the skills to be tested and practical information about the testing session. This helps to reduce anxiety and allows the candidate to prepare constructively for the session.

SHL produces a series of practice leaflets as well as complete practice tests, which organisations can supply to individuals before testing, so that they can familiarise themselves with what is required and prepare accordingly. The use of practice tests is particularly recommended for older candidates, those with less educational experience and others likely to be unfamiliar with formal testing procedures. The offer of a practice test session is highly desirable, and may be a particularly useful way of supporting internal candidates applying for promotion.

At the test session

Make sure all candidates know why and how test scores are going to be used and who will have access to the results. Test administrators should promote a serious but sympathetic atmosphere. It is important to remember that the testing session will be an extremely important event for the candidate, even if it is a routine one for the administrator. Instructions should be clear and not rushed. Administrators should ensure candidates know what they have to do before each test begins.

After the test session

Arrangements should be made to provide candidates with feedback on their results as soon after testing as possible. Personality and motivation questionnaire feedback is critical and will often enhance the interpreter's own understanding of test results. Feedback does not need to be lengthy, indeed with a large number of applicants this might be very time consuming. A face to face interview is preferred, but telephone feedback may be the only option in some circumstances.

Feedback should be given by qualified users and should be accurate and open. Profile charts may be shown to respondents, but they should not be given copies to take away. A short narrative summary may be provided if desired. This is particularly useful where testing is for counselling and development purposes.

Computer generated expert system or narrative report writers can support, but should not replace, the feedback interview. Some may be suitable to give to respondents, but many are intended as aids to interpretation for the trained test user and could easily be misinterpreted by others. Users should follow the guidelines provided by the author or publisher of such systems.

Example

This organisation is committed to dealing fairly with all candidates to be tested. We will be open and honest about the use of tests, provide suitable practice materials and relevant feedback whenever tests are used.

4.10 Retesting

An issue that arises where positions regularly become vacant and unsuccessful applicants reapply is whether they should (be allowed to) retake tests. There is no hard and fast rule but an organisation should have a consistent policy. Results can be allowed to stand for up to 12 months. It is not desirable to allow candidates to be retested regularly unless alternate forms of the test are available. However, it is reasonable to allow an applicant to be retested where there is evidence that he/she might have under-performed the first time (e.g. due to illness). Otherwise, a suitable interval should generally elapse before retesting is allowed.

Example

For selection purposes a candidate's test results are valid for any similar position for 12 months from the date of testing. Candidates may be retested after six months at their request. Candidates may be retested within a shorter period only at the discretion of the relevant personnel manager and should supply details in writing supporting their application.

4.11 Monitoring

Use of tests and other psychometric instruments should be continually monitored to ensure continued appropriateness and effectiveness. In small scale applications, this may amount to ensuring that the techniques remain relevant to the job and that up-to-date test versions and norms are used. Where larger scale use occurs, scores should be monitored at regular intervals to update norms. Monitoring by ethnic group and gender is required to identify any adverse impact. A validation study should be carried out every five years or so, or whenever changes in the job or applicant group are such that initial validity could have been affected.

Monitoring might be the responsibility of local test users, or could be centrally co-ordinated. The performance of a validation study requires detailed knowledge beyond that generally gained in basic level training courses. Advice should be sought from a competent chartered psychologist.

Example

Test monitoring forms should be completed for each exercise involving test use and sent to the Central Testing Unit for processing. Whenever more that 100 people are employed in a job category for which tests are used in selection, a validation study must be performed within three years from the commencement of the use of tests.

4.12 Access to materials

This element of a policy should cover who can buy materials, where they are stored and who has access. The security of materials is paramount. Free circulation leads to over familiarity and devalues psychometric instruments. Responsible test publishers only supply materials to trained users, who, in turn, must ensure untrained users do not gain access to them.

Within an organisation decisions should be taken about who should hold test materials and who should have access. It may not be desirable for all users to have access to all materials. Central storage can help prevent unnecessary duplication of materials but may not be practical in decentralised organisations.

An organisation must supply test users with appropriate storage space where tests can be kept under lock and key. It is highly desirable that all materials are logged in and out of test store. This helps ensure materials are not carelessly left lying around and prevents them going astray. Failure to keep track of materials can be expensive where replacements have to be purchased or annual lease fees paid on missing booklets.

Examples

Test users must ensure that all test materials are securely stored.

An accurate log should be kept of all test materials held. The log should be updated whenever materials are removed, replaced or added to the store.

4.13 Copyright

Test materials are extremely vulnerable to copyright infringement. In most countries, the reproduction of test materials by any means (including computer installations) without the permission of the author is a criminal offence, whether or not the reproduced materials are to be sold. Illegal copying of materials leads to lack of standardisation and poor control of materials, and gives respondents a bad impression. Ultimately, the resulting loss of income will contribute to less new test development, poorer updating services, or higher prices.

The responsibility for obtaining the publisher's permission to install a particular test on a generic computer "shell system" rests with the user, not the supplier, of such a system. The user will be in breach of copyright if permission from the publisher has not been obtained.

All SHL materials, including profile charts and software supplied on computer installations, are subject to copyright. SHL has in the past actively pursued potential breaches of copyright where these are discovered and will continue to do so in the future.

