Simultaneous Optimization of Decisions Using a Linear Utility Function

Hans J. Vos
University of Twente

The purpose of this article is to simultaneously optimize decision rules for combinations of elementary decisions. With this approach, rules are found that make more efficient use of the data than could be achieved by optimizing these decisions separately. The framework for the approach is derived from Bayesian decision theory. To illustrate the approach, two elementary decisions (selection and mastery decisions) are combined into a simple decision network. A linear utility structure is assumed. Decision rules are derived both for quota-free and quota-restricted selection–mastery decisions in case of several subpopulations. An empirical example of instructional decision making in an individual study system concludes the article.

Decision problems in educational and psychological testing can be classified in many ways. An elegant typology of test-based decisions has been given in van der Linden (1985, 1990). Each type of decision making in this typology can be viewed as a specific configuration of three basic elements—namely, a test, a treatment, and a criterion. In general, the following four different types of decision problems can be distinguished: selection, mastery, placement, and classification.

Educational applications of the four types of decision making can be found in such fields as the admission of students to schools (selection), pass–fail decisions (mastery), the aptitude/treatment/interaction paradigm in instructional psychology (placement), and vocational guidance situations where most promising schools must be identified (classification).

In Hambleton and Novick (1973), Huynh (1976a, 1977), Mellenbergh and van der Linden (1981), Novick and Petersen (1976), Petersen (1976), Petersen and Novick (1976), van der Linden (1980, 1981, 1987) and Vos (1988), these elementary decision problems have been studied ex-

309

tensively; these authors also indicate how, analytically or numerically, optimal decision rules can be found by using (empirical) Bayesian decision theory.

The four elementary decisions can be met both in their pure forms or in combinations with each other. The latter is the case, for instance, in test-based decision making in individualized study systems (ISSs). These systems can be conceived of as networks that consist of various types of decisions, as nodes (Vos & van der Linden, 1987). In such systems decision making can be viewed as the processing of students through a network of several of the elementary decisions.

The purpose of this article is the simultaneous optimization of combinations of elementary decisions using a decision–theoretic approach. Compared with separate optimization of elementary decisions, two main advantages can be identified. First, rules making more efficient use of the data can be found. Second, utility structures can be made more realistic. In order to illustrate the approach, here a selection and a mastery decision will be combined into a simple decision network, and it will be indicated how optimal rules for guiding students through such a system can be derived. The first advantage of the simultaneous approach is illustrated using this simple system. For instance, when optimizing acceptance–rejection rules in the combined decision network, pass–fail decisions to be made later can already be taken into account. The second advantage will be explained after the utility function for the combined decision has been specified.

For each elementary decision, one or more of the following restrictions may apply (van der Linden, 1990):

1. *Multiple populations.* The problem of culture-fair decision making may arise because of the presence of subpopulations reacting differently to the test items—that is, for populations defined by race or sex. In such a case, the test items are often assumed to be biased against some of the populations.

2. *Quota restrictions.* For some treatments, due to shortage of resources, the number of vacancies are constrained.

3. *Multivariate test data.* The decisions are based on data from a whole test battery instead of a single test.

4. *Multivariate criteria.* The success of the treatments is measured by multiple criteria.

Here, restrictions will only be made with respect to the presence of subpopulations and the number of students to be accepted for some treatments. First, the problem of culture-fair decision making will be considered for a quota-free selection problem. Next, optimal rules will be derived for quota-restricted selection problems using methods of constrained optimization. The final section presents some empirical examples of optimal cut-off scores for quota-free as well as quota-restricted selection-mastery deci-

310

sions. These are for two subpopulations and are referred to as the *disadvantaged* and the *advantaged* populations.

Statement of the Problem

As noted before, a well-known example of combinations of elementary decisions in education is an individualized instruction system. Figure 1 shows a flowchart of a system in which a selection decision is followed by a treatment, in this case an instructional module. Then a mastery decision follows, after which a placement decision assigns the students to two different routes through a module, both leading to the same learning objective. Real-life ISSs often have more decision points.

Selection-mastery decisions may occur in an ISS, for instance, when decisions on the admission of students to the system should be made. Then a selection test is administered before the treatment takes place and students promising satisfactory results on the criterion are accepted for the first module of the instructional program (see Figure 2). Let us suppose that the criterion is unreliably measured, which is not uncommon in ISSs. If success on the criterion is measured by a threshold value separating masters from nonmasters, then, in fact, after the treatment a mastery decision has to be taken, and the problem is a selection–mastery decision problem. Students who have reached the module objectives may proceed with the next module. However, students who failed are provided with supplemental instruction, extra learning time, corrective feedback, and the like. These students have to prepare themselves for a new mastery test.

In the following, we shall suppose that in the selection–mastery decision problem g ($g \geq 2$) subpopulations reacting differently to the test items can be distinguished. Furthermore, it is assumed that the observed selection test score variable X, the observed mastery test score variable Y, and the true score variable T underlying Y, the criterion score, assume only con-
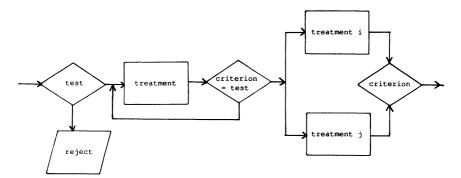

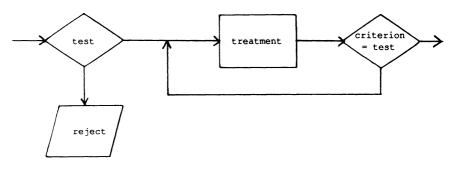FIGURE 1. *Example of an individualized study system*

311

Hans J. Vos

FIGURE 2. *A system of one selection and one mastery decision*

tinuous values. Formally, the presence of populations reacting differently to test items implies different cut-off scores for each population (Gross & Su, 1975; Petersen & Novick, 1976). Therefore, let x_{ci} and y_{ci} denote the cut-off scores for subpopulation i ($i = 1, 2, \ldots, g$) on the observed test score variables X and Y, respectively. However, the cut-off score t_c on the criterion score T is assumed to be equal for each population and is set in advance by the decision maker. The combined decision problem can now be stated as setting simultaneously cut-off scores x_{ci} and y_{ci} that, given the value of t_c, are optimal in some sense.

In the present article, following common practice in criterion-referenced testing, we consider only decisions in which the decision rules δ have a monotone form: Students are admitted to a treatment if their test score is above a certain cutting point and rejected otherwise. They can be defined for our example in the following way:

$$\delta(X, Y) = \begin{cases} a_0 & \text{for } X < x_{ci} \\ a_1 & \text{for } X \geq x_{ci}, Y < y_{ci} \\ a_2 & \text{for } X \geq x_{ci}, Y \geq y_{ci} \end{cases} \qquad (1)$$

where a_0, a_1, and a_2 stand for the actions to reject a student, to retain an accepted student, and to advance an accepted student, respectively.

An appropriate framework for dealing with decision problems such as this is Bayesian decision theory (e.g., DeGroot, 1970; Ferguson, 1967; Keeney & Raiffa, 1976; Lindgren, 1976). There are two fundamental elements in a Bayesian procedure: probabilities and utilities. In case of an ISS, a probability model predicts the outcomes of the several possible routes for the students, and a utility structure evaluates the outcomes predicted. The optimal procedure as prescribed by Bayesian decision theory is to look for a decision rule that maximizes expected utility.

With respect to the first element, it will be assumed that the relation between the measurement X of the selection test, the measurement Y of the

312

mastery test after the treatment, and the criterion variable T can be represented by a joint probability function $\Omega_i(x, y, t)$ of X, Y, and T. The experiment needed for parameter estimation of this joint probability function runs as follows: Students from the same probability function $q_i(x)$ of X are randomly drawn and accepted for the treatment, after which their performances on the mastery test Y are measured. It is important to notice that the necessary statistics for specifying $\Omega_i(x, y, t)$ come from the experiment executed as above and not, for example, from ISSs in which students are already either rejected or accepted for the treatment on the basis of their scores on the selection test in question. This is to guarantee that $q_i(x)$ and the rest of $\Omega_i(x, y, t)$ that arise in this way do not depend on whether or not a student is selected for the treatment on the basis of his or her realized value of X. With respect to this requirement, it should be realized that, although the nature of the actions shown in Figures 1 and 2 is sequential, the cut-off scores x_{ci} and y_{ci} are optimized simultaneously using data coming from the above experiment. Note that, due to the presence of different populations reacting differently to test items, different probability functions for each population should be assumed (Gross & Su, 1975; Petersen & Novick, 1976).

Also, the decision maker may have different utilities associated with different populations (Gross & Su, 1975; Petersen & Novick, 1976). Hence, in addition to separate probability distributions, the decision maker has to specify explicitly his or her utility function for each subpopulation separately.

Monotonicity Conditions

As mentioned before, in a decision-theoretic approach, optimal decision rules are found by optimizing expected utility. However, the restriction to monotone rules in our article is only correct if there are no nonmonotone rules with higher expected utility. It is here that the notion of an essentially complete class of decision rules comes in handy. An essentially complete class is defined as a class of decision rules as good as rules outside this class (e.g., Ferguson, 1967, p. 55).

In case of separate elementary decisions, the monotonicity conditions are known (Ferguson, 1967, section 6.1; Karlin & Rubin, 1956). Two conditions have to be met. First, the probability model relating observed test score Z to true score T should have a monotone likelihood ratio (MLR)— that is, it is required that for any $t_2 < t_1$, the likelihood ratio $f(z|t_1)/f(z|t_2)$ is a nondecreasing function of z. Second, the utility function should be monotone—that is, the actions should be ordered such that for each two adjacent actions the utility functions have at most one point in which the difference between the utilities changes sign. If these conditions are met, a monotone solution is said to exist. It should be noted that for the classifica-

313

tion problem these conditions do not hold without modifications (van der Linden, 1987).

To guarantee that the monotone rule of the combined decision problem belongs to an essentially complete class, the following extra condition (Lehmann, 1959, section 3.3) should hold: For any $t_2 < t_1$, the likelihood ratio $k(x, y|t_1)/k(x, y|t_2)$ is a nondecreasing function in each of its arguments— that is, for any $t_2 < t_1$ and fixed values of $Y = y_{\text{fix}}$ and $X = x_{\text{fix}}$, the likelihood ratios $k(x, y_{\text{fix}}|t_1)/k(x, y_{\text{fix}}|t_2)$ and $k(x_{\text{fix}}, y|t_1)/k(x_{\text{fix}}, y|t_2)$ are nondecreasing functions of x and y, respectively.

It will be shown that, in addition to the conditions of MLR and monotone utility, this condition is sufficient for a monotone solution to exist for the combined decision problem. The condition of monotone utility is elaborated in the next section.

