

PSYCHOMETRIC TESTING IN SOUTH AFRICA: VIEWS FROM ABOVE AND BELOW

*Martin Sehlapelo and Martin Terre Blanche**

Department of Psychology

University of the Witwatersrand

PO Wits

2050

*To whom correspondence should be addressed, now at:

Dept of Psychology

UNISA

Box 392

Pretoria 0001

Email: terremj@alpha.unisa.ac.za

Abstract. Psychometrics played a central role in the development of psychology in South Africa and continues to affect more South Africans than any other branch of psychology. For this reason the transformation of psychology in South Africa will not be complete until psychometric testing practices have been transformed. Interviews with personnel practitioners, political spokespersons, trade unionists, workers and academics, revealed certain convergences and divergences between the views from above and below. Workers tended to have little trust in tests and the testing process, and to see the solution in the formulation of explicit testing policies arrived at through a process of consultation. Personnel practitioners, while acknowledging cultural bias in testing, tended to think that it could be overcome through the construction of non-biased tests. Views from above and below coincided with regard to the popularity of interviewing as an alternative to testing, the idea of learning potential assessment, and the high degree of respect for the technical expertise of psychometricians. This places a heavy burden of responsibility on experts in psychological testing to help redeem psychological testing, and psychology, in the eyes of South Africans.

It is generally agreed that South African psychology is at an historical crossroads as it tries to free itself from the Eurocentric theories, naively positivist methodologies, elitist service delivery mechanisms, and legitimation of the status quo which has characterised it in the past. However, much of the debate around the transformation of South African psychology has been characterised by an emphasis

on academic and clinical issues rather than on psychology's role in industry. The transformation of the discipline will not be complete until transformation permeates all aspects of South African psychology. In particular, we contend that there can be no transformation of psychology without transformation of psychological testing. Traditionally a Cinderella subject in academia, psychometrics has tended to become sidelined from professional debates, as is evidenced by the virtually complete absence of industrial psychologists and assessment professionals from the Forum on the role and function of psychology in a new South Africa (in September 1993), from the PsySSA founding conference (in January 1994), and more generally from the transformation process which has been in operation since early 1993. The irony of this academic apartheid is illustrated by the fact that the psychometrists' own conference in June 1993 was also in a small way concerned with transformation. The conference theme was "Psychometric Assessment: Its relevance in a rapidly changing world" with one of the papers spelling out the bottom line - "Psychometrics in a changing world: Adapt or die" (Reed, 1993).

Why is industrial psychology, and more particularly psychometrics, important? We identify two reasons for psychometrics' importance in the transformation of South African psychology: its central role in the development of South African psychology as a profession; and the extent of testing in South African society.

The development of psychology in South Africa is closely connected to the history of testing in this country (Nzimande, 1995). Arguably the first contact ordinary South Africans had with scientific psychology was through the early research of figures such as Rich (1917), who was concerned with adapting European-designed psychometric tests for use on Zulus. The pre-war history of South African psychology is dominated by psychometric work relating to the educability of "natives" and their intellectual standing relative to "Europeans" (Fick, 1929), while for several decades after the war psychometric testing in industry (particularly as practiced by the NIPR) was a central concern of academic and professional psychologists (Biesheuvel, 1949; Hudson, 1953). Through its early history, psychometrics has come to symbolise the overt and covert racism of South African psychology (cf Nzimande, 1995), and therefore must have a central place in attempts to face up to and overcome that racism.

Although since the 1960s there has been a steady decline in academic publications on psychometrics (Seedat, 1993), the practice of psychometrics has continued to grow. Some indicators of the scope of testing are that the HSRC sells more than 105 000 answer sheets annually (Holburn, 1989); in the 1970s as many as 10 000 workers were tested monthly in the mining industry alone (Veldsman, 1990); and 91,7% of large companies (1000 plus employees) use psychometric tests to select artisans (Holburn, 1989). It is thus entirely possible that psychology touches the lives of more South Africans through testing than in any other way.