Example

Under no circumstances should any test materials be photocopied or installed on computer without the test publisher's express permission.

4.14 Computers in testing

Computers can provide great benefits to testing processes, in standardising administration, developing interpretation as well as managing scoring and handling data. In general, the policy considerations for using computerised tests are the same as for any other assessment medium. The test must be appropriate for its purpose. Computer based administration must still be controlled by a trained administrator. Use of the internet for remote administration can lead to uncontrolled conditions without care. Computer generated interpretation can aid trained test users, but cannot replace them. Even when a computer system takes over much of the testing procedure, its use should always be supervised by a fully trained user.

Example

Computer generated reports should be used to standardise and support interpretations. Only trained users can authorise the use of computer generated reports in a testing process. Appropriate reports can be passed to candidates or line managers, but only with the provision of a named contact in case of questions.

5.Policy Management

Thought should be given to the implementation procedure for the policy, to ensure that test users are aware of the requirements and are committed to working within them. Communication is essential in this process. Test users should understand the function of the policy in preserving high standards.

Commitment from the top is essential in implementing a policy successfully. Managers should ensure that they promote a supportive atmosphere towards the policy and actively implement it themselves.

A review procedure should form part of the policy to ensure it remains relevant to the needs of the organisation. This should specify when and how the workings of the policy will be examined and how changes are to be made.

Example

It is the responsibility of all test users to ensure that this policy is applied at all times. All deviations should be reported to the personnel manager who will take appropriate action. The personnel manager will formally review the functioning of the policy after six months initially and then every two years. Suggestions for changes and amendments should be addressed directly to the personnel manager.

6.Glossary

Ability test
A test designed to measure a specific competence such as verbal reasoning or clerical checking.

Aptitude test
A test designed to measure potential performance in a given area - usually through existing abilities.

Cut-off
The score on a test that separates those selected from those rejected (the 'pass' mark).

Disparate impact
A selection criterion has disparate (or adverse) impact when proportionately fewer members of one ethnic or gender group can meet the criterion.

Feedback
The process of reporting back to test takers their results. At its best, this is an interactive process that enhances the understanding of both the test taker and the test interpreter about the individual.

Interest Inventory
A questionnaire designed to reveal the preferences of an individual for different jobs and job areas. They are most often used in career counselling.

Job analysis
A structured examination of the tasks inherent in a job and the skills required to perform them.

Job description
A structured report of the tasks, functions and responsibilities required in a job.

Local norms
A norm group (see norms) based on a representative sample of people from the organisation in question (e.g. previous applicants for the same job).

Norms
Standard distributions of test scores based on the performance of a representative sample of a given group. Psychometric test scores are interpreted through comparison with relevant norm groups.

Person specification
A structured description of skills, abilities, characteristics and circumstances required to optimally perform a job.

Personality questionnaire
A questionnaire which looks at the typical behaviour, interpersonal style, thoughts and/or feelings of an individual.

Practice leaflet
A leaflet provided to a test taker before a testing session providing a short description of the tests to be administered and a few example items.

Practice test
Test provided to a test taker before a testing session that are similar to those to be used in the session, to allow the test taker to become familiar with test demands and test procedures.

Practice test session	A session where (practice) tests are administered under standard test conditions in order to help a test taker become familiar with test demands, conditions and procedures.
Profile chart	A display of a number of test results from the same individual showing relative high and low scores in the different areas tested.
Psychometric test	A measure of a psychological construct that produces scores which are accurate and valid.
Qualitative interpretation	Interpretation based on an integration of the content of a series of results. The meaning of an individual score will differ according to the constellation of all the other scores.
Quantitative interpretation	Interpretation based on fixed rules applied to numerical scores.
Standard error of measurement	A measure of the accuracy of test scores.
Validation	The process of investigating the extent to which an instrument measures what it is designed to measure. Often the extent to which test scores can predict current or future job performance or some other relevant criterion.

7.Useful Publications

Avoiding Sex Bias in Selection Testing: Guidance for Employers. Equal Opportunities Commission, 1988

Equal Opportunities Guidelines for Best Test Practice in the Use of Personnel Selection Tests. SHL, 1991

The Essentials of Psychological Testing. Cronback, L. Harper and Row, 1990

Guidelines for Best Practice in the Use of Assessment and Development Centres. SHL, 1993

Guidelines for Testing People with Disabilities. SHL, 1992

Guidelines for The Development and Use of Computer Based Assessments. British Psychological Society, 1999.

IPD Code on Psychological Testing. IPD, Wimbledon, 1993

Occupational Testing Course Notes. SHL, 1999

Psychological Testing: A Manager's Guide. Toplis, J, Dulewicz, V and Fletcher, C. IPD, 1997

Psychometric Testing and Visual Impairment: An Employer's Guide. RNIB, 2000

Psychological Testing: Guidance for the User. Steering Committee on Test Standards. The British Psychological Society, Leicester, 1991.

Psychometric Tests and Racial Equality: A Guide for Employers. Commission for Racial Equality, 1992.

Standards for Educational and Psychological Testing. AERA, NCME, APA, 1999.

Towards Fair Selection: A Survey of Test Practices and Thirteen Case Studies. Wood, R and Baron, H. Commission for Racial Equality, 1993.

These guidelines were prepared by Helen Baron.

SHL