Optimising Assessment and Selection with Indigenous People in Remote Communities

G E Kearney and P Davidson

AUSIMM 2006 INTERNATIONAL MINE MANAGEMENT

CONFERENCE

Headings

ABSTRACT

For many years, in numerous settings, the need for a language-fair and culture-fair assessment test has been identified as a significant issue for employment and selection for training of remote indigenous community members.

Pivotal to the successful implementation and application of such a test is for it to gain the acceptance and trust of the local indigenous communities as a fair assessment and for it to provide organisations with an accurate predictive tool. However, when western tests have been used with indigenous groups, inherent language and cultural bias have had a significant negative impact in terms of community acceptance. These have also affected the accuracy of performance predictability.

The Queensland Test was developed to overcome these problems. Originally developed for use in Papua New Guinea, it reduced training failures by 20 per cent. Later refinements were made and it was used with Aboriginal Australians.

Since 2002, assessment tests have been used by the Freeport mine in West Papua to support their recruitment, selection and training processes for the local indigenous community, the Seven Suku Tribes. The initial process included the application of standard 'western' assessment tests which did not provide valid selection predictions, with community members and elders describing these tests as unfair and the organisation being concerned that the results were not accurate.

A trial of The Queensland Test was conducted and a high correlation was established between the performance of Freeport's current Seven Suku employees as assessed by their managers and their Queensland Test results. Based on the results of this pilot the Queensland test was successfully implemented as the primary selection tool for potential employees from these local indigenous groups.

The testing materials and method gained community trust and support, with tribal elders describing the Queensland Test as a fair and suitable selection tool for their community members.

Such results have sparked an interest within the Australian Mining industry, with similar results being reported from a recent trial. The Queensland Test has been implemented at a major site in Queensland and is being investigated for use elsewhere.

INTRODUCTION

When Freeport commenced mining in Papua it encountered some of the most difficult and rugged terrain in the world. Not only is the mine situated on a remote mountainside 4,200 meters above sea level with significant geological and geographical problems to face, but it is also in the homeland of a population of Papuan hunters and gatherers who had previously little or no contact with the outside world. Over time, Freeport has adapted management procedures and strategies to optimise working relationships with the local population.

Today there are approximately 19,500 employees and subcontractors working on-site. Employees are drawn from across the Indonesian archipelago. A significant number of Papuan employees are drawn directly from the province of West Papua and in particular the local tribes that surround the mining lease - Seven Suku. A small number of expatriate professionals, miners and managers support the operations. The company has an aggressive policy of employing local Papuans wherever this is possible.

The company's positive policy regarding the employment and development of local Papuans has resulted in considerable funds being invested in recruiting and developing local people.

Further, the company has established the Nemangkawi Mining Institute at its site operations. This institute is akin to an Australian technical college in the employee development services and other in-house educational programmes it offers.

ValueEdge Pty Ltd, a Melbourne-based firm of Management Consultants and Organisational Psychologists, was asked to be involved in supporting the assessment of existing and potential employees for the mine.

After a period of time, it became apparent that the traditional psychometric tests being used to assess the development potential of local Seven Suku tribe members were proving unreliable and ValueEdge was asked to assist in seeking an alternative.

ValueEdge is unusual in that it consists of experienced engineers and psychologists working together to provide unique and creative solutions to a wide range of problems in industry. When asked to address the problem of selection of people with only modest or perhaps no formal education, they pooled their knowledge to enlist

the services of an experienced psychologist who they knew had dealt with similar problems in Papua New Guinea, other parts of Asia, the Pacific and Australia.

The problems associated with making selection decisions when the applicant population is from a significantly different culture are not new. However, there has been only limited success in finding solutions to this problem. Normally, the closer the fit between the cultural and educational backgrounds of the candidate and the assessor, as well as their similarity in intellectual approach, the greater the degree of accuracy of the assessment decision. In circumstances where the candidate and the assessor come from very different backgrounds or have very different intellectual approaches, the greater the difficulty in making accurate and valid decisions. This principle is observed in a wide variety of human interactions but is especially prevalent in assessment, education and training – particularly when culturally inappropriate material or methods are used. This often leads to a negative stereotype being made of a population or of a subgroup of that population. This principle cannot be overstated, particularly in the context of employment and training of remote indigenous populations.

