Journal of Industrial Psychology, 1995, 21(3), 1-6
Tydskrif vir Bedryfsielkunde, 1995, 21(3), 1-6

THE APPLICABILITY OF FAIR SELECTION MODELS IN THE SOUTH AFRICAN CONTEXT

GK HUYSAMEN

Department of Psychology
The University of the Orange Free State

OPSOMMING

Hierdie artikel verskaf 'n oorsig van verskeie modelle om billike keuring te verkry in situasies waar onderverteenwoordigende groepe geneig is om swakker op keuringstoetse te vaar. Terwyl voorspellingsydigheid 'n statistiese begrip is wat betrekking het op stelselmatige foute in die voorspelling van individue se kriteriumtellings, het keuringsbillikheid te make met die mate waarin keuringsresultate aan sekere sosiaal-politieke vereistes voldoen. Die regressie- en gelyke-risiko-modelle maak aanpassings vir verskille in die kriterium-op-toetsregressielyne van verskillende groepe. Die konstante-verhoudings, voorwaardelike-waarskynlikheids- en gelyke-waarskynlikheidsmodelle manipuleer die toetsafkappunte van verskillende groepe sodat sekere verhoudings wat tussen keuringsresultate (korrekte aanvaardings, verkeerde aanvaardings, korrekte verwerpings, verkeerde verwerpings) gevorm word, vir verskillende groepe dieselfde is. Die besluitnemingsbenadering vereis dat nutwaardes geheg word aan dié verskillende gebeurlikhede by verskillende groepe. Nie alleen is hierdie prosedures by uitnemendheid geskik om oproepe tot regstellende aksie te akkommodeer nie, maar hulle voldoen ook aan die eis van deursigtigheid.

ABSTRACT

This article reviews several models that are aimed at achieving fair selection in situations in which underrepresented groups tend to obtain lower scores on selection tests. Whereas predictive bias is a statistical concept that refers to systematic errors in the prediction of individuals' criterion scores, selection fairness pertains to the extent to which selection results meet certain socio-political demands. The regression and equal-risk models adjust for differences in the criterion-on-test regression lines of different groups. The constant ratio, conditional probability and equal probability models manipulate the test cutoff scores of different groups so that certain ratios formed between different selection outcomes (correct acceptances, correct rejections, incorrect acceptances, incorrect rejections) are the same for such groups. The decision-theoretic approach requires that utilities be attached to these different outcomes for different groups. These procedures are not only eminently suited to accommodate calls for affirmative action, but they also serve the cause of transparency.

At tertiary institutions (e.g., universities, technikons, colleges of education) and in commerce and industry there are frequently more applicants than available vacancies, or not all applicants are likely to be successful in terms of performance on the relevant criteria (e.g., examinations or job success). In situations such as these, selection may take place with a view to admitting or hiring those applicants with the greatest chance of criterion success. Such selection typically takes place on the basis of a selection instrument. For reasons of simplicity, it will be assumed that the latter is a test, but it may also be a battery of tests and/or other variables such as matriculation results or biographical variables.

After the abolition of separate tertiary institutions and job reservation for the various South African population groups, applicants with diverse academic backgrounds are competing for admission to the same tertiary institutions or for the same job vacancies. In view of the generally poorer quality of the school training that some of these groups have been exposed to, competing on an equal footing has placed members from such groups at a disadvantage.

However, if we are to abandon selection tests, there is the real danger that we may resort to procedures that are even more biased than valid tests (cf. Tenopyr, 1981). The problem of predictive bias and selection fairness has been investigated in the United States of America for several decades and various strategies to solve this problem have been proposed. Although local publications (e.g., Raubenheimer, 1983; Taylor & Radford, 1986) have made references to some of these models, to date no published study could be traced in which they have been applied to South African data. The purpose of the present paper is to explain these models in terms of a simple numerical example and to consider their applicability in the South African context.

Requests for reprints should be addressed to G.K. Huysamen, Department of Psychology, The University of the Orange Free State, P.O. Box 339, Bloemfontein 9300.

