PSYCHOMETRIKA—VOL. 61, NO. 1, 155–172
MARCH 1996

A COMPENSATORY APPROACH TO OPTIMAL SELECTION WITH MASTERY SCORES

WIM J. VAN DER LINDEN AND HANS J. VOS

UNIVERSITY OF TWENTE

A Bayesian approach for simultaneous optimization of test-based decisions is presented using the example of a selection decision for a treatment followed by a mastery decision. A distinction is made between weak and strong rules where, as opposed to strong rules, weak rules use prior test scores as collateral data. Conditions for monotonicity of optimal weak and strong rules are presented. It is shown that under mild conditions on the test score distributions and utility functions, weak rules are always compensatory by nature.

Key words: decision theory, mastery testing, monotone Bayes rules, selection decisions.

Introduction

Over the past two decades, Bayesian decision theory has proven to be very useful in solving problems of test-based decision making. Historically, the first decision making problem to draw the interest of psychometricians was the selection problem in education and personnel management. Important milestones in the history of the treatment of selection decisions were the publication of the Taylor-Russell (1939) tables and Cronbach and Gleser's (1956) *Psychological tests and personnel decisions*. However, in spite of some of the theoretical notions in the latter, it was not after an extensive discussion on "culture-fair" selection (Gross & Su, 1975) that selection decisions were fully treated as an instance of Bayesian decision theory (Novick & Petersen, 1976).

With the advance of such modern instructional systems as individualized study systems, mastery learning, and computer-aided instruction (CAI), interest was generated in the possibility to put the problem of mastery testing on sound decision-theoretic footing. In mastery testing, the intent is to classify examinees as "masters" or "non-masters" on the basis of their test scores, using some standard of mastery set on the true-score scale underlying the test scores. Hambleton and Novick (1973) were the first to point at the possibility of applying Bayesian decision theory to mastery testing. Optimal mastery rules for various utility or loss functions are derived in Davis, Hickman and Novick (1973), Huynh (1976, 1977, 1980) and van der Linden and Mellenbergh (1977).

Interest in decision making problems in modern instructional systems has also led to the consideration of two other types of decision making: placement and classification decisions. In either type of decision making, test scores are used to assign examinees to one of the instructional treatments available. However, with placement decisions the success of each of the treatments is measured by the *same* criterion whereas in classification decisions each treatment involves a *different* criterion. The paradigm underlying placement decisions is the Aptitude-Treatment Interaction (ATI) hypothesis,

The authors are indebted to Wilbert Kallenberg for his valuable comments and to Jan Gulmans for providing the data for the empirical example. The names of the authors are alphabetical; they are equally responsible for the contents of this paper.

Requests for reprints should be sent to Wim J. van der Linden, Department of Educational Measurement and Data Analysis, University of Twente, PO Box 217, 7500 AE Enschede, THE NETHERLANDS.

0033-3123/96/0300-94200$00.75/0

which assumes that students may react differentially to instructional treatments, and, therefore, that different treatments may be best for different students. Classification decisions are made if an instructional program has different tracks each characterized by different instructional objectives. Such tracking can be found in systems of comprehensive secondary education or vocational education. Bayesian decision theory for placement and classification decisions is given in Sawyer (1993) and van der Linden (1981, 1987).

Typically, instructional systems as CAI do not involve one single decision but can be conceived of as networks of nodes at which one of the types of decisions above has to be made (van der Linden, 1990; Vos, 1990, 1991, 1993). An example is an instructional network starting with a selection decision, followed by several alternative instructional modules through which students are guided by placement and mastery decisions, and which ends with a summative mastery test. Decisions in CAI networks are usually based on small tests (which often consist of only a few multiple-choice items).

The question is raised how such networks of decisions should be optimized. An obvious approach is to address each decision separately, optimizing its decision rule on the basis of test data exclusively gathered for this individual decision. This approach is common in current design of instructional systems. As an alternative, for a series of linearly related instructional modules Huynh and Perney (1979) proposed a backward computational scheme in which mastery scores at later modules are used to set mastery scores at earlier modules. The purpose of the present paper is to show that multiple decisions in networks can also be optimized *simultaneously*. The advantages of a simultaneous approach are twofold. First, data gathered earlier in the network can be used to optimize later decisions. The use of such prior information can be expected to enhance the quality of the decisions—in particular if only small tests or sets of multiple-choice items are administered at the individual decision points. Second, the option is now available to define utility or loss functions on the ultimate success criterion in the complete network instead of on intermediate criteria measuring the success on individual treatments. This may be realistic in cases where the success of earlier treatments is measured by achievements on later treatments, for example, when remedial teaching is considered a success if students do well in regular courses afterwards. Nevertheless, it is still possible to incorporate utility or costs related to earlier treatments into these functions.

In this paper, a simple decision network of a selection decision followed by one treatment and a mastery decision will be used to make our point. Study of more complicated networks is in progress; results for networks with more than one treatments are given in Vos (1994). First the selection-mastery problem will be formalized. Then important distinctions will be made between weak and strong as well as monotone and nonmonotone decision rules. Next, a theorem will be given showing under what conditions optimal rules will be monotone. Finally, results from an empirical example will be presented to illustrate the differences between a simultaneous and a separate approach.

It should be noted that though the present paper shares some aspects with the theory of multiple-objectives or multiple-criteria decision making (e.g., Keeney & Raiffa, 1976), important differences exist. In an application of the latter, for example, to the mastery testing problem, typically a multivariate criterion of mastery is measured and a multivariate mastery rule is defined on a test battery. In this paper, the criterion and each of the test scores involved are univariate. As will be clear below, the first decision in the network has a univariate decision rule. However, the second decision rule uses earlier test scores as prior information, and takes a multivariate form. More

