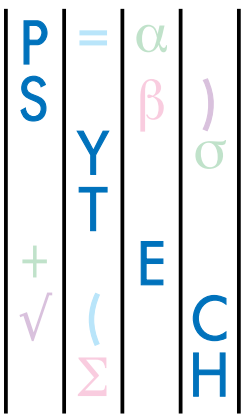
THE TECHNICAL TEST BATTERY

the technical manual

6

Test Batteries



CONTENTS

②

LIST OF TABLES

④

1 THE ROLE OF PSYCHOMETRIC TESTS IN PERSONNEL SELECTION AND ASSESSMENT

One of the main reasons for using reasoning tests to aid selection decisions is that they can provide information that cannot be obtained in other ways. If such tests are not used then what we know about the applicant is limited to the information that can be gleaned from an application form or CV, an interview and references. If we wish to gain information about a person's specific aptitudes and abilities then we have little option but to use psychometric tests. But such tests can do more than simply provide additional information about the applicant. They can add a degree of reliability and validity to the selection procedure that it is impossible to achieve in any other way. How they do this is best addressed by examining the limitations of the information obtained through interviews, application forms and references and exploring how some of these limitations can be overcome by the use of reasoning tests.

THE ROLE OF PSYCHOMETRIC TESTS IN PERSONNEL SELECTION AND ASSESSMENT

While much useful information can be gained from the interview, which clearly has an important role in any selection procedure, it does nonetheless suffer from a variety of weaknesses. Perhaps the most important of these in this respect is that it is not a reliable way to judge a person's level of reasoning ability. While the interview enables us to probe each applicant in depth and discover individual strengths and weaknesses it will not enable us to objectively assess an applicant's aptitudes and abilities.

There are similar limitations on the range and usefulness of the information that can be gained from application forms or CV's. While work experience and qualifications may be prerequisites for certain occupations, in and of themselves they do not determine whether a person is likely to perform well or badly. Past experience and academic achievement is not always a good predictor of ability or future success. While such information is important it may not be sufficient on its own to enable us to confidently choose between applicants. Thus aptitude and ability tests are likely to play a significant role in the selection process as they provide information on a person's potential and not just their achievements to date.

Little needs to be said about the usefulness of references. While past performance is undoubtedly a good predictor of future performance references are often not very good predictors of past performance. If the name of the referee is supplied by the applicant then it is likely that the applicant has chosen someone whom he expects to speak highly of him and has probably avoided supplying the names of those who may have a less positive view of his abilities. Aptitude and ability tests on the other hand give us an indication of the applicant's likely performance which is obtained under exam conditions and thus is likely to be an objective, true reflection of the person's ability.

So what advantages does the use of reasoning tests have over these other forms of assessment? The first advantage they have is that they are standardised. That is to say the same test is given to all the applicants under the same conditions and a standard method is used for scoring and interpreting the test results. Thus the test should produce the same results no matter who administers and interprets it. Moreover, the test results can be represented numerically making it easy both to compare applicants with each other, and with predefined groups (e.g. successful and unsuccess-

ful job incumbents). In addition, as noted above, they provide a range of information which is not easily and reliably assessed in other ways. Such information can fill in important gaps which have not been assessed by application forms, interviews and references and can also raise questions which can later be directly addressed in the interview. It is for this reason that psychometric tests are being increasingly used in personnel selection. Their use adds a degree of objectivity, reliability and breadth to assessment decisions which can not be achieved any other way.

⑧

2 THE CONSTRUCTION OF THE TECHNICAL TEST BATTERY

In an increasingly technical age the importance of being able to accurately assess a persons technical ability and potential cannot be overstated. The Technical Test Battery comprises three separate tests, each designed to assess a different area of technical ability. These areas are the ability to reason with mechanical concepts, the ability to manipulate three dimensional spatial relationships and the ability to quickly and accurately find a path through a complex two dimensional maze. Research has amply demonstrated that these technical abilities are not accounted for by 'general intelligence' but are specific, measurable, abilities in their own right. What is also true, though, is that general reasoning abilities should also be taken into account when considering technical ability. Verbal, numerical and abstract reasoning skills are highly important in most technical occupations and should therefore be assessed alongside technical abilities. Thus it is recommended that a test of general reasoning ability (such as the Graduate Reasoning Test (GRT1) or the General Reasoning Test (GRT2)) should be administered along with the Technical Test Battery.