Linear Utility Function for a Selection-Mastery Decision

Generally speaking, a utility function evaluates the total consequences of all possible decision outcomes. Formally, it is a function $u_{ji}(t)$ that describes the utility incurred when action a_j ($j = 0, 1, 2$) is taken for the student from subpopulation i whose true score is t. Although utility functions can be empirically assessed without making any assumptions about the form of the utility functions (e.g., Vrijhof, Mellenbergh, & van den Brink, 1983), usually the form of the utility function is specified on a priori grounds. In statistical decision theory, several forms of the utility functions have been adopted. Amongst the forms most extensively discussed in the psychometric literature belongs the threshold utility (Gross & Su, 1975; Hambleton & Novick, 1973; Novick & Petersen, 1976; Petersen, 1976; Swaminathan, Hambleton, & Algina, 1975; van der Linden, 1980, 1981, 1987) in which the utility is at a constant low level up to some threshold point t_c and then at a constant high level above that point.

An obvious disadvantage of the threshold utility function is that it assumes constant utility for students to the left or to the right of t_c, no matter how large their distance from t_c. For instance, a misclassified true master with a true score just above t_c gives the same utility as a misclassified true master with a true score far above t_c. It seems more realistic to suppose that for misclassified true masters the utility is a monotonically decreasing function of the true score t.

Moreover, the threshold utility function shows a "threshold" at the point $t = t_c$, and this also seems unrealistic in many cases. In the neighborhood of this point, the utilities for correct and incorrect decisions frequently change smoothly rather than abruptly.

As discussed earlier, the linear utility functions, just like the threshold utility functions, are specified on a priori grounds. It is, however, not known whether decision makers agree with these a priori chosen functions.

314

Fortunately, the linear utility function dealt with here seems to be a realistic representation of the utilities actually incurred in many decision-making situations. In a recent study, for instance, it was shown by Mellenbergh (1989), van der Gaag (1987), and van der Gaag, Mellenbergh, and van den Brink (1988) that many empirical utility structures could be approximated by linear functions.

In view of this, Mellenbergh and van der Linden (1981) and van der Linden and Mellenbergh (1977) propose a linear utility function for determining optimal cutting scores on the separate decisions. Here, their function is restated for the combined decision problem as a linear function in t for subpopulation i (also see Figure 3):

$$u_{ji}(t) = \begin{cases} b_{0i}(t_c - t) + d_{0i} & \text{for } X < x_{ci} \\ b_{1i}(t - t_c) + d_{1i} & \text{for } X \geq x_{ci}, Y < y_{ci} \\ b_{2i}(t - t_c) + d_{2i} & \text{for } X \geq x_{ci}, Y \geq y_{ci} \end{cases} \qquad (2)$$

where $b_{0i}, b_{2i} > 0$.

For each action a_j $(j = 0, 1, 2)$, this function consists of a constant term and a term proportional to the difference between the criterion performance t of a student and the minimum level of satisfactory criterion performance t_c. The parameters $d_{ji}(j = 0, 1, 2; i = 1, 2)$ can represent, for example, the costs of testing or the cost of following an instructional module. The values for d_{ji} are mostly negative because costs of testing are involved. Also, the values for d_{ji} might differ by group, for instance, if special test instructions have to be given for some minority groups. The condition $b_{0i}, b_{2i} > 0$ is equivalent to the statement that for the rejected students and the accepted students who passed the mastery test, utility is a strictly decreasing and increasing function of t, respectively.

Note that it cannot be said beforehand whether the utility associated with action a_1—that is, $u_{1i}(t)$—is increasing or decreasing, because the utility of the combined decision depends on the utilities associated with the selection as well as the mastery decision. Depending on whether the influence of the utility associated with the acceptance or with the fail decision is the most important, $u_{1i}(t)$ is an increasing or decreasing function of t, respectively. The parameters b_{ji} and d_{ji} $(j = 0, 1, 2; i = 1, 2)$ have to be empirically assessed (e.g., Novick & Lindley, 1978; Vos, 1988). Figure 3 displays an example of a combined linear utility function for $b_{1i} > 0$.

At the beginning of the article, I remarked that one of the main advantages of a simultaneous approach was that more realistic utility structures could be used. Formula 3 effectively demonstrates how a utility function defined on the ultimate criterion of the ISS (master or nonmaster) can be properly brought into a previous decision (selection decision).

As noted before, it will be assumed in the combined decision problem that g $(g \geq 2)$ subpopulations can be distinguished because of possible bi-
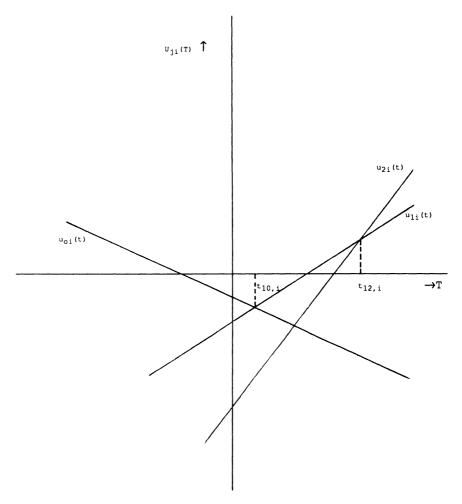
315

FIGURE 3. *Example of a linear utility function for a selection-mastery decision* $(b_{1i} > 0)$

ases in the tests implying that the optimal cutting scores, the probability functions, and the utility functions are allowed to assume a different form for each subpopulation (Gross & Su, 1975; Petersen & Novick, 1976). De Gruijter and Hambleton (1984) have discussed the problems of how inclusive a population should be and that the sample size in each subpopulation may be too small to result in the stable determination of optimal cutting scores.

The point to be noted, however, is that the necessity to distinguish subpopulations arises from the decision problem and its educational or societal

316

context (van der Linden, 1985) and not from the application of statistical decision theory. Any other method of setting cutting scores will meet the same problem as it arises. In contrast, statistical decision theory is quite able to deal with student sampling and subpopulation structures when confronted with the necessity to do so.

As Gross and Su (1975) and Petersen and Novick (1976) have argued convincingly, fair selection is a question of utilities. Whether a selection procedure is believed to be fair to the various subpopulations that can be distinguished depends on the utilities of those involved in the selection process. From this point of view, the linear utility model can be used to allow for the fact that the students might belong to a disadvantaged or advantaged subpopulation by choosing separate parameter values for the subpopulations involved (Mellenbergh & van der Linden, 1981). Suppose, for example, that subpopulation h is considered more advantaged than i. In choosing values of the parameters of the linear utility function, this can be taken into account by requiring that incorrect decisions are considered worse for subpopulation i than for h, whereas correct decisions are considered more valuable for i than for h. This amounts to choosing values of the slope parameters such that $b_{0i} > b_{0h}$ and $b_{2i} > b_{2h}$ for all t.

The possible actions are supposed to be ordered as a_0, a_1, and a_2. Using the fact that, as can be seen from Figure 3, the difference between the utilities changes sign precisely once, the condition of monotone utility for the utility function defined by Formula 2 can be expressed as

$$u_{1i}(t) - u_{0i}(t) = (b_{1i} + b_{0i})(t - t_c) + d_{1i} - d_{0i} \geq 0 \quad \text{for } t > t_{10,i}$$
$$u_{1i}(t) - u_{0i}(t) = (b_{1i} + b_{0i})(t - t_c) + d_{1i} - d_{0i} \leq 0 \quad \text{for } t < t_{10,i'} \tag{3}$$
$$u_{2i}(t) - u_{1i}(t) = (b_{2i} - b_{1i})(t - t_c) + d_{2i} - d_{1i} \geq 0 \quad \text{for } t > t_{12,i}$$
$$u_{2i}(t) - u_{1i}(t) = (b_{2i} - b_{1i})(t - t_c) + d_{2i} - d_{1i} \leq 0 \quad \text{for } t < t_{12,i'} \tag{4}$$

where $t_{10,i}$ and $t_{12,i}(t_{10,i} \leq t_{12,i})$ denote the T coordinates of the intersection of utility line $u_{1i}(t)$ with $u_{0i}(t)$ and $u_{2i}(t)$, respectively. Furthermore, it is assumed that the functions $u_{1i}(t) - u_{0i}(t)$ and $u_{2i}(t) - u_{1i}(t)$ are strictly increasing functions of t, implying that the slope parameters $(b_{1i} + b_{0i})$, $(b_{2i} - b_{1i}) > 0$. Using the fact that $b_{0i}, b_{2i} > 0$, this means that the following condition should hold for the utility parameter b_{1i}:

$$b_{2i} > b_{1i}, \quad \text{if } b_{1i} > 0$$
$$b_{0i} > -b_{1i}, \quad \text{if } b_{1i} < 0. \tag{5}$$

Optimal Cutting Scores for Quota-Free Selection

In this section, optimal cutting scores are derived for the combined decision problem in case of quota-free selection. That is, we are looking for

317

Hans J. Vos

pairs of cutting scores (x_{ci}, y_{ci}) such that the overall expected utility is a maximum.

Overall Expected Utility

In maximizing overall expected utility, first the expected utility of a random student from the ith subpopulation will be calculated, which, as monotone solutions are looked for, can be written as

$$E[u_i(T|x_{ci}, y_{ci})] = \int_{-\infty}^{x_{ci}} \int_{-\infty}^{\infty} [b_{0i}(t_c - t) + d_{0i}]w_i(x, t)\, dt\, dx$$

$$+ \int_{x_{ci}}^{\infty} \int_{-\infty}^{y_{ci}} \int_{-\infty}^{\infty} [b_{1i}(t - t_c) + d_{1i}]\Omega_i(x, y, t)\, dt\, dy\, dx \qquad (6)$$

$$+ \int_{x_{ci}}^{\infty} \int_{y_{ci}}^{\infty} \int_{-\infty}^{\infty} [b_{2i}(t - t_c) + d_{2i}]\Omega_i(x, y, t)\, dt\, dy\, dx,$$

where $w_i(x, t)$ is the joint probability function of X and T in subpopulation i. Let $E_i(T|x)$, $k_i(x, y)$, and $E_i(T|x, y)$ denote the regression function of the criterion variable T on X, the joint probability function of X and Y, and the regression function of the criterion variable T on X and Y in subpopulation i, respectively, and rearranging terms, it follows that (6) can be written as

$$E[u_i(T|x_{ci}, y_{ci})] = \int_{-\infty}^{\infty} \{b_{0i}[t_c - E_i(T|x)] + d_{0i}\}q_i(x)\, dx$$

$$+ \int_{x_{ci}}^{\infty} \{(b_{0i} + b_{1i})[E_i(T|x) - t_c]$$

$$+ d_{1i} - d_{0i}\}q_i(x)\, dx \qquad (7)$$

$$+ \int_{x_{ci}}^{\infty} \int_{y_{ci}}^{\infty} \{(b_{2i} - b_{1i})[E_i(T|x, y) - t_c]$$

$$+ d_{2i} - d_{1i}\}k_i(x, y)\, dy\, dx.$$

Now, the decision procedure is viewed as a series of separate decisions, each of which involves one random student, and it is assumed that the overall expected utility is the sum of the expected utilities for the individual students. Thus, overall expected utility of the combined decision problem is:

$$E[u(T|x_{ci}, y_{ci})] = \sum_{i=1}^{g} p_i E[u_i(T|x_{ci}, y_{ci})], \qquad (8)$$

where p_i, $\sum_{i=1}^{g} p_i = 1$ is the proportion of students from subpopulation i in the total population of students.