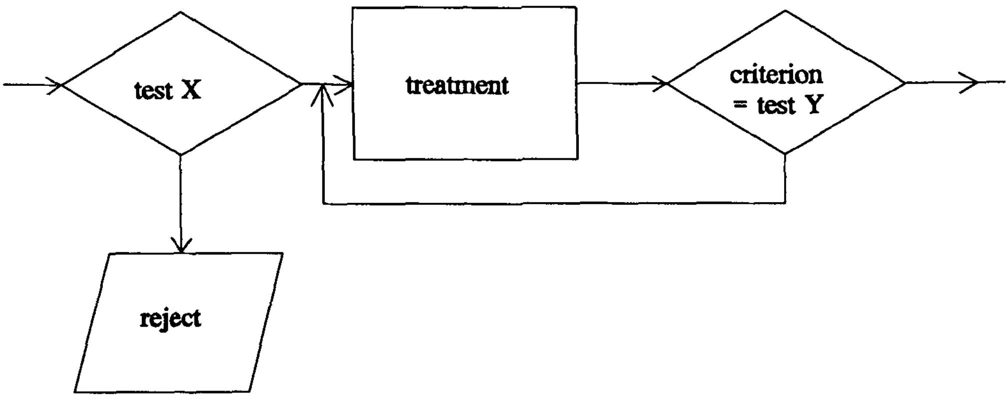

FIGURE 1.
A system of one selection and one mastery decision.

importantly, however, the problem of optimizing more than one rules simultaneously is different from the one of optimizing a single rule which is a function of more than one variable. In principle, it would be possible to generalize the problem addressed in this paper to a problem with multivariate selection and mastery tests. A general treatment of the multivariate mastery problem is given in Huynh (1982).

The Selection-Mastery Problem

A flowchart of the selection-mastery problem is given in Figure 1. An example of the problem is an instructional module with a pretest and a posttest. The pretest is administered to select students for the module. It is assumed that the possible actions are to admit or to reject the student for the module. The posttest is used to decide whether or not the students have mastered the objectives of the module. Typically, the posttest is not perfectly reliable, and the criterion is supposed to be a threshold on the true score underlying the test. The possible actions are to classify a student as a master or a nonmaster. The application addressed here is the case of a pretest used to assess whether students have acquired certain prerequisite skills and a posttest covering the subject matter taught in an instructional module. The two tests are thus not assumed to measure the same variable. Nevertheless, they have a statistical relation which allows for the use of the pretest as a source of additional information in decisions based on the posttest. The empirical example below is presented to show how standard psychometric theory can be used to model this statistical relation. If the two tests do measure the same variable, another model is possible in which the decision rule on the pretest also may lead to a mastery decision. The statistical aspects of this model, which involves different monotonicity problems, are addressed in van der Linden (1995).

The following notation is needed. For a randomly sampled student, the observed scores on the selection and mastery tests are continuous random variables denoted by X and Y, with realizations x and y, respectively. The criterion considered is the classical test theory true score underlying the mastery test. For a randomly sampled student, the true score is denoted by a continuous random variable T with realization t. It is assumed that the standard denoting true mastery is a threshold value t_c on T. Further, it will be assumed that the relation between X, Y, and T can be represented by a joint density function $f(x, y, t)$. The best experiment to estimate the parameters in this density function is the one in which a sample of examinees from the full marginal distribution of X is admitted to the treatment and the performances of these students on the mastery test Y are measured. However, not much efficiency need be lost if the parameters in the density function have to be estimated from a sample censored on the left due to the fact that low performing students are not admitted to the treatment, provided the correct density function is chosen. The statistical theory needed for estimating from censored samples is given in Kendall and Stuart (1979, secs. 32.15–32.21). A useful correction for Pearson's correlation coefficient between X and Y is given in Roe (1979), whereas a Bayesian approach to the problem is provided in Brouwer and Vijn (1979).

Simultaneous Decision Rules

Let each of the possible actions be denoted by a_{ij} $(i, j = 0, 1)$, where $i = 0, 1$ stand for the actions of rejecting and accepting a student and $j = 0, 1$ for the actions of retaining and advancing an accepted student. Since for a rejected student no further mastery decisions are made, the index j will be dropped for $i = 0$.

Generally, a decision rule specifies for each possible realization (x, y) of (X, Y) which action a_{ij} is to be taken.

Weak and Strong Rules

The decision rule for the mastery decision may or may not depend on the score X on the selection test. Intuitively, one would expect a more lenient mastery rule for a student with a high performance on the selection test because this prior information implies that a possible low score on the mastery test is more likely due to measurement error than to a true low performance. Simultaneous rules in which decisions are a function both of the current test score and previous test scores test will be called weak rules in this paper. As a general result, it will be proven that under obvious conditions weak rules will necessarily have a compensatory nature. The title of the paper already alludes to this result.

If decisions are only a function of current test scores, the rules will be called strong (simultaneous) rules.

For the decision network of Figure 1 a weak simultaneous rule δ can be defined as:

$$
\begin{aligned}
&\{(x, y)\colon\ \delta(x, y) = a_0\} = A \times R \\
&\{(x, y)\colon\ \delta(x, y) = a_{10}\} = A^C \times B(x) \\
&\{(x, y)\colon\ \delta(x, y) = a_{11}\} = A^C \times B^C(x),
\end{aligned}
\tag{1}
$$

where A and A^C are the sets of x values for which a student is rejected or admitted for the treatment, and $B(x)$ and $B^C(x)$ are the sets of y values for which a students fails or passes the mastery test. R represents the set of real numbers.

With strong rules, the sets $B(x)$ and $B^C(x)$ are independent of x. Strong simultaneous rules can only be optimal if certain conditions are met. These conditions will be given below.

Monotone and Nonmonotone Rules

Decision rules can take a monotone or a nonmonotone form. A decision rule is monotone if cutting scores are used to partition the sample space into regions for which different actions are taken. For example, a (separate) rule for the selection decision is monotone if there exists a cutting score x_c such that all examinees with $X \geq x_c$ are admitted and those with $X < x_c$ are rejected. All other possible rules are nonmonotone.

For our decision problem, a weak monotone rule δ can be defined as:

$$\delta(X, Y) = \begin{cases} a_0 & \text{for } X < x_c \\ a_{10} & \text{for } X \geq x_c,\ Y < y_c(x) \\ a_{11} & \text{for } X \geq x_c,\ Y \geq y_c(x), \end{cases} \tag{2}$$

with $y_c(x)$ being the cutting score on Y. The fact that this cutting score is written as a mathematical function of x will be justified below proving that $y_c(x)$ is unique for each value of x under reasonable assumptions.

In this paper, the interest will mainly be in monotone rules. The reason for this choice is the fact that the use of cutting scores is common practice in educational and psychological testing, and that, for example, rules which reject students with low or high scores but admit students to a program with scores in the middle of the scale would generally not be considered acceptable. However, the restriction to monotone rules is correct only if it can be proven that for any nonmonotone rule for the problem at hand there is a monotone rule with at least the same value on the criterion of optimality used; that is, if the subclass of monotone rules is essentially complete (Ferguson, 1967, p. 55). Conditions under which the subclass of monotone (simultaneous) rules is essentially complete for the present problem will also be given below.

