DOCUMENT RESUME

ED 389 745 TM 024 378

AUTHOR	Vos, Hans J.
TITLE	A Compensatory Model for Simultaneously Setting Cutting Scores for Selection-Placement-Mastery Decisions. Research Report 94-17.
INSTITUTION	Twente Univ., Enschede (Netherlands). Faculty of Educational Science and Technology.
PUB DATE	Nov 94
NOTE	40p.
AVAILABLE FROM	Bibliotheek, Faculty of Educational Science and Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Bayesian Statistics; *Computer Assisted Instruction; Criteria; *Cutting Scores; Decision Making; Evaluation Criteria; Foreign Countries; *Mastery Tests; Medical Students; Scores; *Selection; *Student Placement
IDENTIFIERS	Monotonicity Analysis; Optimization; Simultaneous Interpretation

ABSTRACT

A method is proposed for optimizing cutting scores for a selection-placement-mastery problem simultaneously. A simultaneous approach has two advantages over separate optimization. First, test scores used in previous decisions can be used as "prior data" in later decisions, increasing the efficiency of the decisions. Then, more realistic utility structures can be defined using final success criteria in utility functions for earlier decisions. An important distinction is made between weak and strong decision rules. Weak rules are allowed to be a function of prior test scores. Conditions for optimal rules to be monotone are presented, and it is shown that optimal weak monotone rules are compensatory by nature. Results from an empirical example of instructional decision making illustrate the differences between simultaneous and separate approaches. Subjects were 71 medical students receiving interactive video or computer-aided instruction. (Contains 2 tables, 2 figures, and 17 references.) (SLD)

A Compensatory Model for Simultaneously Setting Cutting Scores for Selection-Placement-Mastery Decisions

Research Report
94-17

Hans J. Vos

faculty of
EDUCATIONAL SCIENCE
AND TECHNOLOGY

University of Twente

Department of
Educational Measurement and Data Analysis

A Compensatory Model for Simultaneously Setting Cutting Scores for Selection-Placement-Mastery Decisions

Hans J. Vos

To appear in Vos, H.J. (1995). A Compensatory model for simultaneously setting cutting scores for selection-placement-mastery decisions. In Proceedings of the Society for Multivariate Analysis in the Behavioral Sciences (SMABS). Varna, Bulgaria, May 31 - June 5, 1994.

A compensatory model for simultaneously setting cutting scores for a selection-placement-mastery decision , Hans J. Vos - Enschede: University of Twente, Faculty of Educational Science and Technology, November, 1994. - 33 pages.

Abstract

A model is proposed for optimizing simultaneously combinations of test-based decisions using Bayesian theory. The decision problem addressed consists of a selection, a placement, and a mastery decision. Combinations of such decisions can be found, for instance, in computerized adaptive instruction networks. Compared with separate optimization, a simultaneous approach has two advantages. First, test scores used in previous decisions can be used as "prior data" in later decisions and the efficiency of the decisions can be increased. Second, more realistic utility structures can be defined using final success criteria in utility functions for earlier decisions. An important distinction is made between weak and strong decision rules. As opposed to strong rules, weak rules are allowed to be a function of prior test scores. Conditions for optimal rules to be monotone are presented. Also, it will be shown that the optimal weak monotone rules are compensatory by nature. Results from an empirical example of instructional decision making will be presented to illustrate the differences between a simultaneous and a separate approach.

Introduction

Decision problems in educational and psychological testing can be classified in many ways. In van der Linden (1990) the following four types of test-based decisions are distinguished: selection, mastery, placement, and classification. Typically, modern instructional systems as individualized study systems (ISS's), mastery learning, and computer-aided instruction (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, 1994a; Vos & van der Linden, 1987).

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. The purpose of this paper is to show that multiple decisions in networks can also be optimized simultaneously using Bayesian theory (e.g., DeGroot, 1970; Ferguson, 1967; Lindgren, 1976). 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. As a result of the more efficient use of all data available in the network of simultaneous decisions, one might expect that the expected utility for a simultaneous approach is larger than for a separate approach. Second, a more

realistic definition of utility or loss functions is possible, since these functions can now be defined on the ultimate success criterion instead of on intermediate criteria measuring the success on individual treatments.

In this paper, a decision network consisting of a selection, a placement (with two treatments), and a mastery decision will be used to make our point. First, the selection-placement-mastery problem will be formalized. Then important distictions will be made between weak and strong as well as monotone and nonmonotone decision rules. Next, it will be indicated under which conditions optimal rules take a monotone form. Finally, results from an empirical example of instructional decision making will be presented to illustrate the differences between a simultaneous and a separate approach.

The Selection-Placement-Mastery Problem

A flowchart of the selection-placement-mastery problem is given in Figure 1. Real-life systems usually have more decision points.

Insert Figure 1 about here

The first decision to be made is a selection decision. A selection decision is made when a test is administered before a treatment takes place, and only students promising satisfactory results on the criterion are accepted for the treatment. For instance, the treatment may be a remedial module in which some prerequisite

knowledge is offered in preparation of the treatments to come later in the instructional program. After the initial treatment a placement decision follows, where students are assigned to one of two possible treatments based on their placement scores. With placement decisions the success of each of the treatments is measured by the same criterion. The paradigm underlying placement decisions is the Aptitude Treatment Interaction (ATI) hypothesis from instructional psychology, which assumes that students may react differentially to instructional treatments, and, therefore, that different treatments may be best for different students. In general, students with high test scores on the placement test will be assigned to treatment 1. Therefore we will sometimes refer to treatment 1 and 0 as the 'higher' and 'lower' treatment, respectively. Finally, on the basis of the mastery test, it is decided whether the student has mastered the subject matter in the treatment sufficiently and may proceed with the next treatment. Students who fail, however, have to relearn the material offered in the treatment and prepare themselves for a new mastery test.