In quota-free selection there is no restriction as to the number of students that can be accepted for the treatment. Therefore, Equation (8) is maxi-

318

mized if the expected utility of a random student is maximized. This is done by maximizing Formula 7 for each subpopulation separately. The maximum of $E_i[u(T|x_{ci}, y_{ci})]$ now depends only on the second and third term in the right-hand side of Equation (7), because the first term is independent of x_{ci} and y_{ci}. Using a result from decision theory (e.g., Chuang, Chen, & Novick, 1981) stating that for any prior distribution of t, $E[u(T|z)]$ is a nondecreasing function of z if $f(z/t)$ has MLR and $u(t)$ is a nondecreasing function of t, and assuming the monotonicity condition from Lehmann, it follows from (4) that $E_i[u_{2i}(t) - u_{1i}(t)|x, y] = (b_{2i} - b_{1i})[E_i(T|x, y) - t_c] + d_{2i} - d_{1i}$ is a nondecreasing function in each of its arguments. Because $(b_{2i} - b_{1i}) > 0$, this implies that $(\partial/\partial x)E_i(T|x, y)$ and $(\partial/\partial y)E_i(T|x, y) > 0$. Similarly, using (3) instead of (4), it follows that $E_i[u_{1i}(t) - u_{0i}(t)|x] = (b_{1i} + b_{0i})[E_i(T|x) - t_c] + d_{1i} - d_{0i}$ is a nondecreasing function of x, implying that, because $(b_{1i} + b_{0i}) > 0$, $(d/dx)E_i(T|x) > 0$. Using $q_i(x)$, $k_i(x, y) \geq 0$, it follows now that the sign of the sum of the second and third term changes only once from negative to positive, and, therefore, $E[u_i(T|x_{ci}, y_{ci})]$ will reach its maximum for exactly one pair of cutting scores (x_{ci}, y_{ci}).

Maximizing Expected Utility for a Random Student

Necessary conditions for the optimal values of the cutting scores, say x_{ci}' and y_{ci}', optimizing the expected utility for a random student from subpopulation i, $E_i[u(T|x_{ci}, y_{ci})]$, can be obtained by differentiating $E_i[u(T|x_{ci}, y_{ci})]$ with respect to x_{ci} and y_{ci}, setting the resulting expressions equal to zero, and solving simultaneously for x_{ci} and y_{ci}.

Using the property that for any bivariate distribution $f(x, y)$, it holds that

$$\frac{\partial}{\partial s} \int_x^\infty \int_{-\infty}^s f(x, y) \, dy \, dx = -\frac{\partial}{\partial s} \int_x^\infty \int_s^\infty f(x, y) \, dy \, dx = \int_x^\infty f(x, s) \, dx.$$

For the derivative of $E_i[u(T|x_{ci}, y_{ci})]$ with respect to y_{ci} this results in

$$\frac{\partial}{\partial y_{ci}} E_i[u(T|x_{ci}, y_{ci})]$$
$$= s_i(y_{ci}) \int_{x_{ci}}^\infty \{[b_{1i} - b_{2i}][E_i(T|x, y_{ci}) - t_c] + d_{1i} - d_{2i}\} z_i(x|y_{ci}) \, dx = 0, \tag{9}$$

where $z_i(x|y_{ci})$ and $s_i(y)$ denote the posterior probability function of X given $Y = y_{ci}$ and the marginal probability function of Y in subpopulation i, respectively. Because $s_i(y) \geq 0$—the possibility of $s_i(y) = 0$ will be ignored—it follows that Equation (9) can be replaced by

$$\int_{x_{ci}}^\infty \{(b_{2i} - b_{1i})[E_i(T|x, y_{ci}) - t_c] + d_{2i} - d_{1i}\} z_i(x|y_{ci}) \, dx = 0. \tag{10}$$

Similarly, differentiating $E_i[u(T|x_{ci}, y_{ci})]$ with respect to x_{ci}, using $q_i(x) \geq 0$—the possibility of $q_i(x) = 0$ will also be ignored—results in

319

$$(b_{0i} + b_{1i})[E_i(T|x_{ci}) - t_c] + d_{1i} - d_{0i}$$

$$+ \int_{y_{ci}}^{\infty} \{(b_{2i} - b_{1i})[E_i(T|x_{ci}, y) - t_c] + d_{2i} - d_{1i}\}m_i(y|x_{ci})\ dy = 0, \tag{11}$$

with $m_i(y|x_{ci})$ being the posterior probability function of Y given $X = x_{ci}$. Now, solving the system of Equations (10) and (11) for x_{ci} and y_{ci}, one obtains the optimal cutting scores x'_{ci} and y'_{ci}.

Linear Regression

For given regression functions and probability density functions, the optimal quota-free decision strategy is represented by the system of Equations (10) and (11). If the monotonicity conditions are not strict or it does not hold that $s_i(y)$ or $q_i(x) > 0$ in the neighborhood of the solution, the optimal decision strategy may not be unique. Throughout this article it will be assumed that conditions like these are fulfilled.

Because the relations between the test scores and the criterion (true score) in the regression functions are not directly observable, psychometric models are needed to estimate these relations. Possible psychometric models are the linear regression functions $\theta_i + \Gamma_i x$ and $\alpha_i + \beta_i x + \tau_i y$ for $E_i(T|x)$ and $E_i(T|x, y)$, respectively. The monotonicity conditions $(d/dx)E_i(T|x) = \Gamma_i$, $(\partial/\partial x)E_i(T|x, y) = \beta_i$, and $(\partial/\partial y)E_i(T|x, y) = \tau_i > 0$ are fulfilled for these regression functions (e.g., Lord & Novick, 1968). Assuming linear regression, it can be shown from classical test theory that the linear regression of T on X is given by

$$E_i(T|x) = E_i(Y|x) = \mu_{Y,i} + \rho_i(\sigma_{Y,i}/\sigma_{X,i})(x - \mu_{X,i}), \tag{12}$$

$\mu_{Y,i}$, $\mu_{X,i}$, ρ_i, $\sigma_{Y,i}$, and $\sigma_{X,i}$ being the population means of Y_i and X_i, the population correlation between X_i and Y_i, and the population standard deviations of Y_i and X_i, respectively. From Equation (12), it follows that

$$\Gamma_i = \rho_i(\sigma_{Y,i}/\sigma_{X,i}), \theta_i = \mu_{Y,i} - \Gamma_i \mu_{X,i}. \tag{13}$$

Furthermore, using results from classical test theory, it can be shown that the linear regression of T on X and Y can be written as

$$E_i(T|x, y) = \mu_{Y,i} + (\sigma_{Y,i}/\sigma_{X,i})[(\rho_i - \rho_{YY',i}\rho_i)/(1 - \rho_i^2)](x - \mu_{X,i})$$

$$+ [(\rho_{YY',i} - \rho_i^2)/(1 - \rho_i^2)](y - \mu_{Y,i}), \tag{14}$$

$\rho_{YY',i}$ being the reliability coefficient of Y_i. From Equation (14), it follows that

$$\beta_i = (\sigma_{Y,i}/\sigma_{X,i})\{(\rho_i - \rho_{YY',i}\rho_i)/(1 - \rho_i^2)\}$$

$$\tau_i = (\rho_{YY',i} - \rho_i^2)/(1 - \rho_i^2) \tag{15}$$

$$\alpha_i = -\mu_{X,i}\beta_i + \mu_{Y,i}(1 - \tau_i).$$

All quantities appearing in Equations (13) and (15) can be estimated straightforwardly; thus, estimates of the linear regression functions can be calculated.

Because the system of Equations (10) and (11) cannot be analytically solved for x_{ci} and y_{ci}, the determination of the optimal cutting scores may be carried out via numerical approximation methods such as the Newton iterative algorithm for solving nonlinear equations. However, before continuing with this procedure, it is necessary to specify, in addition to the regression and utility functions, the probability functions appearing in Equations (10) and (11). In Appendix A, the iterative solution of the system of Equations (10) and (11) is presented if it is assumed that the variables X and Y have possibly different bivariate normal distributions in each subpopulation.

Special Solutions

The optimal solution for the separate mastery and selection decision both can easily be derived from Equations (10) and (11) by imposing certain restrictions on x_{ci} and y_{ci}, respectively.

First, putting $x_{ci} = -\infty$ in Equation (10), that is, accepting all students, and using $\int_{-\infty}^{\infty} z_i(x|y_{ci})dx = 1$, Equation (10) will take the form

$$(b_{2i} - b_{1i})[E_i(T|y_{ci}) - t_c] + d_{2i} - d_{1i} = 0. \tag{16}$$

Putting both utility lines $u_{1i}(t)$ and $u_{2i}(t)$ in Formula 2 equal to each other, it appears that the t coordinate of the intersection, $t_{12,i}$, is equal to $t_c + (d_{1i} - d_{2i})/(b_{2i} - b_{1i})$, which implies that Equation (16) can be replaced by $E_i(T|y_{ci}) = t_{12,i}$. This solution yields the same optimal cutting score y_{ci}' as the one given by van der Linden and Mellenbergh (1977) for the separate mastery decision. Analogous to the combined decision problem, a psychometric model is needed to specify the regression function $E_i(T|y_{ci})$. For this purpose, the classical test model with linear regression (Lord & Novick, 1968, p. 65) will be assumed, which is known as Kelley's regression line:

$$E_i(T|y_{ci}) = \rho_{YY',i} y_{ci} + (1 - \rho_{YY',i})\mu_{Y,i}. \tag{17}$$

Substituting (17) into (16) gives

$$y_{ci}' = \mu_{Y,i} + [t_c - \mu_{Y,i} + (d_{1i} - d_{2i})/(b_{2i} - b_{1i})]/\rho_{YY',i}. \tag{18}$$

According to Lord and Novick (1968), Equation (17) is

> an interesting equation in that it expresses the estimate of the true score as a weighted sum of two separate estimates—one based upon the student's observed score on the mastery test, y_{ci}, and the other based upon the mean, $\mu_{y,i}$, of the group to which s(he) belongs. If the mastery test is highly reliable, much weight is given to the test score and little to the group mean, and vice versa. (p. 65)

321

Analogous to the derivation of the optimal separate mastery decision from Equation (10), the optimal separate selection decision can be derived from Equation (11) by putting $y_{ci} = -\infty$—that is, advancing all accepted students. Doing so, and using Formula 2, it follows that

$$E_i(T|x_{ci}) = t_{02,i} = t_c + (d_{0i} - d_{2i})/(b_{0i} + b_{2i}), \qquad (19)$$

where $t_{02,i}$ denotes the t coordinate of the intersection of utility line $u_{0i}(t)$ with $u_{2i}(t)$. Also, this optimal solution is the same as the one reached by Mellenbergh and van der Linden (1981) for the separate selection decision. Adopting Kelley's regression line from classical test theory again, it follows from Equation (19) that the optimal cutting score x'_{ci} can be expressed in closed form as

$$x'_{ci} = \mu_{X,i} + [t_c - \mu_{X,i} + (d_{0i} - d_{2i})/(b_{0i} + b_{2i})]/\rho_{XX',i}, \qquad (20)$$

where $\rho_{XX',i}$ denotes the reliability coefficient of X_i.