BASIC CONCEPTS

The situations to which the following selection models apply, have the following in common. There is a common criterion, for example, all students have to obtain a mark of 50% in the same examination in order to pass. Here it will be assumed that the criterion itself is relevant, reliable and unbiased for all groups – an assumption which is usually taken for granted but which may be grossly invalid. Furthermore, all applicants, irrespective of their test scores, should be admitted initially so that the correlation between the test and the criterion can be studied. In all instances it will be assumed that the correlation between the test and the criterion is significantly greater than zero because there would be no point in using an instrument that does not meet this requirement. Consequently, the scatter diagram for the test and the criterion has an elliptical shape which extends from the bottom left to the top right as depicted in Figure 1. The test cutoff score and the criterion cutoff score divide such a scatter diagram into four quadrants representing the following outcomes:

those applicants who would be rejected but who would have been successful had they been accepted (incorrect rejections) (A);

those who would be accepted and who passed the criterion cutoff score (correct acceptances) (B);

those who would be rejected and who would have failed the criterion had they been accepted (correct rejections) (C); and

those who would be accepted but who failed the criterion (incorrect acceptances) (D).

Categories B and C represent the correct selection decisions; the remaining two embody the incorrect decisions.

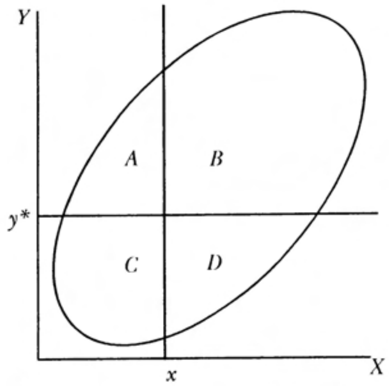

Figure 1: The four outcomes resulting from test and criterion cutoff scores.

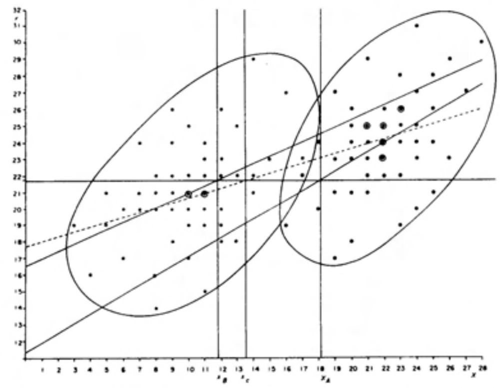

Figure 2: Scatter diagrams for two groups with regression lines differing in slope, intercept and standard error of estimate.

The issue of predictive bias and selection unfairness may arise when different scatter diagrams can be distinguished for different groups. These groups may be formed on the basis of sex, age, language, ethnicity, race, or any such demographic variable.

At the outset, the distinction between *test bias* and *predictive bias* should be clarified. A test of general knowledge containing a large proportion of items dealing with motor car accessories may be said to be biased against women if women as a group obtain a significantly lower mean than men in the test. However, in practical applications we are seldom interested in a test score in itself, but rather we wish to use it to make inferences regarding some or other criterion variable. Thus, if the present test is used to predict performance as a motor mechanic, and if men outperform women to the same degree on this criterion than they do on this test, using the test does not necessarily lead to predictive bias. Predictive bias is defined as the extent to which the regression equation used systematically underpredicts the criterion performance of one group (so that their predicted criterion scores are consistently lower than their actual criterion scores). Thus, predictive bias is present when the criterion-on-test regression line is different for different groups (Cleary, 1968). Notice that it is not the instrument itself that is biased, but an application to which it is put. Some applications of any particular test may be biased whereas others may not be.

Whereas predictive bias is defined statistically, one's definition of *selection fairness* depends on the particular system of values and/or point of departure to which one subscribes. As a result, there can be no single definition of selection fairness than can be determined statistically or objectively. Different models reflect different definitions of fairness, which in turn depend on the points of view of different parties involved in the selection process. As pointed out by Cronbach (1976), there are three such parties, namely the selector (tertiary institution or employer who usually wishes to accept only applicants who are likely to be successful on the criterion), the applicant (who wishes to get accepted), and the group to which the applicant belongs (if applicants identify strongly with the group to which they belong). As will be shown later, the latter party may be of special significance in the South African context.

The various selection models will be demonstrated in terms of the fictitious data set for two demographic groups of 50 applicants each given in Appendix 1. In practice larger samples will be required to yield stable estimates of the required test cutoff scores, but small samples are used here to keep the necessary calculations simple. The scatter diagrams and accompanying criterion-on-test regression lines for these groups appear in Figure 2. Each dot simultaneously represents the position on both the test and the criterion of an individual from the higher-scoring group (Group *A*) and each star performs the same function in respect of the lower-scoring group (Group *B*). A circle around either of these symbols indicates that there are two individuals who are occupying the same position in respect of both the test and the criterion.