1 MECHANICAL REASONING TEST

2 SPATIAL REASONING TEST

3 VISUAL ACUITY TEST

MECHANICAL REASONING TEST

The Mechanical Reasoning Test measures the ability to apply basic mechanical principles. It looks at the ability to grasp the common principles of physics which are evident in everyday life. This ability is tested over a number of different mechanical devices, e.g. gears, pulleys and levers. The mechanical reasoning test was constructed in such a way as to minimise any advantage from the possession of 'specialised knowledge'. Thus the Mechanical Reasoning Test investigates a persons ability to solve problems of a mechanical nature through the application of basic principles which most people could be expected to understand. People who do well on the Mechanical Reasoning test usually like to find out how things work. They often are better than average at learning how to construct, operate, or repair complicated equipment. People who perform poorly on the Mechanical Reasoning Test may find the work rather hard or uninteresting in physical sciences and in shop floor mechanical work which demands thinking and planning, rather than just skill in using one's hands. Many types of work in the construction and manufacturing trades also require one to understand machinery and the use of physical forces as well as to possess manual skills.

SPATIAL REASONING TEST

The Spatial Reasoning Test assesses a persons ability to manipulate and reason about shapes and spatial relationships. Unlike many spatial tests the Spatial Reasoning Test assesses the ability to work with three dimensional relationships. It looks at how well a person can visualize, or form mental pictures of solid objects from looking at flat paper plans. In other words how well can a person think in three dimensions? The Spatial Reasoning Test measures the ability to visualise, to imagine the shape and surfaces of a finished object before it is built, just by looking at the drawings that would be used to guide workers in building it. This ability makes some kinds of mathematics easier, e.g. solid geometry. To a person who does poorly on the Spatial Reasoning Test, an architect's plans for a house or an engineer's plans for a bridge or a machine might look like nothing but several flat drawings. A person who performs well on the Spatial Reasoning Test looking at those same plans can "see" the finished house, bridge or machine. He or she could probably mentally "walk around" the finished structure, looking at it from various angles. People who do well on the Spatial Reasoning Test should have an advantage in work such as draughting, dress designing, architecture, mechanical engineering, die-making, building construction, and some branches of art and decoration.

VISUAL ACUITY TEST

The Visual Acuity Test looks at a persons ability to find a path through complex two dimensional diagrams. The test basically focuses on a persons aptitude for performing tasks which require visual precision. As the test is timed, fifteen mazes have to be completed in eight minutes, it is also a test of a persons ability to perform such precision tasks under a degree of time pressure. With the rise of the new technologies visual precision is becoming an increasingly important aptitude. Thus, for instance, many workers in the microprocessor & electronic industries are required to carry out very detailed work on extremely small components and circuits. Such tasks often need to be carried out with a fair degree of speed and accuracy and the Visual Acuity test can give vital information concerning this ability.

The visual precision assessed by the Visual Acuity Test would be fairly important for a number of technical occupations; e.g. electrical engineers, mechanical and machine shop apprentices, electronic fault diagnosis, engineering draughting, etc.

The Visual Acuity Test is only available in computer-administered format using GeneSys Integrated Assessment Software.

3 PSYCHOMETRIC PROPERTIES OF THE TECHNICAL TEST BATTERY

RELIABILITY

If an ability test is to be used for selection and assessment purposes the test needs to measure each of the aptitude or ability dimensions it is attempting to measure reliably, for the given population (e.g. graduate entrants, senior managers etc.). That is to say, the test needs to be consistently measuring each ability so that if the test were to be used repeatedly on the same candidate it would produce similar results.

TEST-RETEST RELIABILITY

Test-retest reliability statistics estimate the reliability of a reasoning test by administering it repeatedly to the same applicants. If a test is reliable then we would expect it to produce consistent results when repeatedly administered over short periods of time. Thus we would not expect a reliable test to classify someone as being of high intelligence on one occasion and as being of low intelligence on another. Thus repeated test administration can provide an estimate of a test's reliability.