An interesting case arises when $d_{1i} = d_{2i}$ in Equation (18). Whenever this occurs, both utility lines $u_{1i}(t)$ and $u_{2i}(t)$ intersect at t_c, and thus, Equation (18) takes the form

$$y'_{ci} = \mu_{Y,i} + (t_c - \mu_{Y,i})/\rho_{YY',i}. \qquad (21)$$

In other words, if the amounts of constant utility associated with the actions retaining and advancing a student in the separate mastery decision are equal or there are no constant utilities at all, then there is no need to assess the parameters b_{1i} and b_{2i}. In the numerical example, this situation will be further elaborated.

Similarly, all utility function parameters vanish from Equation (20) whenever $d_{0i} = d_{2i}$; thus, Equation (20) can be further simplified to

$$x'_{ci} = \mu_{X,i} + (t_c - \mu_{X,i})/\rho_{XX',i}. \qquad (22)$$

Optimal Cutting Scores for Quota-Restricted Selection

In quota-restricted selection only a fixed number of students can be accepted for the instructional program. The selection constraint can be expressed as

$$p_0 = \sum_{i=1}^{g} p_i [\text{Prob}(X \geq x_{ci})] = \sum_{i=1}^{g} p_i \left[\int_{x_{ci}}^{\infty} q_i(x)dx \right], \qquad (23)$$

where $0 < p_0 < \sum_{i=1}^{g} p_i = 1$ represents the fixed proportion of all students that can be accepted.

Analogous to the quota-free model, optimal cutting scores x'_{ci} and y'_{ci} can be derived by maximization of the overall expected utility from Formula 8 subject to Equation (23). To solve this constrained optimization problem,

322

first introduce the selection constraint into the function to be optimized—Equation (8)—through a Lagrange multiplier λ:

$$L(x_{ci}, y_{ci}, \lambda) = \sum_{i=1}^{g} p_i E[u_i(T|x_{ci}, y_{ci})] + \lambda \left\{ p_0 - \sum_{i=1}^{g} p_i \left[\int_{x_{ci}}^{\infty} q_i(x) dx \right] \right\}, \quad (24)$$

where λ is a constant.

Now optimizing cutting scores for quota-restricted selection proceeds by differentiating $L(x_{ci}, y_{ci}, \lambda)$ with respect to x_{ci} and y_{ci}, setting the resulting expressions equal to zero, and solving for x_{ci}, y_{ci}, and λ using Newton's iterative method again. The system of Equations to be solved numerically is attached as Appendix B with a brief explanation. A numerical example illustrating the procedure is given in the next section.

A Numerical Example

The linear utility model for optimal selection-mastery decisions was applied to a sample of 43 freshmen in medicine. Both the selection and mastery tests consisted of 17 free-response items on elementary medical knowledge with test scores ranging from 0–100. The treatment consisted of a computer-aided instructional (CAI) program.

The students differed with respect to the level at which the subject biology was treated at the secondary education. Some students visited types of secondary education in which some minor attention was given to elementary medical knowledge during the biology lessons. Also, they finished the subject biology with a certifying exam at the end of the secondary education in which some items applied to elementary medical knowledge. Due to this difference in previous education, the total population of 43 students could be distinguished with respect to elementary medical knowledge into a disadvantaged and an advantaged subpopulation of 27 and 16 students, respectively.

Let the disadvantaged and advantaged population be referred to as subpopulation 1 and 2, respectively. The normal models assumed for the distributions X_i and Y_i showed a satisfactory fit to the test data for a Kolmogorov-Smirnov goodness of fit test with p-values of 0.869, 0.934, 0.867, and 0.993 for X_1, Y_1, X_2, and Y_2, respectively. The absolute differences between the theoretical and observed cumulative distribution functions were 0.0686, 0.1035, 0.1495, and 0.1067, respectively. The parameters for the theoretical cumulative distribution functions were estimated from the samples.

The teachers of the course considered students as having mastered the subject matter if their test scores were at least 55. Therefore, t_c was fixed at 55.

The means, standard deviations, and correlation between X and Y were computed for each subpopulation using the maximum likelihood estimates of the unbiased sample means, sample standard deviations, and sample cor-

323

relations, respectively. Furthermore, because the items were not scored as right or wrong, the reliabilities of the test scores were estimated as coefficient α (Cronbach, 1951) for each subpopulation. The results of these computations are shown in Table 1.

It should be emphasized that the necessary data for computing these statistics come from the experiment described earlier in which students exposed to the same selection test were randomly drawn and accepted for the treatment, after which their criterion performances were measured.

First, the quota-free situation is considered. Because population 2 was considered more advantaged than 1, it should hold that $b_{01} > b_{02}$ and $b_{21} > b_{22}$. Besides these conditions for the utility parameters in Formula 2, condition (5) should hold for the utility parameters b_{ji} ($j = 0, 1, 2; i = 1, 2$). Substituting the values of the statistics of Table 1 into Equations (13) and (15), the optimal cutting scores x'_{ci} and y'_{ci} were then obtained by numerically solving the system of Equations (26) and (27) of Appendix A for x_{ci} and $y_{ci}(i = 1, 2)$, and using the computer program NEWTON, available on request from the author, with t_c as starting values and as a criterion for convergence that the absolute differences between $x'_{ci, n+1}$ and $x'_{ci, n}$ as well as between $y'_{ci, n+1}$ and $y'_{ci, n}$ were smaller than 10^{-7}. To illustrate the dependence of the results on the utility structure, optimal cutting scores were computed for 10 different values of the utility parameters b_{ji} and d_{ji} ($j = 0, 1, 2; i = 1, 2$). The absolute values of b_{ji} and d_{ji} for utility function 1 until 5 were the same as the absolute values for utility function 6 until 10. However, the sign of b_{1i} was taken negative in the last five runs, taking into account the fact that the sign of b_{1i} could not be specified beforehand. The results are reported in Table 2 together with the number of iterations required to converge. As can be seen from Table 2, the consequence of increasing the parameters b_{0i} and b_{2i} ($i = 1, 2$) is a decrease of the cutting scores. Furthermore, inspection of Table 2 shows that a decrease of the amount of constant utility, $d_{ji}(j = 0, 1, 2; i = 1, 2)$, implies that the cutting scores have to be raised.

Using Equations (18) and (20), the optimal cutting scores were also computed for the separate mastery and selection decisions. Because no constant

TABLE 1
Statistics selection and mastery tests (X and Y)

	Disadvantaged		Advantaged	
Statistic	X	Y	X	Y
Mean	50.875	62.626	56.453	67.148
Standard deviation	10.981	11.645	11.674	13.344
Reliability	0.762	0.775	0.783	0.813
Correlation	0.8564		0.8685	

324

amounts of utility were assumed for utility functions 2 and 7, Equations (21) and (22) were used to compute the optimal cutting scores for these two utility specifications. The results are reported in Table 2. Table 2 shows that the optimal selection as well as the optimal mastery scores in the separate model have been raised compared to those in the combined model.

This implies that when the combined model is used for decision-making, more individuals will be selected for the treatment and more will be advanced because the cutting points from the simultaneous approach are lower than those obtained with the separate models. Moreover, for all utility structures the cutting points for the selection decision of the combined model are even substantially lower for the following reason. The psychometric portion of this model includes the information provided by the average performance of the group on the mastery test in the estimation of the cutting points for the selection decision—Equation (12)—whereas the psychometric portion of the separate selection model is entirely based on information provided by the selection test, Equation (20), as a weighted sum of individual and group score. Because the mean performance of the groups on the mastery test surpasses their performance on the selection test and the criterion for mastery, the combined model should yield lower cutting points for the selection decision.

In contrast, the cutting points for the mastery decision from the models generally do not have such large differences. These differences heavily depend upon the values of the utility parameters. For instance, for the values of utility structure 3, the difference for the advantaged group is only 2.30.

Inspection of Equation (7) and Table 2 shows that for fixed b_{ji} the optimal cutting scores are dependent upon $d_{2i} - d_{1i}$. The smaller the difference, the higher the optimal cutting scores. This makes sense: A smaller value of d_{2i} relative to d_{1i} means a smaller amount of constant utility for the decision "advance accepted students" relative to "retain accepted students" and, in that case, the optimal cutting score for the mastery decision is higher; a smaller value of d_{1i} relative to d_{2i} would produce the opposite effect.

To give an impression of the gain in reliability of decisions, the reliability both for the simultaneous and separate solution has been calculated. In the psychometric literature, several methods for computing and interpreting reliability indexes for decisions have been proposed (Hambleton & Novick, 1973; Huynh, 1976b; Subkoviak, 1976; Swaminathan, Hambleton, & Algina, 1974; van der Linden & Mellenbergh, 1978). Two reliability indexes that have been recommended are (a) the agreement coefficient and (b) the kappa coefficient. These two coefficients measure the consistency of decisions across two test administrations, and both require somewhat different interpretations than traditional reliability coefficients, which are essentially correlations of test scores on two administrations.

The agreement coefficient is simply the proportion of examinees consis-

325

TABLE 2
Optimal cutting scores quota-free selection

Number	Utility specifications		Cutting scores (x'_{ci}, y'_{ci})				Number of iterations	
	Disadvantaged	Advantaged	Simultaneous		Separate		Disadvantaged	Advantaged
			Disadvantaged	Advantaged	Disadvantaged	Advantaged		
1	$b_{01} = 3.5$ $\quad d_{01} = -2$ $b_{11} = 2$ $\quad d_{11} = -3$ $b_{21} = 3$ $\quad d_{21} = -4$	$b_{02} = 3$ $\quad d_{02} = -2$ $b_{12} = 2$ $\quad d_{12} = -3$ $b_{22} = 2.5$ $\quad d_{22} = -4$	$x_{c1} = 42.80$ $y_{c1} = 44.89$	$x_{c2} = 44.52$ $y_{c2} = 49.25$	$x_{c1} = 56.69$ $y_{c1} = 54.07$	$x_{c2} = 55.06$ $y_{c2} = 54.67$	50	13
2	$b_{01} = 3.5$ $\quad d_{01} = 0$ $b_{11} = 2$ $\quad d_{11} = 0$ $b_{21} = 3$ $\quad d_{21} = 0$	$b_{02} = 3$ $\quad d_{02} = 0$ $b_{12} = 2$ $\quad d_{12} = 0$ $b_{22} = 2.5$ $\quad d_{22} = 0$	$x_{c1} = 42.47$ $y_{c1} = 43.01$	$x_{c2} = 44.20$ $y_{c2} = 44.16$	$x_{c1} = 56.29$ $y_{c1} = 52.78$	$x_{c2} = 54.60$ $y_{c2} = 52.21$	51	17
3	$b_{01} = 3.5$ $\quad d_{01} = -3$ $b_{11} = 2$ $\quad d_{11} = -7$ $b_{21} = 3$ $\quad d_{21} = -10$	$b_{02} = 3$ $\quad d_{02} = -3$ $b_{12} = 2$ $\quad d_{12} = -7$ $b_{22} = 2.5$ $\quad d_{22} = -10$	$x_{c1} = 43.57$ $y_{c1} = 49.98$	$x_{c2} = 45.17$ $y_{c2} = 57.29$	$x_{c1} = 57.70$ $y_{c1} = 56.65$	$x_{c2} = 56.22$ $y_{c2} = 59.59$	12	6
4	$b_{01} = 5$ $\quad d_{01} = -2$ $b_{11} = 2$ $\quad d_{11} = -3$ $b_{21} = 6$ $\quad d_{21} = -4$	$b_{02} = 3.5$ $\quad d_{02} = -2$ $b_{12} = 2$ $\quad d_{12} = -3$ $b_{22} = 5$ $\quad d_{22} = -4$	$x_{c1} = 42.66$ $y_{c1} = 43.37$	$x_{c2} = 44.38$ $y_{c2} = 44.77$	$x_{c1} = 56.53$ $y_{c1} = 53.10$	$x_{c2} = 54.90$ $y_{c2} = 52.62$	53	16