THE REGRESSION MODEL

The first model that specifically addressed the present problem was proposed by Cleary (1968) and subsequently became known as the Cleary or regression model. A difference in test means may come about because of differences in school backgrounds, but by virtue of bridging programmes and/or diligent work and perseverance on the part of the lower-scoring group, the difference in criterion means may be much smaller. In such an instance, the use of the same test cutoff score, x_c, determined by a common regression line, will result in predictive bias according to the Cleary (1968) definition. For instance, for a test score of 17 in the example in Figure 2, the common regression line, indicated by means of an interrupted line, would predict a criterion score of 22,76, whereas the regression line for the higher-scoring group would predict a criterion score of 21,12 and the regression line for the lower-scoring group would predict a criterion score of 24,00. Thus, for a given test score, individuals from the lower-scoring group perform better on the criterion, on average, than those from the higher-scoring group with the same test score, yet are predicted to obtain the same criterion score. Such a common regression line would thus systematically overpredict the criterion scores for members of the higher-scoring group and underpredict the criterion scores for those of the lower-scoring group.

To rectify this bias, the regression model determines separate regression lines yielding different test cutoff scores, x_A and x_B, for the different groups to take differences in slope and/or intercept into account. Notice that the two regression lines differ in terms of slope (0,44 and 0,57 for Groups *B* and *A*, respectively), intercept (16,60 and 11,40, respectively) and standard error of estimate (2,78 and 2,67, respectively). Suppose a criterion cutoff score of 21,75 is used for the data in Appendix 1 and Figure 2. (For example, 21,75 may be considered to be the minimum criterion score that is considered to be indicative of job success.) The regression line determined for the combined groups yields a test cutoff score, x_c, of 13,54 and admits only six applicants from the lower-scoring group and all 50 of the higher-scoring group. However, if regression lines are determined separately for these groups, a test cutoff score, x_B, of 11,83 is obtained for the lower-scoring group and one, x_A, of 18,10 for the higher-scoring group. These cutoff scores now admit 15 applicants from the lower-scoring group and 46 from the higher-scoring group. In the present instance, this model will thus set a lower test cutoff score for the lower-scoring group than for the other group. This selection is fair in the sense that it selects those individuals with the highest predicted criterion scores (and, consequently, the best chance of criterion success) irrespective of their group membership.

Notice that in the present example the difference in slope of the regression lines for the two groups does not cause them to intersect inside the range of possible test scores. If the latter does happen, use of a common regression line would result in the criterion scores being underpredicted for test scores below the point of intersection and being overpredicted for test scores above this point for the very same group. Use of separate regression lines eliminates bias in both regions.

THE EQUAL RISK MODEL

A criticism which may be levelled against the regression model is that if the standard error of estimate has different values for the different groups, this model may be biased towards the group with the smaller standard error of estimate. Whereas the regression model removes bias due to differences in the slope and the intercept of the regression line, the equal risk model (Einhorn & Bass, 1971) also removes bias resulting from differences in the standard error of estimate. Figure 3 shows the distributions of prediction errors associated with a particular value of X for two such groups. If the criterion scores are distributed symmetrically and the criterion cutoff score coincides with the mean criterion score, individuals exceeding the corresponding test cutoff score have exactly the same probability of criterion success, namely, 0,50, irrespective of their group membership. However, as soon as the criterion cutoff score is lower than the criterion mean, individuals from the group with the smaller standard error of estimate have a greater probability of exceeding the criterion cutoff score than do individuals from the other group. This can be seen from Figure 3 in which a greater area of the curve for the distribution with the smaller standard error of estimate lies above (i.e., to the right of) the criterion cutoff score, y^*. (These distributions fit along the vertical axis of a figure such as Figure 2.) The regression model may be fair in the sense that individuals with the same predicted criterion score are selected or rejected irrespective of their group membership. Suppose, however, that two groups have the same predicted criterion performance but that the one group has a smaller standard error of estimate. In such an instance, selection by the regression model may be unfair towards the latter group because it has a higher probability of criterion success than the other group and yet both groups are regarded as equally attractive.