Strong Monotone Rules with Maximum Expected Utility (SMMEU)

To evaluate the use of cutting scores even if conditions for monotonicity are not known to hold, the case of Strong Monotone Rules with Maximum Expected Utility (SMMEU rules) is also considered. An SMMEU rule is a rule with maximum expected utility in the subclass of strong monotone rules. The attention for SMMEU rules is motivated by the fact that educators are familiar with cutting scores as decision rules and that not each of them has a tradition of bothering about conditions for monotonicity.

Thus, if the sets of conditions for both strong and monotone rules to be optimal are satisfied, the subclasses of SMMEU and strong monotone Bayes rules are identical. Otherwise, they differ.

Utility Structure

Generally, a utility function describes the utility of each possible action for the possible true states of nature. Here, the utilities involved in the combined decision problem are defined as the following additive structure

$$u_{ij}(t) = w_1 u_i^{(s)}(t) + w_2 u_j^{(m)}(t), \tag{3}$$

where $u_i^{(s)}(t)$ and $u_j^{(m)}(t)$ represent the utility functions for the separate selection and mastery decisions, respectively, and w_1 and w_2 are nonnegative weights. Since utility is supposed to be measured on an interval scale, the weights in (3) can always be rescaled as follows:

$$u_{ij}(t) = w u_i^{(s)}(t) + (1 - w) u_j^{(m)}(t), \tag{4}$$

where $0 \leq w \leq 1$.

Since no mastery decisions are made for rejected students, it is assumed that such students do not contribute to the utility. This assumption is consistent with the institutional point of view taken here. Hence, it follows from (4) that $u_{0j}(t)$ is equal to $w u_0^{(s)}(t)$ for all j.

It should be noted that the first term of (4) is a function of t and not, for example, of a true score underlying X. This fact illustrates one of the advantages of a simultaneous approach to decision making, namely, that there is no need to resort to intermediate criteria of success but that for all decisions utility can be defined as a function of the ultimate criterion in the network. Also, note that both terms in (4) are a function of the same variable. The utility structure assumed is therefore not an instance of a utility function in a multiple-criteria decision problem. The usual assumptions needed to motivate an additive structure for multiple-criteria functions (Keeney & Raiffa, 1976) are thus not needed here.

The weight factor w in (4) has been introduced to accommodate cases where the utilities for the two decisions are measured in different units. If the units are the same, w is set equal to .5.

Methods for establishing empirical utility functions for test-based decisions have been studied in van der Gaag (1990) and Vrijhof, Mellenbergh, and van den Brink (1983). The dominant conclusion from the series of studies of empirical utilities in selection and mastery decisions in these references is that a choice from the family of linear utility functions is often realistic for both types of decisions. This choice will be made in the empirical example below. Obviously, these functions will be chosen such that utility will be an increasing function of t for the admittance and mastery decision but decreasing functions for the rejection and nonmastery decision. First, however, more general results will be presented.

Expected Utility in the Simultaneous Approach

For the decision rules in (1) and the utility structure in (4), the expected utility for the two decision rules is equal to,

$$
\begin{aligned}
E[U_{sim}(A^C, B^C(x))] \equiv & \int_A \int_R \int_R wu_0^{(s)}(t) f(x, y, t)\, dt\, dy\, dx \\
& + \int_{A^C} \int_{B(x)} \int_R u_{10}(t) f(x, y, t)\, dt\, dy\, dx \\
& + \int_{A^C} \int_{B^C(x)} \int_R u_{11}(t) f(x, y, t)\, dt\, dy\, dx. \qquad (5)
\end{aligned}
$$

In a Bayesian fashion, the expected utility in (5) will be taken as the criterion of optimality in this paper.

Taking expectations, completing integrals, and rearranging terms, (5) can be written as

$$
\begin{aligned}
E[U_{sim}(A^C, B^C(x))] = wE[u_0^{(s)}(T)] + & \int_{A^C} \{E[u_{10}(T) - wu_0^{(s)}(T) | x] \\
& + \int_{B^C(x)} E[u_{11}(T) - u_{10}(T) | x, y] h(y|x)\, dy\} q(x)\, dx, \qquad (6)
\end{aligned}
$$

where $q(x)$ and $h(y|x)$ denote the p.d.f.'s of X and Y given $X = x$.

It is interesting to note that the critical quantities in (6) are the posterior expected utilities given $X = x$ and $(X = x, Y = y)$. It is through these quantities that information from prior tests will play a role in later decisions in the network.

Sufficient Conditions for Monotone Rules

In this section, monotonicity conditions for the simultaneous rules are derived. First, sufficient and necessary conditions for monotone solutions for the separate selection and mastery decisions will be given. Next, sufficient conditions for weak monotone solutions will be derived. Finally, monotonicity conditions for strong simultaneous rules will be derived from the previous case by imposing additional restrictions on the test-score distributions.

Conditions for Separate Selection and Mastery Decisions

Conditions for selection and mastery rules to be (strictly) monotone are given in Chuang, Chen and Novick (1981). Two sets of conditions must be met. First, the families of distributions of the true scores T given $X = x$ and T given $Y = y$ must be stochastically increasing; that is, their cumulative distribution functions (c.d.f.'s) must be decreasing in x and y for all t. Second, the utility functions must be monotone. This condition requires the difference between the utility function for the rejection (nonmastery) and admittance (mastery) decision to change sign at most once.

Both conditions immediately follow from the standard decision problem addressed in statistical decision theory (e.g., Ferguson, 1967; Lindgren, 1976).

Conditions for Weak Simultaneous Rules

Let $V(t|x, y)$ denote the c.d.f. of T given ($X = x$ and $Y = y$) and $H(y|x)$ the c.d.f. of Y given $X = x$. The following theorem gives a set of conditions sufficient for a weak monotone solution:

Theorem. An optimal simultaneous decision rule for the selection-mastery problem with additive utility is (weak) monotone if:

$$u_1^{(m)}(t) - u_0^{(m)}(t) \text{ is strictly increasing in } t, \tag{7}$$

$$u_{10}(t) - wu_0^{(s)}(t) \text{ is strictly increasing in } t, \tag{8}$$

$$V(t|x, y) \text{ is strictly decreasing in } x \text{ and } y \text{ for all } t, \tag{9}$$

$$H(y|x) \text{ is strictly decreasing in } x \text{ for all } y. \tag{10}$$

The first condition guarantees monotone utility for the mastery decision.