It is important to realize that, although the nature of the decisions shown in Figure 1 is sequential, the decision rules are optimized simultaneously. Also, the data used to optimize the rules is supposed to come from the following statistical experiment: First, students exposed to the same selection test are randomly drawn without reference to the score on the selection test in question and accepted for the initial treatment. Next, the accepted students are randomly assigned to either treatment 0 or 1, after which their performances on the mastery test are measured.

In the following, we shall assume that for a randomly sampled individual the observed-score selection test variable X, the observed-score placement test variable Y, the observed-score mastery test variable Z, and the

classical test theory true score variable T underlying Z, i.e., the criterion common to the treatments j (j = 0,1), are continuous random variables. Furthermore, it will be assumed that the relation between X, Y, Z, and T can be represented by a density function $f_j(x,y,z,t)$. Since the treatments take place between the placement and the mastery test, this relation may entail different density functions for each treatment. Therefore, $f_j(x,y,z,t)$ is indexed by j. However, since both the selection and placement test are administered prior to the treatment j, the density functions q(x), s(y), and h(y|x) of respectively X, Y, and Y given X = x will not be indexed by j.

Treatment-dependent Mastery Rules in a Simultaneous Approach

In the case of a simultaneous approach to optimizing decision networks, an important distinction is made with respect to decision rules on the mastery test for students assigned to treatments 0 and 1. These rules will be different to allow for the "collateral information" present in the fact that examinees have followed different previous treatments before they take the mastery test.

Weak Monotone and Strong Monotone Rules

A decision rule specifies for each possible realization (x,y,z) of the sample space XxYxZ which action has to be taken. Here, we only consider monotone rules; that is, rules using cutting scores. Let x_c, y_c, z_{cj}, and t_c denote the cutting scores on X, Y, Z for students assigned to treatment j, and T, respectively, where t_c is set in advance by the decision-maker (j = 0,1). Clearly, the selection-placement-mastery problem now consists of simultaneously setting cutting scores x_c, y_c, and z_{cj} that, given the value of t_c, maximize the expected utility.

In general, when setting y_c as well as z_{cj}, prior achievement on the selection test can or can not be taken into account. In addition, the observed scores on the placement test may or may not influence the cutting scores to be set on the mastery test. Intuitively, students with selection scores equal to or just above x_c must compensate their relatively low selection scores with higher scores on both the placement and mastery test. Reversely, it seems reasonable that students with placement scores far above y_c should be advanced earlier than students with placement scores equal to or just above y_c.

To distinguish between cases where prior achievement has or has not to be taken into account, those rules will be called weak monotone and strong monotone rules, respectively. For each $x \geq x_c$, the weak cutting score on the placement test is defined as a function $y_c(x)$. Similarly, for each $x \geq x_c$ and y, the weak cutting score on the mastery test under treatment j is given as a function $z_{cj}(x,y)$. For strong monotone rules, however, both strong cutting scores y_c and z_{cj} are set independently of observed test scores on prior tests. Since the most general form of the decision rule is a weak simultaneous rule, this type of rule will be treated first. Later on, strong simultaneous rules will be considered as a special form of weak simultaneous rules.

Each action will be denoted by a_{ijk} (i,j,k = 0,1), where i = 0 or 1 stands for rejecting or accepting a student, j = 0 or 1 stands for assigning an accepted student to treatment 0 or 1, and k = 0 or 1 stands for retaining or advancing an accepted student. Since for a rejected student no further placement and mastery decisions are made, the indices j and k will be dropped for i = 0.

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

$$
\begin{aligned}
\{(x,y,z) : \delta(x,y,z) = a_0\} &= A \times R \times R \\
\{(x,y,z) : \delta(x,y,z) = a_{100}\} &= A^C \times B(x) \times D_0(x,y) \\
\{(x,y,z) : \delta(x,y,z) = a_{101}\} &= A^C \times B(x) \times D_0{}^C(x,y) \qquad (1) \\
\{(x,y,z) : \delta(x,y,z) = a_{110}\} &= A^C \times B^C(x) \times D_1(x,y) \\
\{(x,y,z) : \delta(x,y,z) = a_{111}\} &= A^C \times B^C(x) \times D_1{}^C(x,y),
\end{aligned}
$$

where A, A^C, B(x), $B^C(x)$, $D_j(x,y)$, and $D_j{}^C(x,y)$ stand, respectively, for the sets of x, y, and z values for which a student is rejected or admitted for the initial treatment, an accepted student is assigned to treatment 0 or 1, and an accepted student is retained or advanced under treatment j (j = 0,1). R stands for the set of real numbers. Thus, a weak monotone rule δ can be defined for our example as follows:

$$
\delta(X,Y,Z) = \begin{cases}
a_0, & \text{if } X < x_c,\ y \in R, \text{ and } Z \in R \\
a_{100}, & \text{if } X \geq x_c,\ Y < y_c(x), \text{ and } Z < z_{c0}(x,y) \\
a_{101}, & \text{if } X \geq x_c,\ Y < y_c(x), \text{ and } Z \geq z_{c0}(x,y) \\
a_{110}, & \text{if } X \geq x_c,\ Y \geq y_c(x), \text{ and } Z < z_{c1}(x,y) \\
a_{111}, & \text{if } X \geq x_c,\ Y \geq y_c(x), \text{ and } Z \geq z_{c1}(x,y).
\end{cases} \qquad (2)
$$