INTERNAL CONSISTENCY RELIABILITY

Internal consistency statistics estimate the reliability of a test by exploring whether each of the items which measure one ability or aptitude combine to produce a consistent scale. That is to say, we would expect people of superior reasoning ability to do well on all the items which form the test, and not simply on a subset of these items. If the latter were the case then we might suspect that those items which they did not perform well on were in fact not good measures of the underlying reasoning ability. These statistics are the most commonly used ways to estimate a test's reliability.

It is generally recognised that ability tests are more reliable than personality tests and for this reason high standards of reliability are usually expected from such tests. While many personality tests are considered to have acceptable levels of reliability if they have reliability coefficients in excess of .7, ability tests are not usually considered to have acceptable levels of reliability unless they have reliability coefficients in excess of .8.

THE RELIABILITY OF THE TECHNICAL TEST BATTERY

Table 1, below, presents alpha coefficients for the three tests which comprise the Technical Test Battery. Each of these reliability coefficients is greater than .8, clearly demonstrating that the individual components of the Technical Test Battery are highly reliable.

	Cronbach's Alpha
MRT	.81
SRT	.84
Visual Acuity	.80

Table 1: Cronbach's Alpha as an estimate of the reliability of the Technical Test Battery

VALIDITY

Whereas reliability assesses the degree of measurement error of a reasoning test, that is to say the extent to which the test is consistently measuring one underling ability or aptitude, validity addresses the question of whether or not the scale is measuring the characteristic it was developed to measure. This is clearly of key importance when using a reasoning test for assessment and selection purposes. In order for the test to be a useful aid to selection we need to know that the results are reliable and that the test is measuring the aptitude it is supposed to be measuring. Thus after we have examined a test's reliability we need to address the issue of validity. We traditionally examine the reliability of a test before we explore its validity as reliability sets the lower bound of a scale's validity. That is to say a test cannot be more valid than it is reliable.

There are two main ways in which we can say that a test is valid. We call these Construct Validity and Predictive Validity. When tests are used for individual assessment Construct Validity is the more important and when tests are used to predict performance the test's Predictive Validity is the more important.

CONSTRUCT VALIDITY

Construct Validity assesses whether the characteristic which the test is actually measuring is psychologically meaningful and is consistent with the scale's definition.

CRITERION-RELATED VALIDITY

This assesses whether the test is capable of predicting some agreed criterion; for example job performance. Thus while a test may have criterion-related validity it may not have Construct Validity. That is, it may predict a given criterion but may not be measuring a psychologically meaningful construct.

ASSESSING CONSTRUCT VALIDITY

Unlike reliability which can be easily measured Construct Validity is a much more difficult characteristic to assess. Rather than there being one generally agreed way to assess a test's Construct Validity the validity of a test is usually established by presenting a variety of evidence which converges to demonstrate the test's validity. For example, we will want to know that the aptitudes or abilities which the test measures are stable over time and have intuitive, consensual meaning. Moreover we will want to show that a variety of statistical properties hold for the test. These concern the test's:

INTERNAL STRUCTURE

Specifically we are concerned that the test's subscales are correlated with each other in a meaningful way. For example, we would expect the different subscales of a reasoning test to be moderately correlated as each will be measuring a different facet of general reasoning ability. Thus if such subscales are not correlated with each other we might wonder whether each is a good measure of reasoning ability. Moreover, we would expect different facets of verbal reasoning ability (e.g. vocabulary, similarities etc.) to be more highly correlated with each other than they are with a measure of numerical reasoning ability. Consequently, the first way in which we might assess the validity of a reasoning test is by exploring the relationship between the test's subscales.

CONCURRENT VALIDITY:

Here we are concerned to demonstrate that the test produces results which are consistent with those produced by other widely used, recognised, validated tests. To explore the concurrent validity of a test we would usually correlate the candidates' scores on the test which is being validated with their scores on a test which is already known to be valid.

CRITERION VALIDITY:

Here we are concerned to demonstrate that the test discriminates between criterion groups which we would predict to obtain different scores on the test's subscales. For example, we might validate a test measuring verbal and numerical reasoning ability by showing that graduates perform better on the test than non-graduates, and that science students perform better on the numerical reasoning test than arts students.