The second condition stipulates that the difference between the utility functions for the actions a_{10} (acceptance, nonmastery) and a_0 (rejection) be an increasing function of t.

The third condition requires double (strict) stochastic increasingness for the distribution of T given $X = x$ and $Y = y$. Loosely speaking, this condition is met if high true scores on the mastery test coincide with high observed scores on both the selection and mastery tests.

The last condition also requires (strict) stochastic increasingness, and thus that high scores on the mastery and selection test tend to coincide.

Not all conditions in this set are straightforward generalizations of the conditions for the separate decision problems. In particular, the conditions in (8) and (10) are new; they are needed to link the two separate decision problems.

It should be noted that there is no condition analogous to (7) for the selection problem. This is due to the fact that the utility component for this problem is defined on the true score variable for the mastery test.

In the proof of the theorem, the following lemma's are needed:

Lemma 1. Let $f(x)$ be an arbitrary function with $\int |f(x)|\, dx < \infty$, then for any set S of x values it holds that $\int_S f(x)\, dx \leq \int_{S'} f(x)\, dx$ with $S' = \{x: f(x) \geq 0\}$ (e.g., Ferguson, 1967, p. 201).

Lemma 2. For any increasing function $k(t)$, the expectation $E[k(T)|z]$ is an increasing function of z if and only if the c.d.f. of T given $Z = z$ decreases in z for all t (e.g., Lehmann, 1986, p. 116).

For future use, it is observed that if $k(t)$ is a constant, $E[k(T)|z]$ is a constant too. Hence, the nondecreasing version of the lemma also holds.

Lemma 3. If (9) and (10) hold, then the marginal c.d.f. $P(t|x)$ associated with $V(t|x, y)$ is decreasing in x for all t.

Proof of Lemma 3. Let $v(t|x, y)$ be the p.d.f. of T given $X = x$ and $Y = y$. By definition, $1 - P(t|x) = \int_t^\infty \int_{-\infty}^{\infty} v(z|x, y)h(y|x)\, dy\, dz = \int_{-\infty}^{\infty} [1 - V(t|x, y)]h(y|x)\, dy$. From (9)–(10) and Lemma 2, it follows that $1 - P(t|x)$ increases in x for all t, that is, the distribution of T given $X = x$ is stochastically increasing in x. □

For completeness' sake, it is observed that the distribution of T given $Y = y$ is also stochastically increasing in y if (10) is replaced by the stronger condition of monotone likelihood ratio in y w.r.t. x. However, this result is not needed in the remainder of this paper.

Lemma 4. If a function $\kappa(x, y)$ is (strictly) increasing in x and y, then the relation defined by $C = \{(x, y)|\kappa(x, y) = c,\ c \in R\}$ is a decreasing function in x.

Proof of Lemma 4. Assume that there are two pairs $(x_1, y_1) \in C$ and $(x_2, y_2) \in C$ with $x_2 > x_1$, for which $y_2 \geq y_1$. Then, by hypothesis, $\kappa(x_2, y_2) > \kappa(x_1, y_1)$, which contradicts the assumption. □

Proof of Theorem. Let the set $B_0^C(x)$ be defined by

$$B_0^C(x) \equiv \{y: E[u_{11}(T) - u_{10}(T)|x, y] \geq 0\}. \tag{11}$$

Applying Lemma 1 to the second term in the integral in (6), and using $h(y|x) \geq 0$, it follows that for all $B^C(x)$ and an arbitrary but fixed A^C:

$$E[U_{sim}(A^C, B^C(x))] \leq wE[u_0^{(s)}(T)] + \int_{A^C} \{E[u_{10}(T) - wu_0^{(s)}(T)|x] + \int_{B_0^C(x)} E[u_{11}(T) - u_{10}(T)|x, y]h(y|x)\, dy\}q(x)\, dx. \tag{12}$$

Using the following definition of the set A_0^C

$$A_0^C \equiv \left\{ x\colon E[u_{10}(T) - wu_0^{(s)}(T)|x] + \int_{B_0^C(x)} E[u_{11}(T) - u_{10}(T)|x, y]h(y|x)\, dy \geq 0 \right\}, \tag{13}$$

and applying the lemma to the second term in the right-hand side of (6), it follows that for all A^C

$$E[U_{sim}(A^C, B_0^C(x))] \leq wE[u_0^{(s)}(T)] + \int_{A_0^C} \{E[u_{10}(T) - wu_0^{(s)}(T)|x] + \int_{B_0^C(x)} E[u_{11}(T) - u_{10}(T)|x, y]h(y|x)\, dy\}q(x)\, dx. \tag{14}$$

It is now proven that the left-hand sides of the inequalities in (11) and (13) increase in y for all x and in x, respectively.

i. Since $u_{11}(t) - u_{10}(t) = (1 - w)[u_1^{(m)}(t) - u_0^{(m)}(t)]$ and $1 - w \geq 0$, it follows from the condition in (7) that the difference between these two utilities is increasing too. Therefore, (9) and Lemma 2 together imply that

$$E[u_{11}(T) - u_{10}(T)|x, y] \text{ is increasing in } y \text{ for all } x \text{ and in } x \text{ for all } y, \tag{15}$$

and thus that the sets $B_0^C(x)$ take the form $[y_c(x), \infty)$ for all values of x. This result will be used in the following part of the proof.