326

5	$b_{01} = 15$	$d_{01} = -2$	$b_{02} = 14$	$d_{02} = -2$	$x_{c1} = 42.55$	$x_{c2} = 44.25$	$x_{c1} = 56.39$	$x_{c2} = 54.71$	47	16
	$b_{11} = 2$	$d_{11} = -3$	$b_{12} = 2$	$d_{12} = -3$	$y_{c1} = 43.20$	$y_{c2} = 44.47$	$y_{c1} = 52.94$	$y_{c2} = 52.38$		
	$b_{21} = 10$	$d_{21} = -4$	$b_{22} = 9$	$d_{22} = -4$						
6	$b_{01} = 3.5$	$d_{01} = -2$	$b_{02} = 3$	$d_{02} = -2$	$x_{c1} = 42.78$	$x_{c2} = 44.45$	$x_{c1} = 56.69$	$x_{c2} = 55.06$	49	18
	$b_{11} = -2$	$d_{11} = -3$	$b_{12} = -2$	$d_{12} = -3$	$y_{c1} = 43.21$	$y_{c2} = 43.38$	$y_{c1} = 53.04$	$y_{c2} = 52.48$		
	$b_{21} = 3$	$d_{21} = -4$	$b_{22} = 2.5$	$d_{22} = -4$						
7	$b_{01} = 3.5$	$d_{01} = 0$	$b_{02} = 3$	$d_{02} = 0$	$x_{c1} = 42.43$	$x_{c2} = 44.06$	$x_{c1} = 56.29$	$x_{c2} = 54.60$	59	16
	$b_{11} = -2$	$d_{11} = 0$	$b_{12} = -2$	$d_{12} = 0$	$y_{c1} = 43.02$	$y_{c2} = 44.38$	$y_{c1} = 52.78$	$y_{c2} = 52.21$		
	$b_{21} = 3$	$d_{21} = 0$	$b_{22} = 2.5$	$d_{22} = 0$						
8	$b_{01} = 3.5$	$d_{01} = -3$	$b_{02} = 3$	$d_{02} = -3$	$x_{c1} = 43.64$	$x_{c2} = 45.42$	$x_{c1} = 57.70$	$x_{c2} = 56.22$	36	21
	$b_{11} = -2$	$d_{11} = -7$	$b_{12} = -2$	$d_{12} = -7$	$y_{c1} = 43.56$	$y_{c2} = 44.26$	$y_{c1} = 53.55$	$y_{c2} = 53.03$		
	$b_{21} = 3$	$d_{21} = -10$	$b_{22} = 2.5$	$d_{22} = -10$						
9	$b_{01} = 5$	$d_{01} = -2$	$b_{02} = 3.5$	$d_{02} = -2$	$x_{c1} = 42.64$	$x_{c2} = 44.31$	$x_{c1} = 56.53$	$x_{c2} = 54.90$	53	18
	$b_{11} = -2$	$d_{11} = -3$	$b_{12} = -2$	$d_{12} = -3$	$y_{c1} = 43.14$	$y_{c2} = 44.38$	$y_{c1} = 52.94$	$y_{c2} = 52.38$		
	$b_{21} = 6$	$d_{21} = -4$	$b_{22} = 5$	$d_{22} = -4$						
10	$b_{01} = 15$	$d_{01} = -2$	$b_{02} = 14$	$d_{02} = -2$	$x_{c1} = 42.54$	$x_{c2} = 44.22$	$x_{c1} = 56.39$	$x_{c2} = 54.71$	50	17
	$b_{11} = -2$	$d_{11} = -3$	$b_{12} = -2$	$d_{12} = -3$	$y_{c1} = 43.12$	$y_{c2} = 44.38$	$y_{c1} = 52.88$	$y_{c2} = 52.32$		
	$b_{21} = 10$	$d_{21} = -4$	$b_{22} = 9$	$d_{22} = -4$						

tently classified on both administrations, and kappa reflects the proportions of consistent decisions observed beyond that expected by chance. As Berk (1980) has noted in reviewing the literature on criterion-referenced test reliability, the agreement coefficient might be the preferred index of agreement for tests intended for individualized instructional decisions at the classroom level.

Methods for estimating the agreement and kappa coefficient from a single test administration have been proposed, thereby eliminating the need for a second test administration (Huynh, 1976b, Marshall & Haertel, 1976; Subkoviak, 1976, 1988). Because the Kolmogorov-Smirnov goodness of fit tests indicate that the observed distributions of X_i and Y_i ($i = 1, 2$) do not significantly deviate from the norm, it can be shown (Peng & Subkoviak, 1980; Subkoviak, 1976) that the agreement coefficient for the X_i scores, $p_{xi}(i = 1, 2)$, can be calculated as follows: $p_{xi} = 1 - 2[P(z < c) - P(z < c, z' < c)]$. In this equation, $c = (x_{ci} - 0.5 - \mu_{X,i})/\sigma_{X,i}$ where x_{ci}, $\mu_{X,i}$, and $\sigma_{X,i}$ are the optimal cutting score, the mean, and the standard deviation of the X_i scores, respectively (see Huynh, 1976b). The probability that a standardized normal variable is less than c is $P(z < c)$. Finally, $P(z < c, z' < c)$ is the probability that two standardized normal variables, with correlation α, are both less than c, where α is the Cronbach's alpha reliability coefficient of the X_i scores; this quantity is obtained from tables of the bivariate normal distribution provided by Gupta (1963) and others. Similarly, $p_{yi}(i = 1, 2)$ can be obtained for the Y_i scores. The results of these computations for both X_i and Y_i scores are shown in Table 3 in which the ratios of odds of consistent classification, r_{xi} and $r_{yi}(i = 1, 2)$, for the simultaneous and separate solution have also been calculated. As can be seen from these ratios, the simultaneous approach is associated with substantial gains in consistency of classification.

Finally, the quota-restricted situation is considered. The proportions p_i ($i = 1, 2$) of the student population belonging to each subpopulation were estimated as n_i/n, where n represents the total sample size and n_i represents the number of students in the sample in subpopulation i. The proportion p_0 of the total student population that could be accepted for the instructional treatment was arbitrarily set equal to 0.333. The optimal cutting scores x'_{ci} and y'_{ci} were obtained by numerically solving the system of Equations (26) of Appendix A and (36) and (37) of Appendix B for x_{ci} and y_{ci}, and using the computer program LAGRANGE, available on request from the author, with the optimal solution of the quota-free situation as starting values and as a criterion for convergence that the absolute differences between $x'_{c1,n+1}$ and $x'_{c1,n}$, $y'_{c1,n+1}$ and $y'_{c1,n}$, $x'_{c2,n+1}$ and $x'_{c2,n}$, and $y'_{c2,n+1}$ and $y'_{c2,n}$ were smaller than 10^{-7}. The optimal cutting scores were computed again for 10 different values of b_{ji} and d_{ji}; the results are shown in Table 4 together with the number of iterations to convergence.

328

From Table 4 it can be seen that the optimal selection scores x'_{c1} and x'_{c2} in the quota-restricted model have to be raised compared to those in the quota-free model. This result is in accordance with our expectations, because fewer students can be accepted in the restricted situation.

Discussion

In this article an approach to instructional decision making for combinations of elementary decisions has been presented. A useful application of simultaneous decision making can be found in the area of instructional decision making in ISSs. As an example, two elementary decisions (a selection and a mastery decision) were combined into a simple ISS to indicate how, by simultaneous optimization of such networks, optimal rules for proceeding students through ISSs can be designed within a Bayesian decision–theoretic framework. The utility structure adopted in this combined decision problem was a linear utility function.

Further examination of the "best" way to represent more complicated instructional networks of combinations of elementary decisions seems to be a valuable line of research. Such instructional networks can also be formalized with the aid of Bayesian statistics and optimal rules for these simultaneous optimization problems can be found.

Also, more efforts are needed to examine other more realistic forms utility functions might take in certain educational applications. For example, the normal ogive utility function (Novick & Lindley, 1978) takes utility to be a nonlinear function of the true score. Such a utility structure might be adopted, for instance, when it is reasonable to assume a levelling-off effect.

Another interesting line of research seems to be to design "optimal CAI-networks" using the method of simulation. This can be done by first deriving theoretical optimal decision rules using the models advocated in this article. Then, for a simulated distribution of students, it can be determined in which instructional network the shortest time is spent to reach a certain final mastery level.

A final remark is appropriate. The parameters to be estimated are all familiar parameters, such as correlations, covariances, linear regression parameters, and multiple regression parameters. Estimation of these, for which standard statistical theory provides estimators with favorable properties, can yield errors of estimation that propagate in computing the optimal rules of this article. I therefore recommend not using samples as small as those used in the foregoing experiment. The example, however, was used only for illustration. This does not mean, however, that small samples necessarily yield inaccurate results. The way errors of estimation propagate depends on the whole structure of the optimal rules, involving not only distributional parameters but utility parameters as well.