THE STRUCTURE OF TECHNICAL ABILITIES IN THE TTB

Table 2, presents Pearson Product-moment correlations between the three constituent tests of the TTB demonstrates two things. Firstly, the relatively strong correlations between each of the tests indicate that each is measuring one facet of an underlying trait. This is clearly consistent with the design of this test, where each test was intended to assess a different facet of technical ability. Secondly, the fact that each test accounts for no greater than 25% ($r <= .50$) of the variance in the other tests indicates that the Mechanical, Spatial and Visual Acuity tests of the TTB are measuring different facets of technical ability, as they were designed to.

Table 3 shows the correlations between the technical abilities measured by the TTB and Verbal, Numerical and Abstract reasoning abilities as measured by the General Reasoning Test.

Table 4, presents Pearson Product-moment correlations between the three subscales of the GRT2 and the individual tests of the TTB amply demonstrates two points. Firstly, the fairly strong correlations between the reasoning subscales and the TTB tests indicate that reasoning ability, or mental alertness, plays a role in technical abilities. This is not surprising as numerical and verbal skills are important factors in technical fields. Secondly, the fact that each subscale accounts for less than 30% ($r < .55$) of the variance in the TTB indicates that the tests which make up the Technical Test battery are in fact measuring more than just reasoning ability.

Test	MRT	SRT	VA
MRT	—	.44	.50
SRT		—	.45
Visual Acuity			—

Table 2: Product-moment Correlations between the TTB tests (n=83)

Test	VR2	NR2	AR2
MRT	0.48	0.53	.50
SRT	0.31	0.39	.45
Visual Acuity	0.37	0.42	—

Table 3: Correlations between GRT2 Sub-scales & the TTB (n = 83)

THE CONSTRUCT VALIDITY OF THE TTB

The Differential Aptitudes Test (DAT) Technical Battery is one of the most widely respected ability tests. Within this battery there are subtests designed to measure numerical, verbal, spatial and mechanical abilities. The Mechanical Reasoning Test and Spatial Reasoning Test components of the TTB have been validated against the respective components of the DAT Technical Battery.

Table 4 presents product-moment correlations (n=83) between the DAT Space Relations Test and the Spatial Reasoning Test of the TTB. As can be seen the relationship between the two tests is fairly high indicating that the two tests are indeed measuring very similar abilities.

Table 5 presents Pearson Product-moment correlations (n=67) between the DAT Mechanical Reasoning Test and the TTB Mechanical Reasoning Test. As can be seen the two tests are strongly correlated indicating that both tests are measuring very similar abilities.

A major local newspaper group with the largest number of local titles in the United Kingdom sought to examine whether tests could predict the job performance of experienced printers. A sample of 70 completed the General Reasoning Test battery (GRT2) as well as the Mechanical Reasoning Test. Each of the group were assessed on a number of performance criteria by supervisors. In addition, test data were correlated with the results of a job sample print test which was administered at selection stage.

test	SRT
DAT SRT	.73

Table 4: Correlation between the DAT Space Relations and the TTB SRT

test	MRT
DAT MRT	.75

Table 5: Correlations between the DAT MRT and the TTB MRT

	Job sample test	Initiative	Time-keeping
Mechanical (MRT2)	.42	.39	-.29
Verbal	.33	.40	
Numerical	.30	.44	
Abstract		.56	-.32

Table 6: Correlations between job performance measures and aptitudes (n=70)

4 REFERENCES

Gould, S.J. (1981). *The Mismeasure of Man.* Harmondsworth, Middlesex: Pelican.

Heim, A.H. (1970). *Intelligence and Personality.* Harmondsworth, Middlesex: Penguin.

Heim, A.H., Watt, K.P. and Simmonds, V. (1974). AH2/AH3 *Group Tests of General Reasoning; Manual.* Windsor: NFER Nelson.

Johnson, C., Blinkhorn, S., Wood, R. and Hall, J. (1989). *Modern Occupational Skills Tests: User's Guide.* Windsor: NFER-Nelson.