Given South African psychology's intimate historical connection with psychometrics and the continued prevalence of psychometric testing in modern-day South Africa, it should obviously be an important site of transformation. The fact is that psychological tests are used on a large scale to determine who gains access to economic and educational opportunities, and if psychology as a profession is truly

interested in empowerment, the reform of testing practices should be one of its priorities.

METHOD.

What do South Africans think of the current and possible future role of psychometric testing in industry? In trying to formulate provisional answers to this question, we consulted a range of stakeholders such as human resource executives, personnel managers, political spokespersons, trade unionists, workers and academics.

In all we interviewed 46 people in the Gauteng province, of whom more than half were workers with no particular expertise as regards psychometrics, except that the majority had at one time or another been at the receiving end of testing. The racial and gender distribution in the sample of workers and trade unionists reflected the South African population (the view from below), while personnel executives, managers and academics involved in psychometrics were, with a few exceptions, white males (the view from above).

Personnel practitioners and key individuals involved in testing in industry were identified from a table of top companies compiled by the *Financial Mail* (June 1993), as well as through a snowballing process. Some workers were approached in settings such as taxi ranks and during their lunch-hour in city parks, while others were contacted through union officials and by appointment at their places of work. The group included manual labourers, factory workers, and some white collar workers.

Interviews were informal and geared towards eliciting spontaneous perceptions about psychometric testing. Our purpose was to gain an impression of the kinds of issues raised by various stakeholders, rather than to discover the proportions endorsing particular opinions. In order to encourage discussion we as far as possible asked open-ended questions such as: "Could you give us some idea of current testing practices in your company?"; "What do you see as the strong points and weaknesses of psychological testing?"; "How should people be compared to decide who gets a job?"; and "How should the way in which companies test people change in the new South Africa?"

Interviewees were not expected to present formal policy statements, but to speak anonymously in their private capacity. Most interviews were done face-to-face with individuals, although some telephonic interviews and small-group discussions were also conducted.

Interview data were qualitatively analysed by identifying recurrent themes and relating these to current debates in the academic literature. Neither the data collection nor the analysis was intended to yield generalisable conclusions, but to highlight questions for further exploration. We did not conceptualise our study as operating primarily in the "context of justification" of specific facts, but rather in the "context of discovery", and were thus less concerned with Type I and Type II errors (mistakenly rejecting or accepting hypotheses) than with what Kirk and Miller (1986) label Type III errors - namely asking the wrong kinds of questions in the first place. We therefore attempted to counter the social assumptions and

discourses which enter research at the points where scientific problems are first identified and conceptualized by starting our thought from "marginal" lives. Rather than accept the structures and needs of capitalist industry - or traditional research issues in psychometrics - as natural givens, we tried to interrogate these from the point of view of those (black people, the working class) marginalised by the system.

THEMES ARISING FROM THE INTERVIEWS.

Is testing necessary?

Perhaps surprisingly, in the light of psychometrics' entrenched position in industry, the need for testing was not accepted as a given by the majority of our interviewees, but was considered a legitimate question for debate.

In the initial stages interview responses tended to fall into two broad categories - those who rejected testing outright and those who wholeheartedly accepted it. Typical responses in the former category were:
"Tests stink - they reproduce inequality" - Trade unionist;
"Tests are irrelevant and disregard skills" - Clerical worker;
"They should take you for what you are, they should believe your referees who know you for a long time" - Granny at bus-stop.

We did not directly probe the broader philosophical issue underlying testing, namely the idea of measurable individual differences, and none of the respondents raised this as a fundamental objection to testing. Instead, objections centred around the conviction that tests are ineffective and unfair in identifying relevant individual differences.

Those more positively inclined towards testing said things such as:

"Professionals know everything, trust them" - Blue collar worker;
"Having a PhD does not prove that you're clever, I've tested people with PhDs who got lower test scores" - Evaluation consultant.

However as the interviews developed, more nuanced views tended to be articulated, indicating an acceptance of the reality of testing but with a more or less acute awareness of its imperfections. Concerns about testing were voiced on two levels - those relating to the socio-political matrix within which testing takes place, and those relating to the technical quality of tests and testing practices.