Since we confine ourselves to monotone rules in this paper, following common practice in criterion-referenced testing, we are to show that there are no nonmonotone rules with larger expected utility, or, equivalently, that the subclass of monotone rules constitutes an essentially complete class (e.g., Ferguson, 1967; Karlin & Rubin, 1956). Conditions under which the subclass of monotone rules is essentially complete will be given later on. If these conditions are satisfied, a monotone solution is said to exist.

Other Types of Rules for the Combined Decision Problem

The introduction of weak monotone and strong monotone rules implies that optimal rules can be considered which take both a weak monotone and strong monotone form. Since strong monotone rules are special cases of weak monotone rules, this type of rules can only be optimal if along with the conditions under which weak monotone rules are optimal certain additional restrictions are met.

Furthermore, since in educational testing one is accustomed to using strong cutting scores, rules with maximum expected utility in the subclass of strong monotone rules can also be calculated without bothering about monotonicity conditions. To stipulate the difference with the optimal (strong monotone) rules, this type of rules will be termed Strong Monotone Rules with Maximum Expected Utility (SMMEU) rules. It should be emphasized that an optimal rule from the set of all possible simultaneous rules only takes a strong monotone form if the additional (rather strict) restrictions are met.

An Additive Utility Structure for the Combined Decision Problem

Th utility structure of the combined decision problem, $u_{ijk}(t)$, is supposed to be an additive function of the following form:

$$u_{ijk}(t) = w_1 u_i^{(s)}(t) + w_2 u_j^{(p)}(t) + w_3 u_k^{(m)}(t), \qquad (3)$$

where $u_i^{(s)}(t)$, $u_j^{(p)}(t)$, and $u_k^{(m)}(t)$ represent, respectively, the utility functions for the separate selection, placement, and mastery decisions, and w_1, w_2, and w_3 are nonnegative weights. For a rejected student, zero utilities for the separate placement and mastery decisions are assumed. Hence, it follows from (3) that $u_{0jk}(t)$ is equal to $w_1 u_0^{(s)}(t)$ for all j and k.

Since the utility functions $u_i^{(s)}(t)$, $u_j^{(p)}(t)$, and $u_k^{(m)}(t)$ are allowed to assume different forms, Equation 3 offers us a great deal of flexibility to describe utility structures in practical applications. Furthermore, since utility is supposed to be measured on an interval scale, the weights in (3) can be rescaled as follows:

$$u_{ijk}(t) = w_1 u_i^{(s)}(t) = w_2 u_j^{(p)}(t) + (1-w_1-w_2) u_k^{(m)}(t), \tag{4}$$

with $0 \leq w_1, w_2, (1-w_1-w_2) \leq 1$.

As mentioned earlier in the Introduction, one of the main advantages of a simultaneous approach is that more realistic utility structures can be defined. Equations 3-4 demonstrate this fact nicely, since a utility function defined on the ultimate criterion of the decision network is formulated not only for the mastery decision but also both for the selection and placement decision. This choice is in line with the philosophy underlying ISS's.

Expected Utility in a Simultaneous Approach

It is important to realize that the expected utility in a simultaneous approach according to the statistical experiment described earlier is composed of eight terms, whereas according to (1) only five possible actions can be identified. The reason is that action a_0 contributes four terms to the expected utility, namely one for each possible combination of outcomes of the decisions that follow. The placement rules in the calculation of this part of the expected utility are different for students rejected or accepted for the initial treatment to allow for this "collateral information". The mastery rules are not only treatment-dependent but also depend on the information if students are rejected or accepted for the initial treatment.

For the most general form of the decision rules according to (1), the expected utility in a simultaneous approach can be calculated as follows:

$$
\begin{aligned}
&E[U_{sim}(A^C, B^C(x), B_R^C(x), D_0^C(x,y), D_1^C(x), D_{r0}^C(x,y), D_{r1}^C(x,y))] \equiv \\
&\int_A \int_{B_r(x)} \int_{D_{r0}(x,y)} \int_R u_{000}(t) f_0(x,y,z,t)\,dtdzdydx \; + \\
&\int_A \int_{B_r(x)} \int_{D_{r0}^C(x,y)} \int_R u_{001}(t) f_0(x,y,z,t)\,dtdzdydx \; + \\
&\int_A \int_{B_r^C(x)} \int_{D_{r1}(x,y)} \int_R u_{010}(t) f_1(x,y,z,t)\,dtdzdydx \; + \\
&\int_A \int_{B_r^C(x)} \int_{D_{r1}^C(x,y)} \int_R u_{011}(t) f_1(x,y,z,t)\,dtdzdydx \; + \\
&\int_{A^C} \int_{B(x)} \int_{D_0(x,y)} \int_R u_{100}(t) f_0(x,y,z,t)\,dtdzdydx \; + \qquad (5) \\
&\int_{A^C} \int_{B(x)} \int_{D_0^C(x,y)} \int_R u_{101}(t) f_0(x,y,z,t)\,dtdzdydx \; + \\
&\int_{A^C} \int_{B^C(x)} \int_{D_1(x,y)} \int_R u_{110}(t) f_1(x,y,z,t)\,dtdzdydx \; + \\
&\int_{A^C} \int_{B^C(x)} \int_{D_1^C(x,y)} \int_R u_{111}(t) f_1(x,y,z,t)\,dtdzdydx \; ,
\end{aligned}
$$