ii. From (8) through (10) and Lemma's 2 and 3, it follows immediately that the first term in the left-hand side of (13) is increasing in x. For notational convenience, the term $E[u_{11}(T) - u_{10}(T)|x, y]$ is denoted as $\tau(x, y)$. Note that $\tau(x, y)$ is an increasing function of y which is nonnegative for $y \geq y_c(x)$ for all values of x. Now for any $x_2 > x_1$, it follows that

$$\begin{aligned}
&\int_{y_c(x_2)} \tau(x_2, y)h(y|x_2)\, dy - \int_{y_c(x_1)} \tau(x_1, y)h(y|x_1)\, dy \\
&\quad > \int_{y_c(x_1)} \tau(x_2, y)h(y|x_2)\, dy - \int_{y_c(x_1)} \tau(x_1, y)h(y|x_1)\, dy \\
&\quad > \int_{y_c(x_1)} \tau(x_1, y)[h(y|x_2) - h(y|x_1)]\, dy \\
&\quad = \int_{-\infty}^{\infty} \varphi(y)[h(y|x_2) - h(y|x_1)]\, dy,
\end{aligned} \tag{16}$$

where $\varphi(y) \equiv I_{[y_c(x_1),\infty)}(y)\tau(x_1, y)$, and $I_{[y_c(x_1),\infty)}(y)$ is an indicator function which takes the value 1 if $y \in [y_c(x_1), \infty)$ and the value 0 otherwise. In the first step in (16), Lemma 4 is used in combination with (15), whereas the second step follows from the fact that $\tau(x, y)$ is increasing in x for all y. By definition, $\varphi(y)$ is a nondecreasing function of y, and it follows from (10) and Lemma 2 that the final result in (16) is positive. Hence, it can be concluded that the second left-hand term in (13) is increasing in x, and thus that the set A_0^C takes the form $[x_c, \infty)$. Since (6) is maximal for the sets $A_0^C = [x_c, \infty)$ and $B_0^C(x) = [y_c(x), \infty)$, the expected utility is maximized by use of the monotone rules represented by these sets. □

The cutting scores x_c and $y_c(x)$ are the values of x and y for which the inequalities in (11) and (13) become equalities. These values may be infinitely small or large implying that the same decisions have to be made for all examinees.

Monotonicity Conditions for Strong Simultaneous Rules

For strong simultaneous rules, $B_0^C(x)$ is not allowed to depend on x. Therefore, as an additional condition, it must hold for $v(t|x, y)$ and the p.d.f. of T given $Y = y$ that

$$v(t|x, y) = g(t|y). \tag{17}$$

This condition, which immediately follows from (11), implies that all information on T relevant for the decision is contained in $Y = y$, and that, once $Y = y$ is given, the observation $X = x$ does not add any information. If the condition holds, then, obviously, the use of simultaneous rules will not add any efficiency to the decision making procedure.

Calculation of Simultaneous Rules

From the theorem it follows that the optimal weak and strong simultaneous rules can be calculated from the left-hand sides of the inequalities in (11) and (13). To obtain optimal weak rules, first x_c and $y_c(x_c)$ are calculated simultaneously by solving for the values of x and y that render these inequalities to equalities. Then for $x \gtrsim x_c$, $y_c(x)$ is obtained by putting the left-hand side of (11) equal to zero and solving for y.

If the additional conditions for optimal strong rules are met, B_0^C replaces $B_0^C(x)$ in the expression of the expected utility, and x_c and y_c are obtained by simultaneously solving (11) and (13) for the values of x and y that turn these inequalities into equalities.

Likewise, to calculate SMMEU rules the set $B^C(x)$ is replaced by B^C in (6), and the system of equations consisting of the partial derivatives of (6) w.r.t. x_c and y_c equated to zero is solved.

In the empirical example below, for the calculation of all cutting scores Newton's method for solving nonlinear systems was used. The method was implemented in a computer program called NEWTON. Another program, UTILITY, was written to analyze differences in expected utility for the various rules. Copies of the programs are available from the authors of the paper upon request.

Discrete Test Scores

So far the scores on the selection, X, and mastery test, Y, have been taken to be continuous. However, in practice they are discrete. From the monotonicity of the left-hand sides of (11) and (13) it follows that a solution to the discrete problem consists of the largest integer values below or the smallest integer values above x_c and $y_c(x_c)$. Substitution of these values into (6) show for which values the expected utility is maximal.

Optimal Separate Rules

It is observed that optimal rules for the separate decisions can easily be found by imposing certain restrictions on $E[U_{sim}(A^C, B^C(x))]$.

First, substituting $w = 1$ into (6), the expected utility for the separate selection decision $E[U^{(s)}(A^C)]$, can be written as

$$E[U^{(s)}(A^C)] = E[u_0^{(s)}(T)] + \int_{A^C} E[u_1^{(s)}(T) - u_0^{(s)}(T)|x]q(x)\ dx. \tag{18}$$

Next, substituting $w = 0$, $A^C = R$ (i.e., accepting all students for the instructional treatment), and $B^C(x) = B^C$ into (6) gives the following result for the expected utility of the separate mastery decision:

$$E[U^{(m)}(B^C)] = E[u_0^{(m)}(T)] + \int_{B^C} E[u_1^{(m)}(T) - u_0^{(m)}(T)|y]s(y)\ dy, \tag{19}$$

where $s(y)$ denotes the p.d.f. of Y.

Analogous to the simultaneous approach, it can easily be verified from Lemma 1 that upper bounds to $E[U^{(s)}(A^C)]$ and $E[U^{(m)}(B^C)]$ are obtained for the sets of x and y values for which $E[u_1^{(s)}(T) - u_0^{(s)}(T)|x]$ and $E[u_1^{(m)}(T) - u_0^{(m)}(T)|y]$ are nonnegative, respectively. Assuming that the monotonicity conditions for the separate decisions are satisfied, the optimal cutting scores for the separate selection and mastery decisions, say $\tilde{x}_c$ and $\tilde{y}_c$, can be obtained by solving

$$E[u_1^{(s)}(T) - u_0^{(s)}(T)|x] = 0$$

and

$$E[u_1^{(m)}(T) - u_0^{(m)}(T)|y] = 0$$

for x and y, respectively. For further details, see Mellenbergh and van der Linden (1981) and van der Linden and Mellenbergh (1977).

An Empirical Example

Optimal rules were calculated for a selection-mastery decision problem consisting of a CAI module on elementary medical knowledge preceded and followed by a selection and mastery test, respectively. Both tests consisted of 21 items and had possible test scores ranging from 0–100. Data were available for a sample of 76 freshmen in a medical program. A sample of this size is only used to illustrate the techniques in this paper but is not recommended for use with decision problems in the practice of educational measurement. The instructors in the program considered students as having mastered the module if their true scores were larger than 55. Therefore, t_c was fixed at this value. All students in the program were admitted to the instructional module, therefore the samples of the score distributions involved did not suffer from any restriction of range.