History

The desire to understand indigenous people was strong and evidenced from the earliest settlement. Australia was seen as a natural laboratory and almost from the very beginning of European settlement numerous scientists attempted to understand the indigenous population. As early as 1801 an expedition was commissioned by Napoleon under the leadership of Baudin to visit Australia in order to "…. convey the blessings of civilisation in the shape of tools, clothing and ornaments, and especially the most particular inventions in optics, chemistry and natural philosophy were contributed for their advantage or to promote their pleasure." Peron, a naturalist with the expedition used a dynamometer to measure the strength of representative natives of Tasmania (Kearney, 1973). Interestingly, the dynamometer is still used today as a measure of psycho-motor strength, especially by neurologists and therapists seeking to assess the degree of recovery from injury or other trauma.

The Cambridge Anthropological Expedition to the Torres Straits in 1898 was the first occasion on which scientists trained in the newly emerging discipline of psychology were to venture away from the laboratory and conduct a field expedition (Haddon, 1935).

The desire to understand local populations remained of only theoretical and academic interest for a long time. It is only when a decision must be made in relation to the distribution of scarce resources, such as the allocation of educational or training places, that assessment becomes an important consideration. Similarly, the high cost of on-the-job training and assessment may become exorbitant and there is a need to make a selection decision in favour of those candidates who are most likely to profit from the training. This principle is central to most psychological selection, especially in the commercial arena.

Equally important are the negative effects on a person or community when candidates selected for training fail because they do not have the capacity to absorb the training. The subsequent shame and blame can have a negative impact on the relationship between the organisation and the community.

Selection

Cronbach (1960) summarised the issue which is as true today as it was when it was made. He stated "… the outstanding success of scientific measurement of individual differences in behaviour has been that of the general mental test. Despite the over enthusiasm and occasional errors that have attended its development, the general mental test stands today as the most important single contribution of psychology to the practical guidance of human affairs".

Selection is most useful where there is a desire to identify a smaller number of candidates from a large applicant population without having to bear the expense of a long training trial that results in the exclusion of those unsuitable for employment. It is nearly always more efficient to identify those potentially most trainable prior to recruitment. Even in periods of universal national service there is always a considerable wastage prior to enlistment. During World War II, the Australian Army made a conscious decision not to employ those of limited ability in labour battalions, as occurred in a number of other countries. The simple cost benefit analysis demonstrated that it was more costly to supervise and train such applicants and the net drain on the overall organisation would be greater than any benefit derived.

Australian Army psychologists were asked to assist the selection of recruits into the army in Papuan New Guinea. Here, approximately a third of the known languages in the world are spoken and the general level of education was low. Because of the policy of recruiting widely from all areas, the option of limiting recruiting to only the

areas with well developed educational infrastructure was not an option. As a result, there was a relatively high wastage rate at initial training prior to the useful employment of the recruit. Not only was there very considerable cost in the recruiting process from such a wide area, but also unsatisfactory recruits were returned to their village or home as failures and this brought great shame in tribal cultures.

The research team (McElwain and Griffith, 1957) collected empirical evidence and, after discarding a large number of selection devices, were able to devise a programme that reduced the wastage rate at initial training to approximately 2 per cent. This was a highly satisfactory outcome (Ord, 1959). Later modifications were made to the selection instruments and their usefulness was extended to other populations both in New Guinea and in other parts of Asia and the Pacific. This new assessment instrument now eliminated the need for any language in its administration and therefore the need for an interpreter.

THE QUEENSLAND TEST

Utilising the learning from the research, the Queensland Test (McElwain and Kearney, 1970) (known also as the Q Test) was developed so that the task for the candidate to solve was made clearly explicit. Even without any common language or interpretation, there was no doubt about the task requirement. The performance structure of the tasks was made more portable and more robust. The modifications allowed the instrument to be administered to younger groups and children from the age of seven were able to be assessed. Testing was conducted with a variety of population samples including Aboriginal Australians, New Zealand Maori, Fijians, Gilbert and Ellice Islanders, Filipinos as well as Australians of European descent. The test is unique in that it can be administered by a member of one culture to a member of another without there being any need for mutual cultural understanding or the tester and candidate possessing any language in common. Another point of difference with this test is the emphasis on learning that takes place through practice items. The attention to practice items makes it clear that the candidate has properly understood the task required.