where $B_r(x)$, $B_r^C(x)$, $D_{rj}(x,y)$, and $D_{rj}^C(x,y)$ represent, respectively, the sets of y and z values for which a rejected student is assigned to treatment 0 or 1 and failed or passed the mastery test under treatment j ($j = 0,1$).

Taking expectations, completing integrals, rearranging terms, and using

$u_{0jk}(t) = w_1 u_0^{(s)}(t)$ for all j and k, (5) can be written in posterior form as

$$\begin{aligned}
&E[U_{sim}(A^C, B^C(x), B_r^C(x), D_0^C(x,y), D_1^C(x,y))] = w_1 E_0[u_0^{(s)}(T)] + \\
&\int_R \int_{B_r^C(x)} w_1 \{E_1[u_0^{(s)}(T)|x,y] - E_0[u_0^{(s)}(T)|x,y]\} h(y|x) q(x)\,dy\,dx + \\
&\int_{A^C} \Big\{ E_0[u_{100}(T) - w_1 u_0^{(s)}(T)|x] + \\
&\int_R \int_{D_0^C(x,y)} E_0[u_{101}(T) - u_{100}(T)|x,y,z] m_0(z|x,y) h(y|x)\,dz\,dy + \\
&\sum_{j=0}^{1} (2j-1) \int_{B^C(x)} \{E_j[u_{1j0}(T)|x,y] + \\
&\int_{D_j^C(x,y)} E_j[u_{1j1}(T) - u_{1j0}(T)|x,y,z] m_j(z|x,y)\,dz\} h(y|x)\,dy \Big\} q(x)\,dx,
\end{aligned} \tag{6}$$

with $m_j(z|x,y)$ being the p.d.f. of Z given X = x and Y = y under treatment j and where $E_j[.]$ indicates that the expectation has been taken over a distribution indexed by j.

Sufficient Conditions for Monotone Solutions

In this section, sufficient conditions for optimal simultaneous rules to be monotone will be derived. First, conditions under which optimal rules take a weak monotone form will be derived. Next, it will be indicated how sufficient conditions for strong monotone solutions can be obtained from the previous case by imposing certain additional restrictions on $f_j(x,y,z,t)$.

Sufficient Conditions for Weak Monotonicity

To derive the conditions under which optimal rules take a weak monotone form for our example, the following well-known theorem (e.g., Ferguson, 1967, p. 201) is needed:

> For every function f(x) with $\int |f(x)| dx < \infty$
> and any set S of x values,
> it holds that $\int_S f(x)dx \leq \int_{S_0} f(x)dx$ with $S_0 = \{x: f(x) \geq 0\}$. (7)

Applying this theorem first to the inside integrals w.r.t. z, next to the middle integrals w.r.t. y, and finally to the outside integral w.r.t. x in (6), it can be verified that an upper bound to the expected utility is obtained for:

$$D^{*C}_0(x,y) \equiv \{z: E_0[u_{101}(T) - u_{100}(T) | x,y,z] \geq 0\}, \tag{8}$$

$$D^{*C}_1(x,y) \equiv \{z: E_1[u_{111}(T) - u_{110}(T) | x,y,z] \geq 0\}, \tag{9}$$

$$B^{*C}(x) \equiv \{y: \sum_{j=0}^{1} (2j-1)\{E_j[u_{1j0}(T) | x,y] + \int_{D^{*C}_j(x,y)} E_j[u_{1j1}(T) - u_{1j0}(T) | x,y,z] m_j(z | x,y) dz\} \geq 0\}, \tag{10}$$

$$B^{*C}_r(x) \equiv \{y: E_1[u_0^{(s)}(T) | x,y] - E_0[u_0^{(s)}(T) | x,y] \geq 0\}, \tag{11}$$

$$\begin{aligned}
A^{*C} \equiv \{x{:}\ & E_0[u_{100}(T) - W_1 u_0^{(s)}(T)\,|\,x] + \\
& \int_R \int_{D^{*}{}_0^{C}} E_0[u_{101}(T) - u_{100}(T)\,|\,x,y,z]\, m_0(z\,|\,x,y)\, h(y\,|\,x)\, dz\, dy + \\
& \sum_{j=0}^{1} (2j-1) \int_{B^{*C}(x)} \{E_j[u_{1j0}(T)\,|\,x,y] + \\
& \int_{D^{*}{}_j^{C}(x,y)} E_j[u_{1j1}(T) - U_{1j0}(T)\,|\,x,y,z]\, m_j(z\,|\,x,y)\, dz\}\, h(y\,|\,x)\, dy \geq 0\}.
\end{aligned} \qquad (12)$$

We are now able to specify conditions for weak monotonicity. For weak monotone rules, the sets $D^{*}{}_j^{C}(x,y)$, $B^{*C}(x)$, and A^{*C} take the form $[z_{cj}(x,y),\infty)$, $[y_c(x),\infty)$, and $[x_c,\infty)$, respectively. As opposed to these sets, however, it is assumed that the sets $B^{*}{}_r^{C}(x)$ take the form $(-\infty,y_{cr}(x)]$. The reason for assuming a somewhat different form for the sets $B^{*}{}_r^{C}(x)$ is that under quite realistic conditions on the test score distributions and utility functions (Vos, 1994b), a weak monotone solution can be guaranteed. The form of $B^{*}{}_r^{C}(x)$ stated above implies that rejected students, unlike accepted students, with low scores on the placement test will generally be assigned to the 'higher' treatment.