In reference to the last remark, one of the referees of an earlier draft of

329

TABLE 3
Agreement coefficients quota-free selection

Number	Utility specifications — Disadvantaged	Utility specifications — Advantaged	Agreement coefficients Simultaneous — Disadvantaged	Simultaneous — Advantaged	Separate — Disadvantaged	Separate — Advantaged	Odds simultaneous / Odds separate — Disadvantaged	Advantaged
1	$b_{01}=3.5$, $d_{01}=-2$; $b_{11}=2$, $d_{11}=-3$; $b_{21}=3$, $d_{21}=-4$	$b_{02}=3$, $d_{02}=-2$; $b_{12}=2$, $d_{12}=-3$; $b_{22}=2.5$, $d_{22}=-4$	$p_{x1}=0.84$, $p_{y1}=0.94$	$p_{x2}=0.88$, $p_{y2}=0.93$	$p_{x1}=0.80$, $p_{y1}=0.84$	$p_{x2}=0.79$, $p_{y2}=0.88$	$r_{x1}=1.31$, $r_{y1}=2.98$	$r_{x2}=1.95$, $r_{y2}=1.81$
2	$b_{01}=3.5$, $d_{01}=0$; $b_{11}=2$, $d_{11}=0$; $b_{21}=3$, $d_{21}=0$	$b_{02}=3$, $d_{02}=0$; $b_{12}=2$, $d_{12}=0$; $b_{22}=2.5$, $d_{22}=0$	$p_{x1}=0.84$, $p_{y1}=0.95$	$p_{x2}=0.89$, $p_{y2}=0.96$	$p_{x1}=0.80$, $p_{y1}=0.86$	$p_{x2}=0.79$, $p_{y2}=0.90$	$r_{x1}=1.31$, $r_{y1}=3.09$	$r_{x2}=2.15$, $r_{y2}=2.67$
3	$b_{01}=3.5$, $d_{01}=-3$; $b_{11}=2$, $d_{11}=-7$; $b_{21}=3$, $d_{21}=-10$	$b_{02}=3$, $d_{02}=-3$; $b_{12}=2$, $d_{12}=-7$; $b_{22}=2.5$, $d_{22}=-10$	$p_{x1}=0.83$, $p_{y1}=0.89$	$p_{x2}=0.88$, $p_{y2}=0.85$	$p_{x1}=0.81$, $p_{y1}=0.82$	$p_{x2}=0.79$, $p_{y2}=0.84$	$r_{x1}=1.15$, $r_{y1}=1.78$	$r_{x2}=1.95$, $r_{y2}=1.08$
4	$b_{01}=5$, $d_{01}=-2$; $b_{11}=2$, $d_{11}=-3$; $b_{21}=6$, $d_{21}=-4$	$b_{02}=3.5$, $d_{02}=-2$; $b_{12}=2$, $d_{12}=-3$; $b_{22}=5$, $d_{22}=-4$	$p_{x1}=0.84$, $p_{y1}=0.95$	$p_{x2}=0.88$, $p_{y2}=0.95$	$p_{x1}=0.80$, $p_{y1}=0.85$	$p_{x2}=0.79$, $p_{y2}=0.90$	$r_{x1}=1.31$, $r_{y1}=3.35$	$r_{x2}=1.95$, $r_{y2}=2.11$

330

5	$b_{01} = 15$ $b_{11} = 2$ $b_{21} = 10$	$d_{01} = -2$ $d_{11} = -3$ $d_{21} = -4$	$b_{02} = 14$ $b_{12} = 2$ $b_{22} = 9$	$d_{02} = -2$ $d_{12} = -3$ $d_{22} = -4$	$p_{x1} = 0.84$ $p_{y1} = 0.95$	$p_{x2} = 0.89$ $p_{y2} = 0.96$	$p_{x1} = 0.80$ $p_{y1} = 0.86$	$p_{x2} = 0.79$ $p_{y2} = 0.90$
6	$b_{01} = 3.5$ $b_{11} = -2$ $b_{21} = 3$	$d_{01} = -2$ $d_{11} = -3$ $d_{21} = -4$	$b_{02} = 3$ $b_{12} = -2$ $b_{22} = 2.5$	$d_{02} = -2$ $d_{12} = -3$ $d_{22} = -4$	$p_{x1} = 0.84$ $p_{y1} = 0.95$	$p_{x2} = 0.88$ $p_{y2} = 0.96$	$p_{x1} = 0.80$ $p_{y1} = 0.86$	$p_{x2} = 0.79$ $p_{y2} = 0.90$
7	$b_{01} = 3.5$ $b_{11} = -2$ $b_{21} = 3$	$d_{01} = 0$ $d_{11} = 0$ $d_{21} = 0$	$b_{02} = 3$ $b_{12} = -2$ $b_{22} = 2.5$	$d_{02} = 0$ $d_{12} = 0$ $d_{22} = 0$	$p_{x1} = 0.84$ $p_{y1} = 0.95$	$p_{x2} = 0.89$ $p_{y2} = 0.96$	$p_{x1} = 0.80$ $p_{y1} = 0.86$	$p_{x2} = 0.79$ $p_{y2} = 0.90$
8	$b_{01} = 3.5$ $b_{11} = -2$ $b_{21} = 3$	$d_{01} = -3$ $d_{11} = -7$ $d_{21} = -10$	$b_{02} = 3$ $b_{12} = -2$ $b_{22} = 2.5$	$d_{02} = -3$ $d_{12} = -7$ $d_{22} = -10$	$p_{x1} = 0.83$ $p_{y1} = 0.95$	$p_{x2} = 0.87$ $p_{y2} = 0.96$	$p_{x1} = 0.81$ $p_{y1} = 0.85$	$p_{x2} = 0.79$ $p_{y2} = 0.89$
9	$b_{01} = 5$ $b_{11} = -2$ $b_{21} = 6$	$d_{01} = -2$ $d_{11} = -3$ $d_{21} = -4$	$b_{02} = 3.5$ $b_{12} = -2$ $b_{22} = 5$	$d_{02} = -2$ $d_{12} = -3$ $d_{22} = -4$	$p_{x1} = 0.84$ $p_{y1} = 0.95$	$p_{x2} = 0.88$ $p_{y2} = 0.96$	$p_{x1} = 0.80$ $p_{y1} = 0.86$	$p_{x2} = 0.79$ $p_{y2} = 0.90$
10	$b_{01} = 15$ $b_{11} = -2$ $b_{21} = 10$	$d_{01} = -2$ $d_{11} = -3$ $d_{21} = -4$	$b_{02} = 14$ $b_{12} = -2$ $b_{22} = 9$	$d_{02} = -2$ $d_{12} = -3$ $d_{22} = -4$	$p_{x1} = 0.84$ $p_{y1} = 0.95$	$p_{x2} = 0.89$ $p_{y2} = 0.96$	$p_{x1} = 0.80$ $p_{y1} = 0.86$	$p_{x2} = 0.79$ $p_{y2} = 0.90$

$r_{x1} = 1.31$ $r_{y1} = 3.09$	$r_{x2} = 2.15$ $r_{y2} = 2.67$
$r_{x1} = 1.31$ $r_{y1} = 3.09$	$r_{x2} = 1.95$ $r_{y2} = 2.67$
$r_{x1} = 1.31$ $r_{y1} = 3.09$	$r_{x2} = 2.15$ $r_{y2} = 2.67$
$r_{x1} = 1.15$ $r_{y1} = 3.35$	$r_{x2} = 1.78$ $r_{y2} = 2.97$
$r_{x1} = 1.31$ $r_{y1} = 3.09$	$r_{x2} = 1.95$ $r_{y2} = 2.67$
$r_{x1} = 1.31$ $r_{y1} = 3.09$	$r_{x2} = 2.15$ $r_{y2} = 2.67$

331

TABLE 4
Optimal cutting scores quota-restricted selection

Number	Utility specifications				Cutting scores (x'_{ci}, y'_{ci})		No. of iterations
	Disadvantaged		Advantaged		Simultaneous		
					Disadvantaged	Advantaged	
1	$b_{01} = 3.5$	$d_{01} = -2$	$b_{02} = 3$	$d_{02} = -2$	$x_{c1} = 56.47$	$x_{c2} = 59.34$	29
	$b_{11} = 2$	$d_{11} = -3$	$b_{12} = 2$	$d_{12} = -3$	$y_{c1} = 49.31$	$y_{c2} = 49.65$	
	$b_{21} = 3$	$d_{21} = -4$	$b_{22} = 2.5$	$d_{22} = -4$			
2	$b_{01} = 3.5$	$d_{01} = 0$	$b_{02} = 3$	$d_{02} = 0$	$x_{c1} = 56.47$	$x_{c2} = 59.34$	30
	$b_{11} = 2$	$d_{11} = 0$	$b_{12} = 2$	$d_{12} = 0$	$y_{c1} = 48.59$	$y_{c2} = 48.21$	
	$b_{21} = 3$	$d_{21} = 0$	$b_{22} = 2.5$	$d_{22} = 0$			
3	$b_{01} = 3.5$	$d_{01} = -3$	$b_{02} = 3$	$d_{02} = -3$	$x_{c1} = 56.47$	$x_{c2} = 59.34$	23
	$b_{11} = 2$	$d_{11} = -7$	$b_{12} = 2$	$d_{12} = -7$	$y_{c1} = 50.72$	$y_{c2} = 52.49$	
	$b_{21} = 3$	$d_{21} = -10$	$b_{22} = 2.5$	$d_{22} = -10$			
4	$b_{01} = 5$	$d_{01} = -2$	$b_{02} = 3.5$	$d_{02} = -2$	$x_{c1} = 55.96$	$x_{c2} = 60.17$	31
	$b_{11} = 2$	$d_{11} = -3$	$b_{12} = 2$	$d_{12} = -3$	$y_{c1} = 48.49$	$y_{c2} = 48.90$	
	$b_{21} = 6$	$d_{21} = -4$	$b_{22} = 5$	$d_{22} = -4$			

332

5	$b_{01} = 15$ $b_{11} = 2$ $b_{21} = 10$	$d_{01} = -2$ $d_{11} = -3$ $d_{21} = -4$	$b_{02} = 14$ $b_{12} = 2$ $b_{22} = 9$	$d_{02} = -2$ $d_{12} = -3$ $d_{22} = -4$	$x_{c1} = 56.96$ $y_{c1} = 48.94$	$x_{c2} = 58.61$ $y_{c2} = 47.92$	29
6	$b_{01} = 3.5$ $b_{11} = -2$ $b_{21} = 3$	$d_{01} = -2$ $d_{11} = -3$ $d_{21} = -4$	$b_{02} = 3$ $b_{12} = -2$ $b_{22} = 2.5$	$d_{02} = -2$ $d_{12} = -3$ $d_{22} = -4$	$x_{c1} = 56.47$ $y_{c1} = 48.74$	$x_{c2} = 59.34$ $y_{c2} = 48.38$	30
7	$b_{01} = 3.5$ $b_{11} = -2$ $b_{21} = 3$	$d_{01} = 0$ $d_{11} = 0$ $d_{21} = 0$	$b_{02} = 3$ $b_{12} = -2$ $b_{22} = 2.5$	$d_{02} = 0$ $d_{12} = 0$ $d_{22} = 0$	$x_{c1} = 56.47$ $y_{c1} = 48.59$	$x_{c2} = 59.34$ $y_{c2} = 48.21$	30
8	$b_{01} = 3.5$ $b_{11} = -2$ $b_{21} = 3$	$d_{01} = -3$ $d_{11} = -7$ $d_{21} = -10$	$b_{02} = 3$ $b_{12} = -2$ $b_{22} = 2.5$	$d_{02} = -3$ $d_{12} = -7$ $d_{22} = -10$	$x_{c1} = 56.47$ $y_{c1} = 49.03$	$x_{c2} = 59.34$ $y_{c2} = 48.70$	30
9	$b_{01} = 5$ $b_{11} = -2$ $b_{21} = 6$	$d_{01} = -2$ $d_{11} = -3$ $d_{21} = -4$	$b_{02} = 3.5$ $b_{12} = -2$ $b_{22} = 5$	$d_{02} = -2$ $d_{12} = -3$ $d_{22} = -4$	$x_{c1} = 55.96$ $y_{c1} = 48.40$	$x_{c2} = 60.17$ $y_{c2} = 48.75$	31
10	$b_{01} = 15$ $b_{11} = -2$ $b_{21} = 10$	$d_{01} = -2$ $d_{11} = -3$ $d_{21} = -4$	$b_{02} = 14$ $b_{12} = -2$ $b_{22} = 9$	$d_{02} = -2$ $d_{12} = -3$ $d_{22} = -4$	$x_{c1} = 56.96$ $y_{c1} = 48.91$	$x_{c2} = 58.61$ $y_{c2} = 47.89$	29

this article raised the question of robustness of results—that is, whether parameters differing somewhat from the ones that were actually used lead to the same decisions. In order to examine this question, a few computer runs using the programs NEWTON and LAGRANGE have been made with small differences in the utility as well as the distributional parameters. These runs indicated that the optimal cutting scores for both the simultaneous and separate approaches were not very sensitive to small changes in the utility parameters. As far as the sensitivity of the decisions with respect to the distributional parameters was concerned, it appeared that the decisions were somewhat influenced. However, the optimal cutting scores for the simultaneous and the separate approaches changed in the same direction and with about the same magnitude. Moreover, the ratios of odds of consistent classification hardly changed, indicating that the simultaneous optimization technique is still superior in this case to the separate approaches.

