

An outsider's perspective on selection fairness in South Africa

Ype H. Poortinga (Prof Em)
Tilburg University, Netherlands and
University of Leuven, Belgium

Introduction

Professional associations of psychologists like ACSG and SIOPSA are committed to the advancement of South Africa as a Rainbow Nation

The question is raised in how far this commitment should lead in personnel selection to "enforceable rules" of fairness as well as "aspirational goals" (terminology of APA's Ethical Code)

Definition of fairness

Fairness according to SIOPSA guidelines

"There are multiple perspectives on fairness. There is agreement that issues of equitable treatment, predictive bias, and scrutiny for possible bias when subgroup differences are observed, are important concerns in personnel selection; there is not, however, agreement that the term 'fairness' can be uniquely defined in terms of any of these issues" (Guidelines, 2005, p. 26)

Five meanings:

- equal group outcomes (e.g., passing score rates are the same)
- equitable treatment in administration conditions for all testees
- comparable opportunity to learn the subject matter of the test
- **absence of predictive bias** (same test-criterion regression function applies to different groups)
- **perceptions** (of fairness) among applicants

Fairness as a professional issue

Selection decisions/recommendations have three levels:

1. Individual

Maximization of individual interests amounts to placement

2. Organization

Typically there is maximization of employer interests/utility through top-down selection, under the challengeable assumption of equal utility of hiring for all applicants

3. Society at large

Utility considerations include:

(i) socio-political factors, notably promotion of development of "minorities" (e.g., quota hiring, "banding")

(ii) economic factors (e.g., productivity of the workforce)

Fairness as a professional issue (cont)

The responsibility of professionals includes the three levels of individual, organization, and society

Parameters in a fairness equation covering the levels could include:

- rates of unemployment in designated groups
- estimated organizational benefits of selection (Hunter & Schmidt)
- costs of loss of productivity through less qualified workers
- (prior) disadvantages of individuals/groups, e.g., in education

Admittedly, the approach is difficult and cumbersome

- (i) Numerical estimates are needed for parameter weights
- (ii) Weights may vary with categories of jobs and applicants

A societal negotiation process will have to decide on parameters and parameter values to be included in selection equations

Fairness as a professional issue (cont)

It is a responsibility of the psychological profession to engage major stakeholders in a discussion on making fairness concrete, including trade unions (representing workers), employers, selection professionals and government representatives

For professionals the balance of utilities between the three levels should be defined in terms of "enforceable rules" requiring a unique a priori definition of fairness for each selection decision (or category of decisions)

"If the equity objective of the Act is to be reached, we must commit to a specific interpretation of selection fairness and stop hiding behind the protest that it is impossible to produce definitive constitutive and operational definitions of selection fairness" (Theron, 2007)

Choosing variables for selection

Rationale

Is “extraversion” more relevant in SA than “ubuntu”?

Test selection is based on job analysis; requirements tend to be defined in terms of existing (i.e., western) constructs

These do not include “ubuntu”, “interpersonal relatedness”, or other concepts with non-western roots

Ultimately, inclusion of variables has to be justified by work domain relevance and this will require local research

In the meantime there is a dilemma:

- non-inclusion of local concepts can be argued on the basis of what happens to be established psychological knowledge
- inclusion can be argued on the basis of suspected conceptual bias in current conceptualizations

Cross-cultural equivalence

Does a score on an instrument (including an AC) have the same meaning across cultures (in terms of a given interpretation)?

In other words: is the score “equivalent”, is it “unbiased”?

Comparison of scores always requires that there is *conceptual equivalence* (one cannot compare apples and oranges)

Claims of non-identity across cultures have included, for example:

- less "traitedness" of personality in collectivist societies

indigenous concepts (*an-asakti, philotimo, ubuntu, amae*, etc)

Little is known about work performance relevance of indigenous concepts

Measurement equivalence

Beyond conceptual identity three levels of measurement equivalence can be distinguished:

Structural (functional) equivalence: a test measures the same trait, but perhaps not on the same quantitative scale (cf. C and F)

Metric or measurement unit equivalence: steps on a score scale have the same meaning across cultures, but there is no common scale origin (A given difference between two scores then has the same meaning for all applicants; cf. C and K)

Scale or full score equivalence: scores of a given value have the same quantitative meaning cross-culturally and can be interpreted in the same way (cf. C and C)

Table 11.4 An overview of three types of bias and their possible sources

Kind of bias	Source
Construct	<ul style="list-style-type: none">Incomplete overlap of definitions of the construct across culturesDifferential appropriateness of item content (e.g., skills do not belong to the repertoire of either cultural group)Poor sampling of all relevant behaviors (e. g., short instruments covering broad constructs)Incomplete coverage of the psychological construct
Method	<ul style="list-style-type: none">Differential social desirabilityDifferential response styles such as extremity scoring and acquiescenceDifferential stimulus familiarityLack of comparability of samples (e. g., differences in educational background, age, or gender composition)Differences in physical testing conditionsDifferential familiarity with response proceduresTester effectsCommunication problems between subject and tester in either cultural group
Item	<ul style="list-style-type: none">Poor item translationInadequate item formulation (e. g., complex wording)One or a few items may invoke additional traits or abilitiesIncidental differences in appropriateness of the item content (e. g., topic of item of educational test not in curriculum in one cultural group)

From Van de Vijver & Poortinga, 1997

Equivalence of cognitive tests, findings

The structure of intelligence is much the same, at least in literate populations competing in industrial labor markets and education systems

Full score equivalence is doubtful

- Method bias cannot be ruled out (e.g., due to differences in speed-accuracy trade-off)
- There is wide-spread evidence of item bias, or DIF, but adverse impact in terms of job criteria, tends to be limited (in the USA)
- Language is a likely major factor

On cognitive tests with standardized scoring minority group members are probably at a disadvantage, but negative impact is likely to be limited when job demands are school related

Equivalence of personality inventories, findings

Conditions of structural equivalence tend to be met across a wide range of countries (literate samples!) with any inventory, like Eysenck's EPQ ("Giant Three"), or the NEO-PI-R ("Big Five")

With a Chinese instrument (CPAI) another dimension beyond the Big Five, *Interpersonal Relatedness*, was found in China; but when an American sample was tested this factor was also found

State of the art is emphasis on (i) structural identity and (ii) (relatively) small quantitative cross-cultural differences

Potential effects of the latter differences and of method bias (e.g., due to response sets) on selection are not clear

My prediction is: Ubuntu scales will also differentiate between Whites in SA, if there is predictive validity for Blacks

Implications for selection of findings on equivalence

Inclusion of local concepts should have only small effects on job prediction, definitely as long as cognitive variables (with high predictive validity) are the main basis for selection and placement

Use of common concepts is likely to be feasible for all applicants in SA

Use of common instruments (with common norms) may well lead to (some) negative impact

By including concepts of African origin in selection perceived fairness can probably be improved

Conclusions

It is feasible to have fairness uniquely defined if the road of negotiated definition is being followed (there may be some cost in organizational utility)

This requires input from various stakeholders, especially in a heterogeneous society

Professionals in selection and evaluation are well placed to mediate the societal process

Research on equivalence has shown sufficient cultural similarity to justify assessment of the same constructs across cultures with

- parallel instruments for various groups (e.g., SAPI project, Deon Meiring in this congress)
- integrated instruments that sample items from commonly agreed upon domains across groups (e.g., international comparison of educational performance)