It follows that the optimal weak rules take a monotone form if the left-hand sides of the inequalities in (8)-(9), (10), and (12) are increasing functions in z for all x and y, in y for all x and z, and in x for all y and z, respectively, whereas the left-hand side of the inequality in (11) is required to be decreasing in y for all x. In the empirical example below it will be examined if these conditions hold.

The sets $D^{*}{}_0^{C}(x,y)$, $D^{*}{}_1^{C}(x,y)$, $B^{*}{}_r^{C}(x)$, $B^{*C}(x)$, and A^{*C} can be obtained in the following way: For each x and y, first the sets $D^{*}{}_0^{C}(x,y)$ and $D^{*}{}_1^{C}(x,y)$ can be computed from (8) and (9), respectively, whereas for each x

the sets $B^{*}{}_{r}^{C}(x)$ can be computed from (11). Then, for each x, inserting $D^{*}{}_{j}^{C}(x,y)$ into (10), the sets $B^{*C}(x)$ can be computed. Finally, inserting $D^{*}{}_{j}^{C}(x,y)$ and $B^{*C}(x)$ into (12), the set A^{*C} can be computed.

Monotonicity Conditions for Strong Simultaneous Rules

For strong simultaneous rules, $B^{C}(x)$ and $B_{r}^{C}(x)$ are not allowed to depend on x and $D_{j}^{C}(x,y)$ not on x and y. Let $v_j(t \mid x,y,z)$ and $r_j(t \mid x,y)$ denote the p.d.f.'s of T given X = x, Y = y, and Z = z and T given X = x and Y = y under treatment j, respectively (j = 0,1). In addition to the monotonicity conditions for weak simultaneous rules, it then follows from (8)-(11) that an upper bound to the expected utility is reached for a strong monotone rule if $v_j(t \mid x,y,z)$ does not depend on X = x and Y = y, whereas $r_j(t \mid x,y)$ and $m_j(z \mid x,y)$ do not depend on X = x (j = 0,1).

Calculation of Simultaneous Rules

In this section, it will be indicated how the different types of rules discussed earlier can be calculated. First, it will be shown how optimal rules can be obtained being both of a weak monotone and strong monotone form. Next, it will be indicated how SMMEU rules can be derived by maximizing the expected utility over the subclass of strong monotone rules.

Optimal Simultaneous Rules

Assuming the conditions for weak monotonicity are satisfied, optimal weak cutting scores can now be obtained for those values of $z_{cj}(x,y)$, $y_c(x)$, $y_{cr}(x)$, and

x_c for which the inequalities in (8)-(12) turn into equalities. The optimal weak monotone rule is then given by x_c, $\{y_c(x)|x \geq x_c\}$, $\{y_{cr}(x)|x < x_c\}$, $\{z_{c0}(x,y)|x \geq x_c,\ y < y_c(x)\}$, and $\{z_{c1}(x,y)|x \geq x_c,\ y \geq y_c(x)\}$. Assuming the additional restrictions on $f_j(x,y,z,t)$ are also satisfied, optimal strong cutting scores can be obtained by solving the resulting system of equations simultaneously for x_c, y_c, y_{cr}, z_{c0}, and z_{c1}.

SMMEU Rules

The set of SMMEU cutting scores, say x^*_c, y^*_c, y^*_{cr}, z^*_{c0}, and z^*_{c1}, can be computed straightforward by inserting $A^C = [x_c,\infty)$, $B^C(x) = [y_c,\infty)$, $B_r^C(x) = [y_{cr},\infty)$, and $D_j^C(x,y) = [z_{cj},\infty)$ into (6), differentiating w.r.t. x_c, y_c, y_{cr}, z_{c0}, and z_{c1}, setting the resulting expressions equal to zero, and solving simultaneously for x_c, y_c, y_{cr}, z_{c0}, and z_{c1}.

Since no analytical solutions for these systems of equations could be found, the cutting scores can be calculated via a numerical approximation procedure such as Newton's iterative algorithm for solving nonlinear equations. For the present systems of equations, this algorithm 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 author of the paper upon request.

Optimal Separate Rules

It is observed that optimal rules for the separate decisions can easily be found by imposing certain restrictions on the expected utility in a simultaneous approach.