To deal with errors of estimation in the distributional parameters, it should be noticed that still another approach is possible. If only samples of limited size can be used, one can think of establishing "regions of indifference" as in the Johnson and Neyman (1936) technique—that is, intervals on the selection as well as the mastery test score variable for which we are indifferent to the selection and mastery decision because of sampling error. These regions of indifference can be established by first constructing confidence intervals for the distributional parameters, and then simultaneously determining the range of optimal cutting scores associated with these intervals using the procedures presented here.

Appendix A

Iterative Solution in Case of the Bivariate Normal Model

In order to solve the system of Equations (10) and (11) for x_{ci} and y_{ci}, the decision maker must specify the joint probability function of X and Y. It is assumed that the variables X and Y have possibly different bivariate normal distributions in each subpopulation. Assuming that the X and Y scores are in their standardized form, this can be written as

$$k_i(x_N, y_N) = \{(2\pi)^{-1}(1 - \rho_i^2)^{-1/2} \exp[-(x_N^2 - 2\rho_i x_N y_N + y_N^2)/2(1 - \rho_i^2)]\}, \quad (25)$$

where x_N and y_N denote the standardized scores $(x - \mu_X)/\sigma_X$ and $(y - \mu_Y)/\sigma_Y$ of X and Y, respectively. For the standardized bivariate normal distribution in Equation (25), the conditional distribution of X_N given $Y_N = y_n$ is normal with expected value $\rho_i y_N$ and variance $(1 - \rho_i^2)$ (e.g., Johnson & Kotz, 1970). Likewise, the distribution of Y_N given $X_N = x_N$ is normal with expected value $\rho_i x_N$ and variance $(1 - \rho_i^2)$.

Substituting $\alpha_i + \beta_i x + \tau_i y_{ci}$ and $N(\rho_i y_{N, ci}, 1 - \rho_i^2)$ into Equation (10) for

334

$E_i(T|x, y_{ci})$ and $z_i(x_N|y_{N, ci})$, respectively, and using the property that the primitive function of $xe^{-0.5x^2}$ is equal to $-e^{-0.5x^2}$, it follows that Equation (10) will take the form

$$
\begin{aligned}
f(x_{ci}, y_{ci}) = &\,[(b_{2i} - b_{1i})(\alpha_i + \beta_i \mu_{X, i} + \tau_i y_{ci} + \beta_i \sigma_{X, i} \rho_i y_{N, ci} - t_c) \\
&+ d_{2i} - d_{1i}]\sigma_{X, i}\{1 - \Phi[(x_{N, ci} - \rho_i y_{N, ci})/(1 - \rho_i^2)^{1/2}]\} \\
&+ (b_{2i} - b_{1i})\beta_i \sigma_{X, i^2}(1 - \rho_i^2)^{1/2}\, \varphi[(x_{N, ci} - \rho_i y_{N, ci})/(1 - \rho_i^2)^{1/2}] \\
= &\,0,
\end{aligned} \tag{26}
$$

where $\Phi[.]$ and $\varphi[.]$ denote the standard normal distribution function and the standard normal density, respectively.

Similarly, substituting $\theta_i + \Gamma_i x_{ci}$, $\alpha_i + \beta_i x_{ci} + \tau_i y$, and $N(\rho_i x_{N, ci}, 1 - \rho_i^2)$ into Equation (11) for $E_i(T|x_{ci})$, $E_i(T|x_{ci}, y)$, and $m_i(y_N|x_{N, ci})$, respectively, results in

$$
\begin{aligned}
g(x_{ci}, y_{ci}) = &\,[(b_{0i} + b_{1i})(\theta_i + \Gamma_i x_{ci} - t_c) + d_{1i} - d_{0i}] \\
&+ [(b_{2i} - b_{1i})(\alpha_i + \beta_i x_{ci} + \tau_i \mu_{Y, i} + \tau_i \rho_i \sigma_{Y, i} x_{N, ci} - t_c) \\
&+ d_{2i} - d_{1i}]\{1 - \Phi[(y_{N, ci} - \rho_i x_{N, ci})/(1 - \rho_i^2)^{1/2}]\} \\
&+ (b_{2i} - b_{1i})\tau_i \sigma_{Y, i}(1 - \rho_i^2)^{1/2}\, \varphi[(y_{N, ci} - \rho_i x_{N, ci})/(1 - \rho_i^2)^{1/2}] \\
= &\,0.
\end{aligned} \tag{27}
$$

The system of Equations (26) and (27) cannot be analytically solved for x_{ci} and y_{ci}, but it can be iteratively solved using Newton's method for systems of nonlinear equations (e.g., Ortega & Rheinholdt, 1970). Updated estimates $x'_{ci, n+1}$ and $y'_{ci, n+1}$ after iteration $n + 1$ are obtained using the following formulas:

$$
\begin{aligned}
x'_{ci, n+1} &= x'_{ci, n} - \{[f(\partial/\partial y_{ci})g - g(\partial/\partial y_{ci})f]/J(f, g)\} \\
y'_{ci, n+1} &= y'_{ci, n} - \{[g(\partial/\partial x_{ci})f - f(\partial/\partial x_{ci})g]/J(f, g]\},
\end{aligned} \tag{28}
$$

where $J(f, g) = (\partial/\partial x_{ci})f(\partial/\partial y_{ci})g - (\partial/\partial x_{ci})g(\partial/\partial y_{ci})f$ represents the Jacobian of the functions $f(x_{ci}, y_{ci})$ and $g(x_{ci}, y_{ci})$. It is recommended that the cut-off score t_c on the true score scale T is used as a first approximation to x'_{ci} and y'_{ci}.

In order to solve the nonlinear system of Equations (26) and (27) via the iterative procedure given in (28), the partial derivatives of $f(x_{ci}, y_{ci})$ and $g(x_{ci}, y_{ci})$ are needed. They are given as

$$
\begin{aligned}
(\partial/\partial x_{ci})f(x_{ci}, y_{ci}) = &\,-[\sqrt{(1 - \rho_i^2)}]^{-1}\, \varphi[(x_{N, ci} - \rho_i y_{N, ci})/(1 - \rho_i^2)^{1/2}] \\
&[(b_{2i} - b_{1i})(\alpha_i + \beta_i x_{ci} + \tau_i y_{ci} - t_c) + d_{2i} - d_{1i}],
\end{aligned} \tag{29}
$$

335

$$(\partial/\partial y_{ci})f(x_{ci}, y_{ci})$$

$$= (b_{2i} - b_{1i})\sigma_{X,i}/\sigma_{Y,i}\{(\tau_i\sigma_{Y,i} + \beta_i\rho_i\sigma_{X,i})(1 - \Phi[(x_{N,ci} - \rho_i y_{N,ci})/(1 - \rho_i^2)^{1/2}])$$

$$+ \rho_i/[(1 - \rho_i^2)^{1/2}]\varphi[(x_{N,ci} - \rho_i y_{N,ci})/(1 - \rho_i^2)^{1/2}](\alpha_i + \beta_i x_{ci} + \tau_i y_{ci} - t_c)\}$$

$$+ (d_{2i} - d_{1i})\rho_i/[(1 - \rho_i^2)^{1/2}]\sigma_{X,i}\varphi[(x_{N,ci} - \rho_i y_{N,ci})/(1 - \rho_i^2)^{1/2}]/\sigma_{Y,i}, \qquad (30)$$

$$(\partial/\partial x_{ci})g(x_{ci}, y_{ci}) = (b_{0i} + b_{1i})\Gamma_i$$

$$+ (b_{2i} - b_{1i})/\sigma_{X,i}\{(\beta_i\sigma_{X,i} + \tau_i\sigma_{Y,i}\rho_i)(1 - \Phi[(y_{N,ci} - \rho_i x_{N,ci})/(1 - \rho_i^2)^{1/2}])$$

$$+ \rho_i/[(1 - \rho_i^2)^{1/2}]\varphi[(y_{N,ci} - \rho_i x_{N,ci})/(1 - \rho_i^2)^{1/2}](\alpha_i + \beta_i x_{ci} + \tau_i y_{ci} - t_c)\}$$

$$+ (d_{2i} - d_{1i})\rho_i/[(1 - \rho_i^2)^{1/2}]\varphi[(y_{N,ci} - \rho_i x_{N,ci})/(1 - \rho_i^2)^{1/2}]/\sigma_{X,i}, \qquad (31)$$

$$(\partial/\partial y_{ci})g(x_{ci}, y_{ci}) = -[(1 - \rho_i^2)^{1/2}]^{-1}\varphi[(y_{N,ci} - \rho_i x_{N,ci})/(1 - \rho_i^2)^{1/2}]$$

$$[(b_{2i} - b_{1i})(\alpha_i + \beta_i x_{ci} + \tau_i y_{ci} - t_c) + d_{2i} - d_{1i}]/\sigma_{Y,i}. \qquad (32)$$

An interesting special case of the combined linear utility function arises when $d_{0i} = d_{1i} = d_{2i}$. In that case, all utility parameters d_{ji} ($j = 0, 1, 2$) vanish from Equations (26) and (27). In other words, if the amount of constant utility, d_{ji}, for each action is equal, then there is no need to choose values for d_{ji} in determining the optimal cutting scores x'_{ci} and y'_{ci}.

It can be shown (Lord & Novick, 1968, sect. 17.2) that the standard normal distributions appearing in Equations (25)–(32) are almost interchangeable with logistic functions for a scale parameter equal to 1.7. The logistic model will be preferred in the iterative procedure because it is easier to work with mathematically than the standard normal model. Using this approximation, we may rewrite the standard normal distribution as follows:

$$\Phi[(x_{N,ci} - \rho_i y_{N,ci})/(1 - \rho_i^2)^{1/2}] = \{1 + \exp[-1.7(x_{N,ci} - \rho_i y_{N,ci})/(1 - \rho_i^2)^{1/2}]\}^{-1},$$

$$\Phi[(y_{N,ci} - \rho_i x_{N,ci})/(1 - \rho_i^2)^{1/2}] = \{1 + \exp[-1.7(y_{N,ci} - \rho_i x_{N,ci})/(1 - \rho_i^2)^{1/2}]\}^{-1}.$$

The iterative procedure is implemented in a computer program called NEWTON.