These two perceptions coincided roughly with what, in the psychometrics literature, would respectively be termed fairness and bias (Taylor, 1987b). Although we do not explicitly organise the remainder of our results in terms of these two concepts, they underlie most of the controversies canvassed.

Using relevant tests.

A very common theme in our interviews, especially with workers, was that tests are perceived as irrelevant to the actual skills required in jobs. Workers could not see the sense in being judged on their performance in pencil and paper tests which had no bearing whatsoever on day-to-day job activities. The National Union of Metalworkers (NUMSA) explicitly refer to this issue in a policy document, which states that "Tests shall be relevant ... for the job for which the applicants are being considered".

Suggestions as to how testing could become more relevant to actual job requirements, were numerous, and included reference to more direct measures of job experience and competence, interviews, and the measurement of more appropriate constructs including learning potential.

Job experience.

The first stipulation of the NUMSA code on testing states that "job related experience is more important than tests", and this sentiment was echoed by almost all the workers we spoke to. Although the "view from above" much more rarely incorporated this perspective, some test construction professionals we spoke to were interested in the idea of making more explicit provision for past experience and perhaps developing systems through which this could be quantified and combined with test scores.

Although it would be inaccurate to suggest that the personnel practitioners ignored past experience, they seemed largely unaware of what may be a widespread perception among workers that this is not given enough weight relative to test scores. Workers' unhappiness in this regard relates not only to psychometric test results, but also to what is perceived as an undue emphasis on formal qualifications rather than proven competence. An emphasis on formal but "irrelevant" criteria of this sort is often seen as simply a means of entrenching (white) privilege.

Interviewing.

As a means of gathering relevant information in arriving at employment decisions, workers tended to see interviews as having greater legitimacy than tests. In this they were supported by a considerable proportion of the personnel practitioners we spoke to, some of whom were quite ready to defend their interviewing systems as rigorous, reliable and valid. This relative harmony between views from above and below, is contradicted by much of the academic literature on tests as opposed to interviewing (Tenopyr, 1981; Anastasi, 1982), which tends to portray interviews as subjective, unreliable and lacking in validity.

A trust in interviewing is also reflected in the section of the NUMSA code dealing with interviews, which is considerably shorter than that on testing, and which makes no reference to controversial issues raised in the academic literature such as the biasing effects of the race and gender of the interviewer.

Other constructs.

To the extent that the idea of testing was accepted, a number of alternate constructs to those commonly measured were suggested. Chief among these were standardised tests of task skills in specific jobs, such as typing, bricklaying, and so on. It is interesting to contrast this with the very low frequency of use of work sample tests (Holburn, 1989) relative to psychometric tests, especially in larger companies. A possible reason for this may be the labour intensiveness of administering work sample tests - which could provide a challenge to test constructors.

Another construct sometimes mentioned is racism. Screening out white applicants on the grounds of elevated scores on tests designed to measure racism would

represent an ironic reversal of fortunes for those who, it could be argued, have thus far benefited most from psychometric testing. Some interviewees also mentioned constructs with a somewhat "African" flavour on which white applicants might fare less well. These included the ability to network, and the ability to function effectively as a member of a group rather than purely as an individual.

Potential.

Several workers mentioned the idea of trainability, which is quite similar to that of learning potential, a concept which has gained increasing popularity among academics (Shochet, 1986; Boeyens, 1989; Prinsloo, 1993; Retief, 1993; Taylor, 1987a, 1994), unionists and personnel practitioners. Such is the popularity of the learning potential approach that one company of which we are aware has suspended its psychometric testing until an adequate learning potential instrument becomes available. The popularity of learning potential in South Africa is due to its promise of providing a means of fair assessment despite unequal and inadequate educational opportunities. Even Nzimande (1995), who is highly critical of testing in general, is of the opinion that the assessment and identification of potential could be useful in the context of affirmative action and the need to redress historical imbalances.

Learning potential assessment involves pre-testing, a period of instruction and post testing. The extent to which testees benefit from the instruction is taken as indicative of their potential also to benefit from future learning opportunities. Despite its popularity there are numerous conceptual and technical measurement issues which have yet to be adequately resolved (Boeyens, 1989). Some of these are addressed in Taylor (1994).