First, for the separate selection decision it holds that $m_0(z|x,y) = m_1(z|x,y) = m(z|x,y)$ and $w_2 = (1-w_1-w_2) = 0$, since both treatments coincide and there are zero placement and mastery utilities in this case, respectively. Substituting these restrictions into (6), the expected utility for the separate selection decision, $E[U_{sep}^{(s)}(A^C)]$, becomes

$$E[U_{sep}^{(s)}(A^C)] = E[u_0^{(s)}(T)] + \int_{A^C} E[u_1^{(s)}(T) - u_0^{(s)}(T)|x]q(x)dx. \qquad (13)$$

Next, for the separate placement decision it holds that $A^C = R$ and $w_1 = (1-w_1-w_2) = 0$, since all students are accepted for the initial treatment and there are zero selection and mastery utilities in this case, respectively. Substitution of these restrictions as well as $B^C(x) = B^C$ into (6) results for the expected utility of the separate placement decision, $E[U_{sep}^{(p)}(B^C)]$, in

$$E[U_{sep}^{(p)}(B^C)] =$$

$$E_0[u_0^{(p)}(T)] + \int_{B^C} \{E_1[u_1^{(p)}(T)|y] - E_0[u_0^{(p)}(T)|y]\}s(y)dy. \qquad (14)$$

Finally, for the separate mastery decision it holds that $A^C = R$, $m_0(z|x,y) = m_1(z|x,y) = m(z|x,y)$, and $w_1 = w_2 = 0$, since all students are accepted for the initital treatment, both treatments coincide, and there are zero selection and placement utilities in this case, respectively. Inserting these restrictions as well as $D_j^C(x,y) = D^C$ into (6), the expected utility for the separate mastery decision, $E[U_{sep}^{(m)}(D^C)]$, becomes

$$E[U_{sep}^{(m)}(D^C)] = E[u_0^{(m)}(T)] + \int_{D^C} E[u_1^{(m)}(T) - u_0^{(m)}(T)|z]p(z)dz, \qquad (15)$$

where p(z) denotes the marginal distribution of Z.

Analogous to a simultaneous approach, it can easily be verified that upper bounds to $E[U^{(s)}_{sep}(A^C)]$, $E[U^{(p)}_{sep}(B^C)]$, and $E[U^{(m)}_{sep}(D^C)]$ are obtained for the sets of x, y and z values for which the integrands in (13), (14), and (15) are nonnegative, respectively. For monotone rules, these sets take the form $[x_C,\infty)$, $[y_C,\infty)$, and $[z_C,\infty)$, respectively. Assuming the monotonicity conditions for the separate decisions are satisfied, it follows that optimal cutting scores for the separate selection, placement, and mastery decisions, say $\tilde{x}_C$, $\tilde{y}_C$, and $\tilde{z}_C$, can be obtained by solving the integrands in (13), (14), and (15) for x_C, y_C, and z_C, respectively. For further details, see Mellenbergh and van der Linden (1981), van der Linden (1981), and van der Linden and Mellenbergh (1977).

An Empirical Example

The procedures for computing the optimal rules were applied to a sample of 71 freshmen studying medicine. Treatments 0 and 1 consisted of an interactive video (IV) and a computer-aided instructional (CAI) program. Ordering the treatments in this way was motivated by the fact that the IV-program contained more examples and exercises than the CAI-program, implying that students with high scores on the placement test were generally assigned to treatment 1.

The selection, placement, and mastery tests consisted each of 25 free-response items with test scores ranging from 0 to 100. It should be emphasized that the example given in this paper was used only to illustrate the models. The use of small samples is generally not recommended, because the estimated model parameters may have large errors which tend to propagate when

computing optimal rules.

The teachers of the course considered a student as having mastered the subject matter if he/she could answer correctly at least 60% of the total domain of items. Therefore, t_c was fixed at 60.

Multivariate Normal Distribution

It was assumed that the variables X, Y, Z, and T followed (possibly) different multivariate normal distributions under both treatments. Under this assumption, the means, variances, correlations, and reliability coefficients of the mastery test scores were estimated using maximum likelihood estimates and coefficient alpha, respectively. Table 1 shows the results of the computations.

Insert Table 1 about here

When applying the procedures in this paper, it should always be checked whether the assumed multivariate normal distribution for $f_j(x,y,z,t)$ holds. This assumption was tested by examining whether a trivariate normal distribution for (X,Y,Z) under both treatments as well as the linearity of the regression functions $E_j(T|x)$, $E_j(T|y)$, and $E_j(T|x,y)$ did hold against the data. The trivariate normal distribution for (X,Y,Z) under treatment j was tested using a Chi-square test by partitioning the sample space into 20 intervals of (x,y,z) observations (df = 20-9-1). Furthermore, the null hypotheses of "no linear relation" for the regression functions $E_j(Z|x) = E_j(T|x)$ and $E_j(Z|y) = E_j(T|y)$ were tested for a usual t-test (df = n_j-2), with n_j denoting the number of students in the sample

assigned to treatment j. Finally, the null hypotheses of "no linear relation" for the regression functions $E_j(Z|x,y) = E_j(T|x,y)$ were tested using the standard F-test ($df = [m, n_j - m - 1]$), with $m = 2$ denoting the number of explanatory variables. All p-values showed a satisfactory fit ($\alpha = 0.05$).