Appendix B

Numerical Solution for Quota-Restricted Selection

Differentiating $L(x_{ci}, y_{ci}, \lambda)$—Equation (23)—with respect to y_{ci} and x_{ci}, setting the resulting expressions equal to zero, and using $p_i > 0$, yields

$$(\partial/\partial y_{ci})E[u_i(T|x_{ci}, y_{ci})] = 0, \qquad (33)$$

$$(\partial/\partial x_{ci})E[u_i(T|x_{ci}, y_{ci})] + \lambda q_i(x_{ci}) = 0. \qquad (34)$$

As can be noticed from Equation (9), the solution to Equation (33) is the same as the solution for the case of quota-free selection given by Equation (26).

The first term in the left-hand side of Equation (34) represents the derivative of the expected utility of a random student with respect to x_{ci} in the unrestricted situation—Equation (11). Substituting this partial derivative into Equation (34), and using $q_i(x) \geq 0$, it follows that

$$
\begin{aligned}
(b_{0i} + b_{1i})[E_i(T|x_{ci}) - t_c] + d_{1i} - d_{0i} - \lambda \\
+ \int_{y_{ci}}^{\infty} \{(b_{2i} - b_{1i})[E_i(T|x_{ci}, y) - t_c] + d_{2i} - d_{1i}\} m_i(y|x_{ci})\, dy = 0.
\end{aligned}
\tag{35}
$$

Inserting the expressions for the linear regression functions and $N(\rho_i x_{N,ci}, 1 - \rho_i^2)$ for $m_i(y_N|x_{N,ci})$, and integrating Equation (35), results in

$$
\begin{aligned}
h(x_{ci}, y_{ci}) =\ & [(b_{0i} + b_{1i})(\theta_i + \Gamma_i x_{ci} - t_c) + d_{1i} - d_{0i} - \lambda] \\
& + [(b_{2i} - b_{1i})(\alpha_i + \beta_i x_{ci} + \tau_i \mu_{Y,i} + \tau_i \rho_i \sigma_{Y,i} x_{N,ci} - t_c) \\
& + d_{2i} - d_{1i}](1 - \Phi[(y_{N,ci} - \rho_i x_{N,ci})/(1 - \rho_i^2)^{1/2}]) \\
& + (b_{2i} - b_{1i})\tau_i \sigma_{Y,i}(1 - \rho_i^2)^{1/2}\, \varphi[(y_{N,ci} - \rho_i x_{N,ci})/(1 - \rho_i^2)^{1/2}] = 0.
\end{aligned}
\tag{36}
$$

Because it has been assumed that the joint distribution of X and Y is a possibly different bivariate normal distribution in each population, it follows that $q_i(x)$ is a normal distribution with mean $\mu_{X,i}$ and variance $\sigma_{X,i}^2$ (e.g., Johnson & Kotz, 1970). Hence, the restriction of Equation (23) can be written as

$$
v(x_{c1}, x_{c2}, \ldots, x_{cg}) = \sum_{i=1}^{g} p_i[1 - \Phi(x_{N,ci})] - p_0 = 0.
\tag{37}
$$

Now, the solution for the quota-restricted selection model is found by solving the system of Equations (26), (36), and (37) for the $(2g + 1)$ unknown parameters x_{ci}, y_{ci}, and λ. Note that with quota-restricted selection, unlike quota-free selection, the optimal cutting scores x_{ci}' and y_{ci}' $(i = 1, \ldots, g)$ are dependent upon each other.

In order to apply Newton's iterative method to solve the given system of nonlinear equations, the partial derivatives are required again. From Equations (36) and (27), it can easily be verified that $(\partial/\partial x_{ci})h(x_{ci}, y_{ci}) = (\partial/\partial x_{ci})g(x_{ci}, y_{ci})$, $(\partial/\partial y_{ci})h(x_{ci}, y_{ci}) = (\partial/\partial y_{ci})g(x_{ci}, y_{ci})$, and $(\partial/\partial\lambda)h(x_{ci}, y_{ci}) = -1$. The derivatives of $f(x_{ci}, y_{ci})$ and $g(x_{ci}, y_{ci})$ with respect to x_{ci} and y_{ci} were given in Equations (29) until (32), respectively. Furthermore, it follows from Equation (37) that

$$
(\partial/\partial x_{ci})v(x_{c1}, x_{c2}, \ldots, x_{cg}) = (-p_i/\sigma_{X,i})\varphi(x_{N,ci}).
\tag{38}
$$

337

Hans J. Vos

Analogous to the quota-free selection model, as can easily be seen from Equations (26), (36), and (37), no values for the utility function parameters $d_{ji}(j = 0, 1, 2)$ have to be specified when the amount of constant utility for each action is equal.

A computer program called LAGRANGE has been written to obtain an optimal decision rule. In the program, the optimal solution (x_{ci}', y_{ci}') of the quota-free selection model can be used as a first approximation in the iterative procedure.

References

Berk, R. A. (1980). A consumer's guide to criterion-referenced test reliability. *Journal of Educational Measurement, 17,* 323–349.

Chuang, D. T., Chen, J. J., & Novick, M. R. (1981). Theory and practice for the use of cut-scores for personnel decisions. *Journal of Educational Statistics, 6,* 129–152.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297–334.

DeGroot, M. H. (1970). *Optimal statistical decisions.* New York: McGraw-Hill.

De Gruijter, D. N. M., & Hambleton, R. K. (1984). On problems encountered using decision theory to set cutoff scores. *Applied Psychological Measurement, 8,* 1–8.

Ferguson, T. S. (1967). *Mathematical statistics: A decision theoretic approach.* New York: Academic Press.

Gross, A. L., & Su, W. H. (1975). Defining a "fair" or "unbiased" selection model: A question of utilities. *Journal of Applied Psychology, 60,* 345–351.

Gupta, S. S. (1963). Probability integrals of multivariate normal and multivariate t. *Annals of Mathematical Statistics, 34,* 792–828.

Hambleton, R. K., & Novick, M. R. (1973). Toward an integration of theory and method for criterion-referenced tests. *Journal of Educational Measurement, 10,* 159–170.

Huynh, H. (1976a). Statistical considerations of mastery scores. *Psychometrika, 41,* 65–79.

Huynh, H. (1976b). On the reliability of decisions in domain-referenced testing. *Journal of Educational Measurement, 13,* 253–264.

Huynh, H. (1977). Two simple cases of mastery scores based on the beta-binomial model. *Psychometrika, 41,* 65–78.

Johnson, N. L., & Kotz, S. (1970). *Distributions in statistics: Continuous univariate distributions.* Boston: Houghton Mifflin.

Johnson, P. D., & Neyman, J. (1936). Tests of linear hypotheses and their application to some educational problems. *Statistical Research Memoirs, 1,* 57–93.

Karlin, S., & Rubin, H. (1956). Distributions possessing a monotone likelihood ratio. *Journal of the American Statistical Association, 1,* 637–643.

Keeney, D., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value trade-offs.* New York: Wiley.

Lehmann, E. L. (1959). *Testing statistical hypothesis.* New York: Wiley.

Lindgren, B. W. (1976). *Statistical theory* (3rd ed.). New York: Macmillan.

338

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Marshall, J. L., & Haertel, E. H. (1976). *The mean split-half coefficient of agreement: A single administration index of reliability for mastery tests.* Unpublished manuscript, University of Wisconsin-Madison.

Mellenbergh, G. J., & van der Linden, W. J. (1981). The linear utility model for optimal selection. *Psychometrika, 46,* 283–293.

Mellenbergh, G. J. (1989, March). *Empirical specification of utility functions.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Novick, M. R., & Petersen, N. S. (1976). Towards equalizing educational and employment opportunity. *Journal of Educational Measurement, 13,* 77–88.

Peng, C. Y. J., & Subkoviak, M. J. (1980). A note on Huynh's normal approximation procedure for estimating criterion-referenced reliability. *Journal of Educational Measurement, 17,* 359–368.

Ortega, J. M., & Rheinholdt, W. C. (1970). *Iterative solution to nonlinear equations in several variables.* Orlando: Academic Press.

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. (1976). An expected utility model for "optimal" selection. *Journal of Educational Statistics, 4,* 333–358.

Petersen, N. S., & Novick, M. R. (1976). An evaluation of some models for culture-fair selection. *Journal of Educational Measurement, 13,* 3–31.

Subkoviak, M. J. (1976). Estimating reliability from a single administration of a mastery test. *Journal of Educational Measurement, 13,* 265–271.

Subkoviak, M. J. (1988). A practitioner's guide to computation and interpretation of reliability indices for mastery tests. *Journal of Educational Measurement, 25,* 47–55.

Swaminathan, H., Hambleton, R. K., & Algina, J. (1974). Reliability of criterion-referenced tests: A decision-theoretic formulation. *Journal of Educational Measurement, 11,* 263–268.

Swaminathan, H., Hambleton, R. K., & Algina, J. (1975). A Bayesian decision theoretic procedure for use with criterion-referenced tests. *Journal of Educational Measurement, 12,* 87–98.

van der Gaag, N. L. (1987, June). *Applications of empirical linear utility functions to mastery decisions.* Paper presented at the European meeting of the Psychometric Society, University of Twente, The Netherlands.

van der Gaag, N. L., Mellenbergh, G. J., & van den Brink, W. P. (1988). Empirical utility functions for pass/fail situations. *Methodika, 2*(1), 40–52.

van der Linden, W. J. (1980). Decision models for use with criterion-referenced tests. *Applied Psychological Measurement, 4,* 469–492.

van der Linden, W. J. (1981). Using aptitude measurements for the optimal assignment of subjects to treatments with and without mastery scores. *Psychometrika, 46,* 257–274.

van der Linden, W. J. (1985). Decision theory in educational research and testing. In T. Husen & T. N. Postlethwaite (Eds.), *International encyclopedia of education: Research and studies* (pp. 1328–1333). Oxford: Pergamon Press.

van der Linden, W. J. (1987). The use of test scores for classification decisions with threshold utility. *Journal of Educational Statistics, 12,* 62–75.

van der Linden, W. J. (1990). Applications of decision theory to test-based decision making. In R. K. Hambleton & J. N. Zaal (Eds.), *New developments in testing: Theory and applications* (pp. 129–155). Boston: Kluwer.

van der Linden, W. J., & Mellenbergh, G. J. (1977). Optimal cutting scores using a linear loss function. *Applied Psychological Measurement, 1,* 593–599.

van der Linden, W. J., & Mellenbergh, G. J. (1978). Coefficients for tests from a decision-theoretic point of view. *Applied Psychlogical Measurement, 2,* 119–134.

Vos, H. J. (1988). The use of decision theory in the Minnesota Adaptive Instructional System. *Journal of Computer-Based Instruction, 15*(2), 65–71.

Vos, H. J., & van der Linden, W. J. (1987). Designing optimal rules for instructional decision making in CAI systems. In J. Moonen & T. Plomp (Eds.), *Developments in educational software and courseware* (pp. 291–299). Oxford: Pergamon Press.

Vrijhof, B. J., Mellenbergh, G. J., & van den Brink, W. P. (1983). Assessing and studying utility functions in psychometric decision theory. *Applied Psychological Measurement, 7,* 341–357.

Author

HANS J. VOS is Assistant Professor, Dept. of Education, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands. He specializes in psychometric methods, data analysis, and research methodology.