The terms “culture free” and "culture fair” had been used to describe assessment instruments which were claimed to be free of any cultural bias. The authors of the Queensland Test questioned whether such a description was valid or even theoretically achievable. They chose to describe the test as analogous to a broad

communication channel which allows the maximum clarity in communicating what is required to complete a cognitive task.

This led to further study being conducted on samples drawn from a population of deaf students (Kearney JE, 1969). The studies proved successful and the test was found to be quite useful with a group who had been difficult to assess previously.

The Queensland Test is best described as a non-cyclic omnibus, individually administered performance test of general cognitive capacity. It consists of six sub-tests. The individual items are constructed of portable coloured beads and tiles with which the candidate is required to construct, manipulate or recall a pattern.

The six sub-tests are:

Subtest 1: Sequential Memory

Subtest 2: Visual Memory

Subtest 3: Planning

Subtest 4: Abstract Manipulation

Subtest 5: Pattern Matching

Subtest 6: Design Sequencing

Each sub-test consists of items of increasing difficulty in which a candidate is required to demonstrate his level of cognitive performance in finding a solution. The test takes about one hour to administer and requires that the psychologist administering the test is a skilled tester. The test does not rely on formal education or the need to complete items quickly. Although there are time limits, these are generous and simply meant to find a convenient time for discontinuation.

Confirmatory Studies

It seemed that this psychological test might be useful to assist in the assessment of Papuan applicants for employment in the Freeport operation. In 2004 a sample of approximately 100 Papuan employees of Freeport were assessed using the Queensland Test. The purpose of this trial was to ensure:

- that there was nothing unusual in the perceptual or cognitive styles of the local community that would create problems in the use of the Queensland Test

- that the distribution of test results met certain psychometric standards

- that the test was valid when compared against established criteria measures

There had been certain cultural differences found in the early stages of the construction of the Queensland Test. With some cultural groups, e.g. Papuans and New Guineans, an analysis of the error patterns showed that whenever an error was made in the recall of the pattern, this error was a symmetrical response. In other cultural groups, such as Aboriginal Australians, errors made in the response sequence were asymmetrical. The utilisation of the strategy in recognizing symmetry was an important difference, but it was not the cognitive process which was being examined. Because of this difference, caused by the intrusion of a second order cognitive process, the test was redesigned to eliminate this process.

One of the most common reasons for failure in educational assessment and psychological testing occurs when the person completing the assessment does not fully understand what is required. This occurs not only when there are language differences but also whenever there are characteristically different styles of response from one culture to another. This may be caused by one population having been taught a particular paradigm as part of a formal learning process (Kearney and McElwain, 1976). For example, in the Philippines Intelligence Scale, a person is asked to identify the odd one out in the following sequence: cow, buffalo, chicken and pig. The expected answer is chicken. However a local farmer may select the cow. Alternatively the buffalo may be chosen because it is the only working animal on a farm. To the Muslim population in Mindanao the pig is an unclean animal and would be identified as the exception.

A micro analysis of a test, item by item, demonstrates which items are performing properly and which do not predict accurately. In addition, if the overall distribution of results is skewed in one direction it will indicate that the test is too easy; or if skewed in the other direction it will mean that the test is too difficult or perhaps not properly understood. This could occur if the test was given in an unfamiliar language or if unfamiliar concepts are employed. In the Freeport trial the results were distributed

normally, indicating that the test was an appropriate instrument to use for that population.

In the absence of an external criterion such as strictly comparable school results or performance on other tests of cognitive ability, the most acceptable criterion for selection is performance measured by supervisor's ratings. The measure of rating of performance by an on-the-job supervisor or in a training programme by an instructor is always a useful measure. However, some trainees perform well for one supervisor but not for another. Similarly some respond well to one instructor but not to another. Overall such measures tend to be reliable and often are the only measures that can be used to establish a satisfactory criterion. There will always be some degree of error which will cause some contamination in the analysis of validity coefficients.