Utility Functions for the Separate Decisions

For the separate selection and placement decisions, it will be assumed that the utility functions can be represented as linear functions of the criterion variable T (Mellenberg & van der Linden, 1981; van der Linden & Mellenbergh, 1977):

$$u_i^{(s)}(t) = \begin{cases} b_0^{(s)}(t_c - t) + d_0^{(s)} & \text{for } i = 0 \\ b_1^{(s)}(t - t_c) + d_1^{(s)} & \text{for } i = 1, \end{cases} \qquad (16)$$

$$u_j^{(p)}(t) = \begin{cases} b_0^{(p)}(t_c - t) + d_0^{(p)} & \text{for } j = 0 \\ b_1^{(p)}(t - t_c) + d_1^{(p)} & \text{for } j = 1, \end{cases} \qquad (17)$$

where $b_i^{(s)}$, $b_j^{(p)} > 0$ and $[b_1^{(p)} - b_0^{(p)}] > 0$ ($i,j = 0,1$). The condition $b_1^{(s)}$, $b_0^{(s)} > 0$ states that utility must be an increasing function of the criterion for the acceptance decision, but a decreasing function for the rejection decision. Furthermore, the condition $b_0^{(p)}$, $b_1^{(p)} > 0$ implies that both for assigning students to treatment 0 and 1, utility is an increasing function of t. Finally, the condition $[b_1^{(p)} - b_0^{(p)}] > 0$ implies that $u_0^{(p)}(t)$ must be a more slowly increasing function in t than $u_1^{(p)}(t)$. This condition is needed to guarantee a monotone solution for the separate placement decision (cf. van der Linden, 1981).

For the separate mastery decision the well-known threshold utility function is assumed (e.g., Hambleton & Novick, 1973; Huynh, 1976; Novick & Lindley, 1978), which is defined by the following four constants:

$$u_k^{(m)}(t) = \begin{cases} d_{00} & \text{for } t < t_c \text{ and } k = 0 \\ d_{01} & \text{for } t < t_c \text{ and } k = 1 \\ d_{10} & \text{for } t \geq t_c \text{ and } k = 0 \\ d_{11} & \text{for } t \geq t_c \text{ and } k = 1, \end{cases} \qquad (18)$$

where $d_{10} < d_{00}$ and $d_{01} < d_{11}$. The condtions $d_{10} < d_{00}$ and $d_{01} < d_{11}$ express the assumptions that incorrect decisions represent a smaller utility than correct decisions.

Monotonicity Conditions

It should always be checked whether the conditions for weak monotonicity are satisfied. Doing so, it turned out that the left-hand sides of the inequalities in (8)-(9), (10), (11), and (12) were increasing in z for all x and y, increasing in y for all x and z, decreasing in y for all x, and increasing in x for all y and z, respectively, with some minor exceptions at the lower ends of the range of test scores ($0 \leq x,y,z \leq 100$).

Finally, it remained still to be tested if the additional conditions on $f_j(x,y,z,t)$ for strong monotone solutions were met. First, the conditions of $r_j(t|x,y)$ and $m_j(z|x,y)$ being independent of $X = x$ were tested comparing the linear regression functions $E_j(T|x,y)$ and $E_j(T|y)$. The null hypothesis "the variable X does not deliver a significant contribution to the explanation of T" was tested with an F-test (df = $[1,n_j-3]$). Second, the condition of $v_j(t|x,y,z)$ being

independent of $X = x$ and $Y = y$ ws tested comparing the linear regression functions $E_j(T \mid x,y,z)$ and $E_j(T \mid z)$. The null hypothesis "the variables X and Y do not deliver a significant contribution to the explanation of T" was tested using an F-test with $df = [2,n_j-4]$. All p-values did not show a satisfactory fit to the test data, however, implying that the optimal rules did not take a strong monotone form ($\alpha = 0.05$). Therefore, only SMMEU rules and no optimal strong rules were considered.

An absolute maximum appeared to exist for the expected utility in the subclass of strong monotone rules, because the Hessian matrix was negative definite for all nonnegative test scores.

Results for both Simultaneous and Separate Rules

To illustrate the dependence of the results on the chosen utility structures, the SMMEU and the set of weak cutting scores $(x_c, y_c(x_c), z_{cj}(x_c,y_c(x_c)))$ were computed for three different values of the utility parameters as well as for $w_1 = 0.6$ and $w_2 = 0.2$, $w_1 = 0.1$ and $w_2 = 0.8$, and $w_1 = w_2 = 0.333$ using the program NEWTON. The results are displayed in Table 2. The optimal cutting scores $\tilde{x}_c$, $\tilde{y}_c$, and $\tilde{z}_c$ for the separate selection, placement, and mastery decisions are also reported in Table 2. It should be noted that the weak and SMMEU cutting scores $y_{cr}(x_c)$ and y^*_{cr} stayed nearly constant at approximately 39.48 and 39.89, respectively. These cutting scores are not displayed in Table 2, however, because they were only used to compute the optimal rules but are not used for taking decisions in an existing ISS.

Insert Table 2 about here

Since $(\partial/\partial x)z_{cj}(x,y) = -\beta_{TXj.YZ}/\beta_{TZj.XY}$, it followed that $z_{c0}(x,y)$, and $z_{c1}(x,y)$ decreased by 1.85 and 1.05 per unit increase in x for $y < y_c(x)$ and $y \geq y_c(x)$, respectively. Similarly, $(\partial/\partial y)z_{cj}(x,y) = -\beta_{TYj.XZ}/\beta_{TZj.XY}$ implied that $z_{c0}(x,y)$ and $z_{c1}(x,y)$ decreased by 2.05 and 1.32 per unit increase in y for $y < y_c(x)$ and $y \geq y_c(x)$, respectively. The behavior of $y_c(x)$ was examined using the program NEWTON. The results of the computations indicated that $y_c(x)$ was decreasing in x for $x \geq x_c$. Hence, our expectations of the functions $y_c(x)$ and $z_{cj}(x,y)$ being decreasing in x and in both x and y for $x \geq x_c$ could be confirmed ($j = 0,1$). Furthermore, to illustrate the combined effect of both x and $y_c(x)$ on $z_{cj}(x,y_c(x))$, the graphical displays of $y_c(x)$, $z_{c0}(x,y_c(x))$, and $z_{c1}(x,y_c(x))$ for $x \geq x_c$ are shown in Figure 2 for utility structure (2).