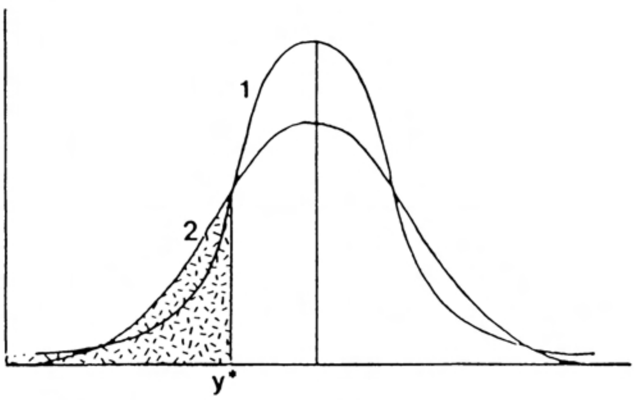

Figure 3: Conditional criterion-score distributions for a given test-score value for two groups with a common regression line but different standard errors of estimate.

The equal risk model requires that the following be specified: (a) the minimum acceptable criterion cutoff score (say, an examination mark of 50%) and (b) the lowest probability at which this criterion score would have to be passed in order to get selected. Separate test cutoff scores are then determined for the two groups so that individuals in either group whose test scores exceed these respective scores will be predicted to be above the specified criterion cutoff score with a probability that is at least equal to the specified value. For this model, the conditional distribution of criterion scores, given any test score, is assumed to be normal. Thus, to apply this model, a much larger sample of cases than that in the present numerical example is required at each test score to make this assumption tenable. With this in mind, suppose that for the present example it has been decided that only applicants who have a probability of 0,67 of passing a criterion score of 21,75 be admitted. In the case of Group A, those applicants who have a test score of 20 have a probability of 0,67 of obtaining a criterion score of 21,75 (because two-thirds of them obtained a criterion score greater than 21,75). For all test scores greater than 20, this probability exceeds 0,67, whereas for scores smaller than 20, this probability falls below 0,67. The test cutoff score which occupies the same position in respect of Group B is 13.

This model is fair in the sense that those applicants with the chosen (or higher) probability of passing the criterion cutoff score are selected irrespective of their group membership. If the groups have the same standard error of estimate, the regression model and the equal risk model yield the same results.

THE CONSTANT RATIO MODEL

Suppose the group means differ to a greater extent on the test than on the criterion as is the case for the data listed in Appendix 1 and graphically represented in Figure 2. In this example, the two groups show markedly different test means (22,00 and 10,06 for Groups A and B, respectively) but much closer criterion means (23,98 and 20,98, respectively). Because the standard deviations for both the test and the criterion for both groups are much the same (ranging from 2,56 to 3,05), the 12-point test-score difference may be said to be about four times the size of the 3-point criterion-score difference. In such an instance, the regression model will be fair as far as individuals are concerned in the sense that individuals with the required predicted criterion score are admitted irrespective of their group membership. However, it may be viewed as unfair in terms of the groups to which these individuals belong. For example, the ratio of the proportion selected to the base rate (the proportion who would have been successful if everybody was selected) (i.e., the ratio $[B + D/A + B]$ in terms of Figure 1) may be higher for the higher-scoring group than for the lower-scoring group. Notice that half as many applicants from Group B as from Group A passed the criterion. However, the test cutoff score, x_c, determined by means of the common regression line, selects only six individuals from the lower-scoring group as opposed to all 50 applicants (i.e., fewer than one-eighth as many) from Group A. More specifically, the ratio of the proportion selected to the proportion successful is equal to $0,12/0,40 = 0,30$ for Group B and $1,00/0,80 = 1,25$ for Group A. Consequently, the ratio of the proportion selected to the proportion successful overwhelmingly favours the higher-scoring group. If the cutoff scores established above by means of the regression model are used, the situation improves only slightly with the ratio involved being $0,30/0,40 = 0,75$ and $0,92/0,80 = 1,15$, for Groups B and A, respectively.

The constant ratio model proposed by Thorndike (1971) and formulated by Cole (1973), selects different test cutoff scores for the two groups so that the ratio of the proportion selected to the proportion successful is the same for both groups. Thus, if one group has twice as large a probability of criterion success as another group, then this model will select twice as many from that group as from the other. It has already been calculated that a test cutoff score of 18,01 for Group A yields 1,15 for the relevant ratio. The test cutoff score which yields the same ratio for Group B is 11,00 because for this score this ratio equals $0,46/0,40 = 1,15$.

Petersen and Novick (1976) pointed out that if the distribution of criterion scores is the same for both groups and if their regression lines are parallel (same slope but not necessarily the same intercept), then the regressin model and the constant ratio model will dictate the same test cutoff score for the two groups.