Score Distributions

It was assumed that (X, Y, T) followed a trivariate normal distribution. Under this assumption, the bivariate distribution of (X, Y) is also normal. Further, the regression function $E(Y|x)$ is linear. It should be noted that bivariate normality is sufficient for linear regression.

The two observable consequences were tested against the data using a chi-square and a t-test. The probabilities of exceedance were 0.219 and 0.034, showing a satisfactory fit which confirmed our visual inspection of various plots of the distributions. The result for the chi-square test was obtained partitioning the sample space into 20 regions of (x, y) values which were, except for the regions at the tails, taken to be of equal size. The number of degrees of freedom was equal to $\text{df} = 20 - 5 - 1 = 14$.

Standard results from classical test theory (Lord & Novick, 1968, chap. 2) were used to express the conditional expectations and variances of T given x and/or y as functions of observable quantities. It was assumed that $E(T|y)$ followed Kelley's regression line (Lord & Novick, 1968, sect. 3.7). Therefore, only the following two regression equations needed special attention:

$$E(T|x) = E(Y|x) = \alpha_{TX} + \beta_{TX}x; \tag{20}$$

$$E(T|x, y) = \alpha_{TXY} + \beta_{TXY}x + \gamma_{TXY}y. \tag{21}$$

The parameters in these regression equations can be written as:

$$\begin{aligned}
\beta_{TX} &= \rho_{XY}\left(\frac{\sigma_Y}{\sigma_X}\right);\\
\alpha_{TX} &= \mu_Y - \beta_{TX}\mu_X;\\
\beta_{TXY} &= \rho_{XY}\left[\frac{\sigma_Y}{\sigma_X}\right]\left[\frac{1-\rho_{YY'}}{1-\rho_{XY}^2}\right];\\
\gamma_{TXY} &= \frac{\rho_{YY'} - \rho_{XY}^2}{1-\rho_{XY}^2};\\
\alpha_{TXY} &= \mu_Y(1-\gamma_{TXY}) - \beta_{TXY}\mu_X.
\end{aligned} \tag{22}$$

In addition, assuming homoscedasticity, it holds for the conditional variances Var $(T|x)$, Var $(T|y)$, and Var $(T|x, y)$ that:

$$\begin{aligned}
\text{Var }(T|x) &= (\rho_{YY'} - \rho_{XY}^2)\sigma_Y^2;\\
\text{Var }(T|y) &= \sigma_Y^2\rho_{YY'}(1-\rho_{YY'});\ \text{and}\\
\text{Var }(T|x, y) &= \frac{\sigma_Y^2(1-\rho_{YY'})(\rho_{YY'} - \rho_{XY}^2)}{1-\rho_{XY}^2}.
\end{aligned} \tag{23}$$

The statistics in Table 1 are estimates of the means, variances, and correlation in the above equations whereas the reliability estimates are Cronbach's alpha's. The estimates were substituted into the equations above to get the required estimates of the regression parameters and the conditional variances.

Utility Structure

The following choice was made for the functions $u_i^{(s)}(t)$ and $u_j^{(m)}(t)$ in (4):

$$u_i^{(s)}(t) = \begin{cases} b_0^{(s)}(t_c - t) + d_0^{(s)} & \text{for } i = 0\\ b_1^{(s)}(t - t_c) + d_1^{(s)} & \text{for } i = 1 \end{cases} \tag{24}$$

$$u_j^{(m)}(t) = \begin{cases} b_0^{(m)}(t_c - t) + d_0^{(m)} & \text{for } j = 0\\ b_1^{(m)}(t - t_c) + d_1^{(m)} & \text{for } j = 1 \end{cases} \tag{25}$$

TABLE 1

Some Statistics of the Selection and Mastery Tests (X and Y)

Statistics	X		Y
Mean	50.679		62.436
Standard Deviation	8.781		9.456
Reliability	0.773		0.802
Correlation		0.751	

where $b_i^{(s)}$, $b_j^{(m)} > 0$ $(i, j = 0, 1)$. The parameters $d_i^{(s)}$ and $d_j^{(m)}$ can represent, for example, the fixed amount of costs involved in following an instructional module and testing the examinees. The condition $b_0^{(s)}$, $b_1^{(s)} > 0$ states that utility be a decreasing function for the rejection decision, but an increasing function for the acceptance decision. Similarly, the condition $b_0^{(m)}$, $b_1^{(m)} > 0$ expresses that the utilities associated with failing and passing the mastery test be decreasing and increasing functions in t, respectively.

The same utility functions were used in an analysis of separate selection and mastery decisions in Mellenbergh and van der Linden (1981) and van der Linden and Mellenbergh (1977). For other possible utility functions, see Novick and Lindley (1979).

Monotonicity Conditions

The condition in (7) is met since $b_j^{(m)} > 0, j = 0, 1$.

It can easily be verified that the condition in (8) is satisfied if the weight w and the parameters $b_0^{(s)}$, $b_1^{(s)}$, and $b_0^{(m)}$ are chosen such that

$$w > \frac{b_0^{(m)}}{b_0^{(s)} + b_1^{(s)} + b_0^{(m)}}. \tag{26}$$

All numerical values for the utility parameters in the example were chosen to meet these two requirements.

Under the model of a trivariate normal distribution for (X, Y, T) in this example, the conditions in (9) and (10) were met by the positive slopes of the regression lines and planes in this distribution.

Finally, the additional condition for solutions to be strong monotone in (17) was tested comparing the two regression lines $E(T|x, y)$ and $E(T|y)$ using an F-test. The probability of exceedance was 0.038, indicating that the result was significant at $\alpha = .05$. Therefore, only SMMEU rules and no optimal strong rules were considered.

Results for the Simultaneous Rules

For several values of the utility parameters, weak monotone and SMMEU rules were calculated. The results are reported in Table 2, where the cutting scores for the SMMEU rules are denoted as x_c^* and y_c^*. It is reminded that the results were obtained from a small number of examinees, and thus are sensitive to sampling error.

As is clear from the results, the consequence of increasing the values of the parameters $b_i^{(s)}$ was a decrease of the optimal weak and SMMEU cutting scores on the selection test. This relation holds generally. Larger values for these parameters represent the fact that acceptance decisions are evaluated relatively stronger, and therefore result in lower cutting scores.