In this Freeport study a correlation of $r = 0.598$ ($p < 0.001$) was established. This is a good level in such trials.

Another important feature of this test is that the candidates enjoy doing it. Tribal elders and senior members of the group reported that they considered this testing method as fair and judged it to be an acceptable and trusted process. Such endorsement had not been given to other assessment methods and examinations.

Including the trial applications of the Queensland Test and since its successful implementation as the primary selection tool for potential employees from these local indigenous groups, over 2500 candidates have been assessed.

There is an additional cost in using such a test as the Queensland Test. Because the test is individually administered by a trained psychologist, it is much more labour intensive and demanding than pencil and paper tests which can often be administered in a group format. However, there are a number of benefits that outweigh the initial cost. They are:

- that the test allows successful employee selection and retention and has an established predictive validity for populations who would not otherwise be able to be assessed

- the avoidance of selecting unsuitable candidates of limited ability who will be more costly to supervise and train, resulting in a net drain on the overall organisation

- a more positive engagement with the local community who have been shown to accept the results of assessments they consider to be fair

Second Trialling Study

Subsequent to the introduction of the Queensland Test into the selection procedure at Freeport, it was introduced to a mining company in northern Australia who have a commitment to the employment and training of a significant Aboriginal Australian population. Exactly the same procedure was followed. A pilot study was conducted on site and psychologists from the personnel department were trained in the administration of the test. The same criteria were met. The correlation between the demonstrated performance in employment and Queensland Test results was $r = 0.67$ ($p < 0.05$). As a result, the Queensland Test has been introduced as part of the selection process and is reported to be making a positive contribution to the ongoing selection process.

The Queensland Test is a proprietary test licensed by George Kearney to ValueEdge for commercial application.

Summary

These two studies together with the considerable early work with a variety of indigenous populations, demonstrate the usefulness of the Queensland Test as an assessment instrument. The test is useful for populations with limited or no formal education. There is no necessity for there to be a common language shared by the tester and the candidate. It is also a useful test for standard populations particularly those who may also have communication difficulties.

Another important feature of the test is that it has found ready acceptance by candidates who were unhappy with other forms of testing that they considered to unfairly discriminate against them.

References

Cronbach, L J 1960. *Essentials of Psychological Testing* (Harper and Row: New York)

Haddon, A C 1935, *Reports of the Cambridge University Expedition to the Torres Straits* (Cambridge University Press: Cambridge) 6 vols.

Kearney, G E 1973. Early psychological studies, in *The Psychology of Aboriginal Australians* (ed: G E Kearney, P R de Lacey, and G R Davidson. (John Wiley and Sons: Sydney)

Kearney, G E, with de Lacey, P R and Davidson, G R, (Eds.) 1973 *The Psychology of Aboriginal Australians*. (John Wiley and Sons: Sydney)

Kearney, G E, and McElwain, D W, (Eds.) 1976. Aboriginal Cognition: Retrospect and Prospect. *Proceedings of the Australian Institute of Aboriginal Studies Conference on Cognition* (Australian Institute of Aboriginal Studies: Canberra, also, Humanities Press: New Jersey)

Kearney, Jacqueline E 1969. A new performance scale of cognitive capacity for use with deaf subjects, *American Annals of the Deaf*, 114: 2-14

McElwain, D W and Griffiths, D M 1957 *Report on the possibility of using psychological procedures as an aid to recruiting in the Pacific Islands Regiment* (unpublished report), Headquarters, Northern Command: Brisbane

McElwain, D W and Kearney, G E 1970. *The Queensland Test Manual: A manual for use with assessment of general cognitive capacity under conditions of reduced communication*. (Australian Council for Educational Research: Melbourne)

Ord, I G 1959. Development of a test of cognitive capacity for indigenes of Papua-New Guinea, MA thesis (unpublished), University of Queensland, Brisbane.