THE CONDITIONAL PROBABILITY AND THE EQUAL PROBABILITY MODELS

The conditional probability model expresses the number of correct acceptances as a proportion of the base rate, whereas the equal probability model expresses this number as a proportion of the number of applicants accepted. These models require that the respective proportions so formed be the same for the different groups. Thus, the conditional probability model (Cole, 1973) sets the test cutoff scores so that the proportion $B/(A + B)$ is the same for all groups. If a test cutoff score of 18,10 is retained for Group *A* as has been done above, this ratio equals 38/40 = 0,95 for this group. The test cutoff score for Group *B* for which the present ratio would also be 0,95, happens to be 8. (The relevant ratio for this test cutoff score for Group *B* is 19/20 = 0,95). This model is fair in the sense that it selects equal proportions of potentially successful persons from different groups.

The equal probability model discussed (but not particularly promoted) by Linn (1973) sets the test cutoff scores so that the ratio $B/(B + D)$ is equal for all groups. For a test cutoff score of 18,10 for Group *A*, this ratio would be equal to 38/46 = 0,83. The test cutoff score which yields the same ratio for Group *B* turns out to be 14, for which the present ratio equals 5/6 = 0,83. This model is fair in the sense that among those selected, the probability of criterion success is the same for all groups.

A COMPARISON BETWEEN THE PRECEDING MODELS

The regression model and equal risk model define fairness in terms of equality of opportunity for individuals in the sense that they select those individuals with the highest predicted criterion performance irrespective of their group membership. The constant ratio model, the conditional probability model and the equal probability model, by contrast, define fairness in terms of equality of some ratio of proportions of acceptees, rejectees and/or successes. As such, these three models may be characterized as quota models. The regression and equal risk models set fairness equal to lack of (statistical) bias. Such bias presents no problem to the quota models, however, because these models focus on the equality of certain quotas that need to be met irrespective of whether the tests are biased or not.

Petersen (1980) points out that in the United States the minority group tends to have lower mean criterion and mean test scores than does the majority group and that the difference between mean criterion scores tends to be smaller than the difference between mean test scores. Under these conditions, the constant ratio model will set a lower test cutoff score for the lower-scoring group than does the regression model, and the conditional probability model will set an even lower test cutoff score for this group than does the constant ratio model. This is borne out by the data in Appendix 1: For a test cutoff score of 18,10 for Group *A*, the regression model, the constant ratio model and the conditional probability model require test cutoff scores of 11,84, 11,00 and 8,00, respectively, for Group *B*. Consequently, applicants from the lower- and higher-scoring groups who have the same predicted criterion score will be equally attractive in terms of the regression model, whereas the constant ratio model will favour the applicant from the lower-scoring group and the conditional probability model will do so to an even greater extent. In the process, the latter models will accept individuals from the lower-scoring group whose predicted criterion scores fall below the criterion cutoff score and hence would not have been accepted in terms of the regression model. As a result, these models will accept individuals from the lower-scoring group who fare poorer on the criterion, on average, than do applicants from the higher-scoring group. (However, if the difference between criterion means exceeds the difference between test means, these models will assign a lower test cutoff score to the higher-scoring group). In the present numerical example, 8/46 = 17% of the individuals from Group *A* admitted by these methods in terms of a test score of 18,10 fail the criterion. By contrast, the proportions of incorrect acceptances for Group *B* in terms of the regression model, the constant ratio model and the conditional probability models are 33,33%, 43,48% and 53,66%, respectively. For these reasons, the constant ratio model and the conditional probability model were popular among institutions with policies of preferential treatment to candidates from lower-scoring racial and ethnic communities. However, the risk of damage to the self-esteem of individuals from lower-scoring groups due to failure at a tertiary institution or at a job, should be borne in mind. At the same time, it should be realized that the constant ratio model, for example, will reject some applicants from Group *A* in favour of applicants from Group *B* that have a lower probability of success on the criterion. Some would be inclined to describe such a result as being tantamount to reverse discrimination (Schmidt & Hunter, 1974), while others would argue that such actions are necessary to redress past inequities.

Petersen and Novick (1976) pointed out that the selection rule established in terms of the equal probability model usually will not coincide with that based on the regression model, the constant ratio model or the conditional probability model. For the present example, this model specifies test cutoff scores (for the two groups) that are much closer to each other than those dictated by any of the other models. Jensen (1980) pointed out that this model becomes inapplicable when the lower-scoring and higher-scoring groups differ too widely and a very high criterion cutoff score is set. Under these circumstances there simply may be no test cutoff scores that will effect equality in the ratios as required by this model.