A decrease of the amount of constant utility, $d_i^{(s)}$ and $d_j^{(m)}$, resulted in increases of the optimal weak and SMMEU cutting scores on the selection test. For the parameter w it was found that the optimal weak and SMMEU cutting scores on the selection test increase in w for utility structures (1) through (3) and (4) through (6) in Table 2, whereas the opposite holds for utility structures (7) through (9) in the table. These effects are the result of the specific values of the parameters in the study and cannot be generalized.

The effect of the reliability of the mastery test on the behavior of the cutting score for the mastery decision is discussed below.

Results for the Separate Approach

The monotonicity conditions for the separate selection and mastery decision problems are implied by those for the simultaneous problem as well as the condition that $b_i^{(s)} > 0$, $i = 0, 1$. They were therefore satisfied. The optimal cutting scores $\tilde{x}_c$ and $\tilde{y}_c$ for the two separate decision problems are also reported in Table 2. In particular for $w = 0.3$, the weak cutting scores $y_c(x_c)$ on the mastery test generally were high compared with $\tilde{y}_c$.

For the selection test the results did not differ much from those obtained for the weak monotone rules. This fact can be explained as follows: Students who were just accepted in the case of a weak monotone rule had to compensate their rather low cutting scores on the selection test with relatively high scores on the mastery test compared with students accepted in the case of separate rules. However, the decreasing character of $y_c(x)$ in x implied that only students accepted with selection scores equal to or just above x_c did need these rather high scores on the mastery test to reach the mastery status.

Comparison of the Expected Utilities

For the simultaneous approach a gain in expected utility relative to the separate approach was expected. To check whether this expectation could be confirmed, the weighted sum of the expected utilities for the optimal separate rules was compared with the expected utilities for the optimal weak monotone rules. The results are also displayed in Table 2.

It can be seen that the expected utilities for the optimal weak monotone rules yielded the largest values for all utility structures. This result was in accordance with our expectations. Furthermore, Table 2 indicates that the expected utilities for the optimal weak monotone rules were only slightly larger than for the SMMEU rules. This result follows immediately from an analysis of (6). The first term of this expression is always the same constant for the two rules. The fact that x_c and x_c^* were nearly equal implies that the first part of the second term in (6) cannot differ much for the two rules, whereas, for the current data, the last term of (6) appeared to contribute hardly to its

TABLE 2
Optimal Rules and Expected Utilities for both Separate and Simultaneous Rules

No.	Utility Specifications		w	Cutting Scores			Expected Utility		
				Simultaneous		Separate			
				Weak Monotone	SMMEU		Weak Monotone	SMMEU	Separate
(1)	$b_0^{(s)} = 3$ $b_1^{(s)} = 5$ $b_0^{(m)} = 2$ $b_1^{(m)} = 3$	$d_0^{(s)} = -2$ $d_1^{(s)} = -3$ $d_0^{(m)} = -4$ $d_1^{(m)} = -5$	0.3	$x_c = 40.85$ $y_c(x_c) = 55.80$	$x_c^* = 41.25$ $y_c^* = 51.88$	$\tilde{x}_c = 41.64$ $\tilde{y}_c = 53.41$	27.84	27.76	26.32
(2)			0.6	$x_c = 41.41$ $y_c(x_c) = 55.42$	$x_c^* = 41.51$ $y_c^* = 51.79$	$\tilde{x}_c = 41.64$ $\tilde{y}_c = 53.41$	32.64	32.60	31.70
(3)			0.9	$x_c = 41.60$ $y_c(x_c) = 55.29$	$x_c^* = 41.62$ $y_c^* = 51.76$	$\tilde{x}_c = 41.64$ $\tilde{y}_c = 53.41$	37.61	37.59	37.08
(4)	$b_0^{(s)} = 3$ $b_1^{(s)} = 5$ $b_0^{(m)} = 2$ $b_1^{(m)} = 3$	$d_0^{(s)} = 0$ $d_1^{(s)} = 0$ $d_0^{(m)} = 0$ $d_1^{(m)} = 0$	0.3	$x_c = 38.95$ $y_c(x_c) = 56.72$	$x_c^* = 39.65$ $y_c^* = 52.02$	$\tilde{x}_c = 41.49$ $\tilde{y}_c = 53.16$	32.08	31.98	30.54
(5)			0.6	$x_c = 40.81$ $y_c(x_c) = 55.46$	$x_c^* = 40.93$ $y_c^* = 51.67$	$\tilde{x}_c = 41.49$ $\tilde{y}_c = 53.16$	36.38	36.35	35.32
(6)			0.9	$x_c = 41.37$ $y_c(x_c) = 55.08$	$x_c^* = 41.39$ $y_c^* = 51.52$	$\tilde{x}_c = 41.49$ $\tilde{y}_c = 53.16$	40.67	40.66	40.11
(7)	$b_0^{(s)} = 2$ $b_1^{(s)} = 4$ $b_0^{(m)} = 1$ $b_1^{(m)} = 2$	$d_0^{(s)} = -2$ $d_1^{(s)} = -3$ $d_0^{(m)} = -4$ $d_1^{(m)} = -5$	0.3	$x_c = 41.83$ $y_c(x_c) = 55.38$	$x_c^* = 42.06$ $y_c^* = 51.82$	$\tilde{x}_c = 41.70$ $\tilde{y}_c = 53.58$	19.14	19.05	17.58
(8)			0.6	$x_c = 41.74$ $y_c(x_c) = 55.44$	$x_c^* = 41.82$ $y_c^* = 51.90$	$\tilde{x}_c = 41.70$ $\tilde{y}_c = 53.58$	24.08	24.05	23.03
(9)			0.9	$x_c = 41.70$ $y_c(x_c) = 55.47$	$x_c^* = 41.72$ $y_c^* = 51.93$	$\tilde{x}_c = 41.70$ $\tilde{y}_c = 53.58$	29.04	20.02	28.49

value. In addition, it should be noted that, though the statistical tests suggest a satisfactory fit of the data to the assumption of trivariate normality, such assumptions are never perfectly met. Finally, the table shows that for all three approaches, the expected utility yielded the largest value for $w = 0.9$. In other words, the utility for the selection decision contributed most to the expected utility for the optimal rules in this study.