There is also a paradox in the simultaneous endorsement of the importance of learning potential and job specific skills. While job specific skills expressly depend on previous learning, and therefore disadvantage those who have not had learning opportunities, learning potential is assumed to equalise opportunity. That these two concepts are nevertheless simultaneously supported is illustrated by the NUMSA code which in one breath calls for the use of "Tests of learning potential as well as job specific tests such as work sample tests".

Although it is encouraging to see such a high degree of agreement among workers (who like the idea of trainability), trade unionists, academics, and personnel practitioners, a danger inherent in the learning potential discourse is that it may come to be seen as a technical solution to problems which are fundamentally socio-political in nature. To be labeled as lacking in potential is, if anything, more damning than to be considered unintelligent, and it is hard to imagine any procedure which can assign such value-laden characterisations in a neutral, "scientific" way.

Coaching.

Another approach to equalising opportunities in psychometric testing is to coach testees in test-taking techniques. The NUMSA code states: "Coaching in test-taking shall be provided for applicants unfamiliar with testing procedures. Past tests should be made available to applicants". We were intrigued by this possibility and put it to several interviewees. Most workers did not seem to have a problem

with this, but personnel practitioners were somewhat ambivalent and we heard comments such as:

"The examples provided in tests are good enough coaching" - Test administrator;

"Certain kinds of coaching are not practical for certain types of tests - it would be giving the whole thing away" - Personnel executive;

"Maybe coaching could be used in the interim until people are more used to tests" - Personnel manager.

Trauma of testing.

The academic literature on testing occasionally refers to "the dehumanising effects of mental testing" (Prinsloo, 1993) and similar sentiments were voiced by some of the professionals, who spoke in abstract terms of the traumatic effects of testing. However, we were astounded by the frequency and vehemence with which workers broached this topic. Some examples are:

"Testing adults reminds most workers of the horrible school days" - Municipal worker;

"It makes me feel as if I'm being treated like a child" - Taxi commuter;

"You spend hours and hours writing tests" - Office worker;

"They inform us late about when we will be tested and then it comes as a shock" - Worker.

Trust in testing process.

Related to this is a profound distrust of the whole testing enterprise. It is seen as a matter of routine, rather than as a serious attempt to find out more about the individual. It is also seen as a smoke-screen, designed to keep black people out. Even at a concrete level the integrity of the process is doubted. We often heard comments like:

"Why should we write in pencil? They rub it out and change our answers";

"They pass whites even if they fail";

"We get the results late because they use the time to change our scores";

"We write for white people - they swop the answer sheets around".

Feedback.

A contributory factor to the absence of trust in testing and workers' reports of feeling ill at ease with the testing process is the lack of adequate feedback. Workers who had some faith in the accuracy of test scores, but were rejected for jobs on the basis of these scores described themselves as being placed in a particularly distressing position - while the company possessed information about their weaknesses, they themselves were not informed and were thus unable to rectify their presumed inadequacies. A municipal worker explained: "If a person is lacking he or she should be improved"; while a programmer thought that tests should not be used to exclude people but to "show a person areas that can be improved".

The importance of "diagnostic" testing, aimed at capitalising on strengths and correcting weaknesses rather than blanket rejection or acceptance of applicants, is fairly frequently mentioned in the literature (Taylor, 1987a) and it is clear that this idea would find considerable favour with workers.

Bias.

In the light of the abundant literature on test bias, we will not discuss the issue at any length here apart from recording that there was virtual unanimity amongst the people we spoke to that many psychometric tests are biased against blacks, with rather obvious forms of bias such as inappropriate language being most often mentioned. A rather worrying tendency we noted, however, was for some personnel practitioners to believe that they had in their possession a system which was immune against bias. Whether it was a battery of imported American tests or an interviewing system which supposedly reveals "potential", several practitioners believed that the techniques they used had somehow overcome or circumvented the problems of bias. Others were eager to lay their hands on such technology, which they believed would soon be invented.