Petersen and Novick (1976) showed that models corresponding to the converse of the constant ratio, the conditional probability and the equal probability model may be formulated in terms of the probabilities of being rejected rather than the probabilities of being accepted. These authors criticize these models because selection decisions based on them will typically not be the same as those based on their converse counterparts. For example, if the conditional probability model is considered fair towards the lower-scoring group in terms of the proportion of acceptees, the converse conditional probability model would be unfair towards the very same group in terms of the proportion of rejectees.

DECISION-THEORETIC MODELS

More recently, the above models, have been superseded by decision-theoretic approaches (Crocker & Algina, 1986). In the first place, decision-theoretic models require that consensus be reached on the relative desirabilities, called *utilities*, for each of the four different kinds of outcomes for each of the groups involved and that these be expressed quantitatively. In other words, in the case of two groups there are eight different outcomes each of which may be assigned a different utility. These approaches thus make provision for the possibility that the errors resulting from incorrect rejections (*A*) and incorrect acceptances (*D*) are not viewed as equally serious or equally objectionable, and that any one of these errors (say, incorrect acceptances) is not regarded as equally serious or equally objectionable for different groups.

Each of the models discussed above implicitly assigns utilities to various outcomes and aims to maximize the overall utility. These utilities are implicit in the characteristic chosen to be equated across groups. Thus, the equal risk model sets selection errors to be equal for the various groups. Similarly, the constant ratio and equal probability models assign lower utilities to incorrect rejections from lower-scoring groups than for incorrect rejections from higher-scoring groups, but assign the same utilities for incorrect acceptances from both groups. In the decision-theoretic models these implicit utility values have to be stated explicitly. For example, in terms of affirmative-action demands for higher numbers of admissions from previously disadvantaged groups, it may be argued that incorrect acceptances among the academically disadvantaged groups represent a less serious error than incorrect acceptances among the advantaged group. Similarly, the utility for a correct acceptance among the academically disadvantaged group may be set higher than that among the academically advantaged group. As a result, a candidate from the disadvantaged group may be accepted and one from the advantaged group may be rejected even though they have the same test score and even though candidates from the former group have a lower probability of success than candidates from the latter group with the same score. Understandably, the utilities for correct decisions are chosen to be greater than those associated with incorrect decisions. From the above, it is clear that decisions about the particular utilities that are to be used cannot be derived logically or mechanically by means of some formula. They have to be based on considerations that are mainly socio-political rather than psychometric in nature. Saywer, Cole and Cole (1976) and Novick and Lindley (1978) may provide useful suggestions in this regard.

Decision-theoretic strategies may be implemented by means of two kinds of analyses, namely, extensive-form analyses and normal-forms analyses of which the latter appears to be the more popular approach. In normal-forms analyses the probability of each outcome (*A, B, C* and *D,* in terms of Figure 1) for each of the groups involved is multiplied by its utility and summed over all four outcomes for all groups to obtain the overall expected utility. The combination of test cutoff scores for the different groups that maximizes the overall expected utility is determined. Suppose utilities of 0, 1, 1, and 0,5 are assigned to these four outcomes for the higher-scoring group and utilities of 0, 1,50, 1 and 0,25 are attached to them for the lower-scoring group in the present numerical example. These utilities imply that a correct acceptance of a candidate from Group *B* is valued more highly ($1\frac{1}{2}$ to 1) than a correct acceptance of a candidate from Group *A*, and that an incorrect acceptance from Group *A* is considered half as serious an error as an incorrect acceptance from Group *B*. Similarly, a correct acceptance of a Group *B* candidate is judged to be six times more acceptable than an incorrect acceptance from this group.

For a test cutoff score of 17 for Group *A*, the probabilities are 0,00, 0,80, 0,02 and 0,18 for the four outcomes and the corresponding probabilities for a test cutoff score of 9 for Group *B* are 0,04, 0,36, 0,24 and 0,36, respectively. By multiplying these probabilities by the above utilities, and summing the eight products so obtained, an overall expected utility of

$$1(0{,}80) + 1(0{,}02) + (0{,}50)(0{,}18) + 1{,}5(0{,}36) + 1(0{,}24) + 0{,}25(0{,}36) = 1{,}78$$

is found, which is the highest value for any combination of test cutoff scores for the two groups. Consequently, these are the cutoff scores dictated by the utilities chosen.

In extensive-forms analysis (Crocker & Algina, 1986), the average (or expected) utility is calculated for each individual for both the decision to select and the decision to reject and the one with the larger expected utility indicates the decision in respect of that individual. A candidate is accepted or rejected depending on which of these two decisions has the greater expected utility. As a result, if the utility of an incorrect acceptance is set sufficiently higher than the utility of an incorrect rejection, a candidate may be accepted even though he or she has a greater probability of failing than of succeeding on the criterion.