Concluding Remarks

For a monotone utility structure, Lemma 4 shows that under the natural condition of the distributions of selection and mastery test scores being stochastically increasing in the true score on the mastery test, optimal cutting scores for the mastery test under the weak simultaneous rules are a decreasing function of the scores on the selection test. As already explained, this feature introduces an element of compensation in the decision procedure: It is possible to compensate low scores on the mastery test by high scores on the selection test. A quantitative estimate of this effect can be calculated for the data set in the empirical example above. Substituting the estimated regression plane (21) into the left-hand-side in (11) and solving for y yields

$$y_c(x) = \frac{\dfrac{d_0^{(m)} - d_1^{(m)}}{b_0^{(m)} + b_1^{(m)}} + t_c - \alpha_{TXY} - \beta_{TXY}x}{\gamma_{TXY}}. \tag{27}$$

The derivative of this equation w.r.t. x is equal to $-\beta_{TXY}/\gamma_{TXY}$, which for the data set was estimated as $-.675$. It follows for all utility structures in this example that the cutting score $y_c(x)$ on the mastery test has to be lowered by .675 for each score point above x_c on the selection test. Obviously, the size of this effect is a function of the reliability of the mastery test. If the reliability is low, the compensatory effect of the selection test is high, whereas for high reliability less compensation is necessary. For a perfectly reliable mastery test with $\rho_{YY'} = 1$, it is easy to show from (22) that $\gamma_{TXY} = 1$, and $\alpha_{TXY} = \beta_{TXY} = 0$, and hence that (27) becomes identical to $y_c = [d_0^{(m)} - d_1^{(m)})/(b_0^{(m)} + b_1^{(m)}] + t_c$, which is independent of x.

The question whether a compensatory mastery rule is fair is touched only briefly here. On the one hand, one may argue that if two examinees have the same observed score on the mastery test, the same decision should be made for either of them. On the other hand, as argued earlier, in the realistic case of a mastery test not perfectly reliable, examinees with the same observed score on the mastery test but different observed scores on the selection test have different expected *true scores* on the mastery test, and it seems unfair to the examinee with the higher expected true score not to take this information into account. A comparable issue can be raised with respect to the use of group-based statistical information in determining (noncompensatory) cutting scores in a separate treatment of the selection and mastery decision problems. For a discussion of this problem, see de Gruijter and Hambleton (1984a, 1984b) and van der Linden (1984). The underlying issue is whether an institutional viewpoint is allowed or a more individual viewpoint should be taken. Internationally, legal systems show different jurisprudence related to this issue.

Although the area of individualized instruction is a useful application of simultaneous decision making, it should be emphasized that the optimization models advocated in this paper have a larger scope of application. For any situation in which subjects are accepted for a certain treatment on the basis of their scores on a selection test with attainments evaluated by a mastery test, the optimal rules presented in this paper can improve the decisions. An example is psychotherapy where clients accepted have to pass a success criterion before being dismissed from the therapy.

References

Brouwer, U., & Vijn, P. (1979). Bayesiaanse schatter voor de correlatiecoëfficiënt in "restriction of range" (geval 1) [Bayesian estimators for the correlation coefficient with restriction of range (Case 1)]. *Tijdschrift voor Onderwijsresearch, 4*, 281–290.

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., & Gleser, G. C. (1965). *Psychological tests and personnel decisions* (2nd ed.). Urbana, IL: University of Illinois Press.

Davis, C. E., Hickman, J., & Novick, M. R. (1973). *A primer on decision analysis for individually prescribed instruction* (ACT Technical Bulletin No. 17). Iowa City, IA: American Testing College Program.

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

de Gruijter, D. N. M., & Hambleton, R. K. (1984b). Reply to van der Linden's "Thoughts on the use of decision theory to set cutoff scores". *Applied Psychological Measurement, 8*, 19–20.

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.

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

Huynh, H. (1977). Two simple classes of mastery scores based on the beta-binomial model. *Psychometrika, 42*, 601–608.

Huynh, H. (1980). A nonrandomized minimax solution for passing scores in the binomial error model. *Psychometrika, 45*, 167–182.

Huynh, H. (1982). A Bayesian procedure for mastery decisions based on multivariate normal test data. *Psychometrika, 47*, 309–319.

Huynh, H., & Perney, J. C. (1979). Determination of mastery scores when instructional units are hierarchically related. *Educational and Psychological Measurement, 39*, 317–323.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs.* New York: Wiley.

Kendall, M., & Stuart, M. A. (1979). *The advanced theory of statistics. Volume 2: Inference and relationship.* London: Charles Griffin & Co.

Lehmann, E. L. (1986). *Testing statistical hypotheses* (2nd ed.). New York: John Wiley and Sons.

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

Lord, F. M., & Novick, M. R. (1986). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

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

Novick, M. R., & Lindley, D. V. (1979). Fixed-state assessment of utility functions. *Journal of the American Statistical Association, 74*, 306–311.

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

Roe, R. A. (1979). The correction for restriction of range and the difference between intended and actual selection. *Educational and Psychological Measurement, 39*, 551–560.

Sawyer, R. (1993, April). *Decision theoretic models for validating course placement systems.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Taylor, H. C., & Russell, J. T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection discussion and tables. *Journal of Applied Psychology, 23*, 565–578.

van der Gaag, N. (1990). *Empirische utiliteiten voor psychometrische beslissingen* (Empirical utilities for psychometric decisions). Unpublished doctoral dissertation, University of Amsterdam, The Netherlands.

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. (1984). Some thoughts on the use of decision theory to set cutoff scores: Comment on de Gruijter and Hambleton. *Applied Psychological Measurement, 8*, 9–18.

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. (1995). Some decision theory for course placement. Manuscript submitted for publication.

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

Vos, H. J. (1990). Simultaneous optimization of decisions using a linear utility function. *Journal of Educational Statistics*, *15*, 309–340.

Vos, H. J. (1991). Simultaneous optimization of the aptitude treatment interaction decision problem with mastery scores. In M. Wilson (Ed.), *Objective measurement: Theory into practice* (Vol. 1, pp. 313–331). Norwood, NJ: Ablex.

Vos, H. J. (1993). Simultaneous optimization of selection-placement decisions followed by a mastery decision. In F. J. Maarse et al. (Eds.), *Computers in psychology: Tools for experimental and applied psychology* (Vol. 4, pp. 235–244). Lisse, The Netherlands: Swets & Zeitlinger.

Vos, H. J. (1994). *Simultaneous optimization of test-based decisions in education*. Unpublished doctoral dissertation, University of Twente, The Netherlands.

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.

Manuscript received 4/11/94
Final version received 2/28/95