An interesting question which arose in this regard was who should be responsible for constructing and funding a new generation of unbiased tests? We found general agreement that there should be some form of partnership between the state and private enterprise in test development. The state was most often identified as the appropriate agency to determine standards and provide funding, while some respondents thought that private enterprise should perform the actual research required. Although the need for a central watch-dog and funding body was widely recognised, the HSRC and TCRSA were often portrayed as illegitimate organisations. People (including HSRC researchers) spoke of the white male dominated structure of the HSRC (a situation which has recently started to change) and of its tainted history.

Some interviewees were willing to contemplate the restructuring of the HSRC but many felt that a new body should be formed or that the HSRC's functions in respect to the development of tests for industry should be taken over by other organisations such as the National Training Boards (NTBs) or universities.

Testing Policy.

Despite recent expositions in the psychometrics literature (Taylor & Radford, 1986; Donald et al, 1990; Veldsman, 1990) on the importance of formulating explicit company policies regarding testing, most of the personnel practitioners we spoke to candidly admitted that their companies did not have any policy, apart from a general commitment to the efficacy of testing. Workers were unanimous that if their companies had testing policies, they were not aware of them. A few companies appeared to have informal policies, determined by top management and assessment professionals.

In contrast to this state of affairs, unionists and workers were clear in their insistence that in future each company should have a clearly articulated testing policy and that this should be formulated in consultation with all affected parties. From our discussions with union officials it appears that an opinion emerging in labour circles is that national policies should be negotiated at the level of NTBs, policies for particular industries should be framed by Industrial Training Boards (ITBs), and policies for individual companies should be negotiated between the company and worker representatives. The NUMSA code on testing quoted earlier is an example of the kind of position which worker leaders are likely to adopt in such negotiations.

Interestingly, interviewees spent considerably less time discussing the possible content of policies than the process of policy formulation. With the exception of a few individuals who were happy to leave the matter to experts, workers and trade-unionists emphasised the importance of a consultative process and the need for creating balanced negotiating forums in terms of workers vs management, as well as race and gender. This is in contrast to the expert literature on the topic (Taylor & Radford, 1986; Donald et al, 1990; Veldsman, 1990) and several of the personnel practitioners we spoke to, who had much to say about the content of policy and were at most luke-warm about instituting a process of consultation.

CONCLUSION.

The assumption that personnel practitioners are uniformly pro-testing and that workers reject testing as always contrary to their class interests could not be supported from the opinions expressed by the stakeholders in our study. Workers were highly critical of many tests and testing practices - the irrelevance of the constructs measured, the dehumanising effects of testing, the lack of feedback, and so on - but were willing to accept rational testing policies which result from an intensive process of consultation.

The view from above (as we could reconstruct it from our interviews) was somewhat different. Almost without exception personnel practitioners appear to realise the need to transform psychometrics. However at times they seem to view the transformation in overly technical terms. There appears to be a real expectation that many of the problems can be solved by switching to interviews, constructing unbiased tests, using learning potential tests, developing more intricate selection strategies, and so on. At a deeper level, however, we sensed an appreciation of the fact that as cultural products tests can never be constructed to be truly "culture free", and that a technical solution to the problems of testing in a multi-cultural society will always remain elusive.

What both workers and managers had in common was an acceptance of the role of psychological assessment experts. This obviously places an enormous burden of responsibility on such experts, who - especially if they function as paid consultants to management or unions - could easily fall into the trap of purveying "easy solutions". On the positive side, this places them in a unique position to begin to redeem psychometrics and with it the practice of psychology in South Africa. However, following Nzimande (1995), we would argue that broader questions regarding the place of psychometrics in industry and society cannot be answered by psychometricians alone. The limited stakeholder survey reported on here has shown that it is worthwhile consulting those not normally considered to have informed opinions on psychometric testing, but who are nevertheless affected by it. This is not only important from the point of view of democracy and fair labour practice, but can have direct utility in improving testing policies.

REFERENCES.

Anastasi, A (1982) *Psychological testing*, 6th ed. New York: MacMillan.