IMPLICATIONS FOR THE SOUTH AFRICAN SITUATION

For several reasons the above models are eminently suited to selection procedures in presentday South Africa. Evidence (e.g., Taylor & Radford, 1986) suggests that South African population groups obtain significantly different mean scores on a variety of psychological tests which are not necessarily reflected in relevant criterion measures. In the preceding sections the above models have been demonstrated in terms of two groups comparable to the so-called majority and minority groups in the United States of America. In South Africa there are at least four such 'population' groups and the majority group may be considered to be environmentally the most disadvantaged. However, the number and the relative sizes of the groups involved and their respective positions in terms of the test and the criterion have no bearing on the principles involved. (In terms of Figure 2, the numerical size of a group does not necessarily affect the two-dimensional size of its ellipse, but the numbers of cases in the cells – a feature which needs to be represented graphically along a third dimension.)

Much has been said and written about the inadequate high-school training offered by especially the former Department of Education and Training and about the disadvantage of pupils from this department when they compete for admission to tertiary institutions or for jobs on an equal footing with pupils from other education departments. Taylor and Radford (1986, p. 80) postulated that apartheid practices such as 'unequal per capita government spending on educational and social facilities, statutory and informal restrictions on the occupational and geographic mobility of certain groups, and various other consequences of the political dominance by one ethnic group of others in a heterogeneous society, are considered likely to have contributed to a differentiation in opportunities for cognitive development . . .' for different population groups. Points of view such as this virtually constitute the *raison d'etre* for calls for affirmative action. In the name of affirmative action, appeals are made for the admission of a greater number of applicants from among those who have completed their high-school training under the former Department of Education and Training. Because of the poorer selection-test performance of this very group (which is most underrepresented at tertiary institutions and in high-level jobs), calls are frequently made for the abolition of such selection tests. However, as demonstrated above, test bias does not necessarily have to be synonymous with predictive bias and/or selection fairness. There are models to ensure unbiased and fair selection procedures when valid tests with significantly different means for different demographic groups are used.

Not only are the above models eminently suited to accommodate the demands of affirmative action, they also allow such programmes to be implemented in a transparent manner. Essentially the demand for the admission of more applicants from underrepresented groups boils down to a call for assigning higher utilities to the correct acceptances and incorrect acceptances of applicants from these groups than to those from other groups. Although there may be little opposition to the call for such an increase, the question remains as to the basis on which this should be done. The decision-theoretic approach described above requires that the considerations that go into determining utilities are debated and are publicly spelled out by the relevant stakeholders. It should be realized that even if institutions do not formally disclose such utilities, they necessarily subscribe to utilities that are implicit in their selection procedures. Given that the political climate dictates that different utilities be used for different groups, the purpose of transparancy will be served by generating and explicitly stating these utilities rather than keeping them undisclosed, if not underfined altogether.

Moreover, the above models need not identify disadvantaged persons solely on the basis of race or ethnicity. In the admission of students to tertiary institutions, racial or ethnic background may simply be replaced by the education department in which an individual has matriculated. If eventually all applicants come from the same education department, degree of environmental deprivation or disadvantage may be assessed individually by means of a questionnaire such as that developed by Van den Berg (1985). (Also compare Novick & Ellis, 1977, in this regard.)

Usually a combination of predictors, which may include matriculation performance and biographical variables, is developed in an effort to increase the multiple correlation with criterion performance. It has been suggested that matriculation performance is a poorer predictor of academic success among candidates from the former Department of Education and Training than among other candidates. Another attractive feature of the above models is that the predictors used do not necessarily have to be the same for the various groups. (I.e., provided that one does not ascribe to what Schmidt &

Hunter, 1976, has termed qualified individualism, which rejects any form of differential treatment of different groups, even if this would result in poorer prediction for some groups.)

All that these models require is that the variable or combination of variables that has the maximum correlation with the criterion for a particular group be used with that group and these variables or combinations of variables may be different for various groups (Cronbach, 1976). Thus, optimum prediction of first-year performance may be achieved by using, say, test results and matriculation performance for students who have matriculated under the education departments of the former House of Assembly, and test performance and biographical variables (e.g., number of siblings) for students coming from the former Department of Education and Training.

In this country we are in the fortunate position not to have to develop the above models or to investigate their usefulness from scratch. At the same time, it would be foolish to ignore their relevance and usefulness in the South African context. Current practices such as simply accepting a much higher proportion from the black than from the white applicant group has neither any psychometric basis nor agrees with any defined notion of fairness.

APPENDIX 1: PREDICTOR (*X*) AND CRITERION (*Y*) SCORES OF GROUPS OF LOW-SCORING AND HIGH-SCORING APPLICANTS

Low-scoring group				High-scoring group			
X	Y	X	Y	X	Y	X	Y
3	19	10	21	16	19	22	24
4	16	10	22	17	22	22	25
5	19	11	21	18	20	22	25
5	21	10	25	18	24	22	26
6	17	11	15	19	17	23	19
6	20	11	19	19	21	23	22
7	21	11	20	19	23	23	24
7	20	11	21	19	27	23	25
7	24	11	22	20	18	23	26
8	14	11	23	20	21	23	26
8	16	11	24	20	23	23	28
8	20	12	18	20	24	24	20
8	21	12	20	20	25	24	23
8	22	12	22	20	26	24	24
9	18	12	23	21	21	24	25
12	19	12	26	21	22	24	27
9	20	13	22	21	23	24	31
9	21	13	25	21	24	25	21
9	22	13	18	21	25	25	24
9	24	14	21	21	25	25	26
9	26	14	22	21	29	25	28
10	17	14	29	22	22	26	23
10	19	15	23	22	23	26	29
10	20	16	27	22	23	27	27
10	21	17	23	22	24	28	30

REFERENCES

Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: CBC College Publishing.

Cronbach, L.J. (1976). Equity in selection – where psychometrics and political philosophy meet. *Journal of Educational Measurement*, **13**, 31-41.

Cleary, T.A. (1968). Test bias: Prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement*, **5**, 115-124.

Cole, N.S. (1973). Bias in selection. *Journal of Educational Measurement*, **10** 237-255.

Einhorn, H.J. & Bass, A.R. (1971). Methodological considerations relevant to discrimination in employment testing. *Psychological Bulletin*, **75**, 262-269.

Jensen, A.R. (1980). *Bias in mental testing*. London: Methuen.

Linn, R.L. (1973). Fair test use in selection. *Review of Educational Research*, **43**, 139-161.

Novick, M.R. & Ellis, D.D. (1977). Equal opportunity in educational and employment selection. *American Psychologist*, **32**, 306-320.

Novick, M.R. & Lindley, D.V. (1978). The use of more realistic utility functions in educational applications. *Journal of Educational Measurement*, **15**, 181-191.

Petersen, N.S. (1980). Bias in the selection rule – bias in the test. In L.J. Th van der Kamp & D.N.M. de Gruijter (Eds), *Psychometrics for educational debates*. New York: John Wiley.

Petersen, N.S. & Novick, M.R. (1976). An evaluation of some models for culture-fair selection. *Journal of Educational Measurement*, **13**, 3-29.

Raubenheimer, I. van W. (1983). Probleemareas by die gebruik van dieselfde of afsonderlike gestandaardiseerde toetse vir verskillende bevolkingsgroepe. In J.F. Vorster (Red.), *Simposium oor die problematiek wat ontstaan by die gebruik van dieselfde of afsonderlike toetse vir verskillende bevolkingsgroepe* (pp. 8-37). Pretoria: RGN.

Saywer, R.L., Cole, N.S. & Cole, J.W. (1976). Utilities and the issue of fairness in a decision theoretic model for selection. *Journal of Educational Measurement*, **13**, 59-76.

Schmidt, F.L. & Hunter, J.E. (1974). Racial and ethnic bias in psychological tests: Divergent implications of two definitions of test bias. *American Psychologist*, **29**, 1-8.

Schmidt, F.L. & Hunter, J.E. (1976). Critical analysis of the statistical and ethical implications of various definitions of test bias. *Psychological Bulletin*, **83**, 1053-1071.

Taylor, J.M. & Radford, E.J. (1986). Psychometric testing as a unfair labour practice. *South African Journal of Psychology*, **15**, 79-86.

Tenopyr, M.C. (1981). The realities of employment testing. *American Psychologist*, **36**, 1120-1122.

Thorndike, R.L. (1971). Concepts of culture-fairness. *Journal of Educational Measurement*, **8**, 63-70.

Van den Berg, A.R. (1985). *Using the Junior South African Individual Scales (JSAIS) (1981) for testees from South African population groups which were not included in the norm population*. Pretoria: Human Sciences Research Council.