- Biesheuvel, S (1949) Recent advances in personnel research in South Africa. *Journal of the South African Institute of Personnel Management*, 4(1), 6-20.
- Boeyens, J (1989) Learning potential: A theoretical perspective. PERS-432. Pretoria: Human Science Research Council.
- Donald, C et al (1990) Turning fairness in selection and placing into practical company policy. *South African Journal of Labour Relations*, 14(4), 52-77.
- Duckitt, J (1991) The development and validation of a subtle racism scale in South Africa. *South African Journal of Psychology*, 21(4), 233-239.
- Duckitt, J (1993) Further validation of a subtle racism scale in South Africa. *South African Journal of Psychology*, 23(3), 116-119.
- Fick, M L (1929) Intelligence test results of poor white, Native (Zulu), Coloured and Indian school children and the educational and social implication. *South African Journal of Science*, 26, 904-920.
- Financial Mail (1993) Top Companies. Supplement to the *Financial Mail*, June 1993.
- Forum (1993) Provisional list of organisations to be invited. Forum on the role and function of psychology in a new South Africa. Holiday Inn Garden Court, 2-4 September 1993.
- Holburn, P (1989) Apprentice selection: An HSRC/NTB survey of policies and methods used in the RSA with an emphasis on psychometric testing. C/PERS 406. Pretoria: Human Sciences Research Council.
- Hudson, W (1953) The occupational classification of Africans. Unpublished doctoral dissertation. University of the Witwatersrand, Johannesburg.
- Kirk, J & Miller, M L (1986) *Reliability and validity in qualitative research*. Beverly Hills: Sage.
- NUMSA (1992) Code of practice to end unfair discrimination in employment practices. Johannesburg: NUMSA.
- Nzimande, B (1995) To test or not to test? Paper presented at the Congress on Psychometrics, CSIR, Pretoria, 5-6 June 1995.
- Prinsloo, M (1993) The measurement of cognitive processes and potential. Paper presented at the Congress on Psychometrics for Psychologists and Personnel Practitioners, Midrand, 7 & 8 June 1993.
- Reed, R (1993) Psychometrics in a changing world: Adapt or die. Paper presented at the Congress on Psychometrics for Psychologists and Personnel Practitioners, Midrand, 7 & 8 June 1993.

Retief, A (1993) Assessing managerial potential in a multicultural society. Paper presented at the Congress on Psychometrics for Psychologists and Personnel Practitioners, Midrand, 7 & 8 June 1993.

Rich, S G (1917) Binet-Simon tests on Zulus. *South African Journal of Science*, 14, 477-482.

Shochet, I M (1986) Manifest and potential performance in advantaged and disadvantaged students. Unpublished PhD dissertation, University of the Witwatersrand, Johannesburg.

Seedat, M (1993) Topics, trends and silences in South African Psychology, 1948-1988: Ethnocentrism, crisis and liberatory echoes. Unpublished doctoral dissertation, University of the Western Cape, Cape Town.

Taylor, T R (1987a) The future of cognitive assessment. PERS-420. Pretoria: Human Sciences Research Council.

Taylor, T R (1987b) Test bias: The roles and responsibilities of test user and test publisher. PERS-424. Pretoria: Human Sciences Research Council.

Taylor, T R (1994) A review of three approaches to cognitive assessment, and a proposed integrated approach based on a unifying theoretical framework. *South African Journal of Psychology*, 24(4), 184-193.

Taylor, J M & Radford, E J (1986) Psychometric as an unfair labour practice. *South African Journal of Psychology*, 16(3), 79-87.

Tenopyr, M L (1981) The realities of employment testing. *American Psychologist*, 36(10), 1120-1127.

Tollman, S G & Msengana, N B (1990) Neuropsychological assessment: Problems in evaluating the higher mental functioning of Zulu-speaking people using traditional western techniques. *South African Journal of Psychology*, 20(1), 20-24.

Veldsman, T H (1990) Necessity for a comprehensive psychometric policy within an organization. *Journal of Industrial Psychology*, 16(1), 24-30.