Insert Figure 2 about here

As Figure 2 shows, $y_c(x)$, $z_{c0}(x,y_c(x))$, and $z_{c1}(x,y_c(x))$ decreased with approximately 0.5, 0.8, and 0.36 per unit increase in x for $x \geq x_c$. Apparently, the decreasing character of $z_{cj}(x,y)$ in x does have a stronger influence on $z_{cj}(x,y_c(x))$ than the combined effect of the decreasing character of both $z_{cj}(x,y)$ in y and $y_c(x)$ in x for utility structure (2). For all other utility structures, the same pattern could be observed in this study.

Inspection of Table 2 shows that $z_{c0}(x_c,y_c(x_c))$ is larger than $z_{c1}(x_c,y_c(x_c))$. Combined with the fact that both $z_{c0}(x,y)$ (when $y < y_c(x)$) and $z_{c1}(x,y)$ (when $y \geq y_c(x)$) are decreasing functions in y, it follows that students just accepted for the initial treatment and assigned to treatment 0 are always confronted with higher weak cutting scores on the mastery test than students just accepted for the initial treatment and assigned to treatment 1.

As can be seen from Table 2, $\tilde{x}_c$ did not show large differences compared to x_c. However, in particular for $w_1 = 0.6$ and $w_2 = 0.2$, $\tilde{y}_c$ and $\tilde{z}_c$ were substantially higher and lower compared to $y_c(x_c)$ and $z_{cj}(x_c,y_c(x_c))$, respectively ($j = 0,1$).

Obviously, students with selection scores $X = x_c$ were sooner assigned to the 'higher' treatment in the case of a weak monotone approach but had to compensate their relatively low weak cutting scores on the placement test with higher optimal weak cutting scores on the mastery test. However, the decreasing character of $z_{cj}(x,y_c(x))$ in x for $x \geq x_c$ implies that for students with selection scores far above x_c these rather high weak cutting scores on the mastery test under treatment j decreased. Also, students assigned to the 'higher' treatment with selection and placement scores far above x_c and $y_c(x)$ needed only low scores on the mastery test.

Comparison of the Expected Utilities

As earlier noted, one might expect that in a case with empirical data the expected utility for a simultaneous approach will be larger than for a separate approach. This expectation will now be examined comparing the expected utilities for the optimal weak monotone and SMMEU rules with the weighted sum of the expected utilities for the optimal separate rules using the program UTILITY. The

results are also reported in Table 2.

As can be seen from Table 2, the expected utilities for both the optimal weak monotone and SMMEU rules yielded larger values than for the weighted sum of the expected utilities for the optimal separate rules for all nine utility structures. Some utility structures, such as, for instance, utility structures (3), (6), and (9), even showed substantial gains in expected utility for a simultaneous approach.

Furthermore, inspection of Table 2 shows that for all three approaches, the expected utility yielded the largest value for $w_2 = 0.8$ in this study. In other words, the utility for the placement decision contributed most to the expected utility in our example.

Also, comparing utility structures (1)-(3) with (4)-(6), it can be concluded from Table 2 that raising the utilities for the correct mastery decisions resulted in an increase of the expected utilities for all three approaches. Obviously, correct mastery decisions have a relatively strong positive influence on the specification of the utility structure.

Finally, the expected utilities for the optimal weak monotone and SMMEU rules were compared to each other. As Table 2 shows, the expected utilities for the optimal weak monotone rules were larger than for the SMMEU rules. This result does not contradict our predictions, because the expected utility for an optimal weak monotone rule must yield the largest expected utility of all simultaneous rules if the conditions for weak monotonicity are met.

Discussion

In this paper, cutting scores for a selection-placement-mastery problem were optimized simultaneously using Bayesian decision theory. The optimal decision procedures were illustrated empirically using data from the area of instructional decision making. It turned out that in some cases considerable gains in expected utility could be achieved by the optimal weak monotone rules compared to the weighted sum of the expected utilities for the optimal separate rules.

The results indicated that the optimal weak monotone rules $y_c(x)$ and $z_{cj}(x,y)$ were decreasing in x and both in x and y. In Vos (1994b) it is shown that under the same rather mild conditions on the test score distributions and utility functions which guarantee a weak monotone solution, optimal weak monotone rules are always compensatory by nature. The title of the paper already alludes to this result. As already explained, this feature introduces an element of compensation in the decision procedure: It is possible, for instance, to compensate low scores on the placement test by high scores on the selection test.

A final note is appropriate. Although instructional decision making is a useful application of simultaneous decision making, the models advocated in this paper, however, are not limited to this area. Other useful applications may be found in such areas as psychotherapy in which it can be expected that accepted patients for the program react differentially to a certain kind of therapy followed by a success criterion, which has to be passed before being dismissed from the therapy.

References

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

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

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 consideration of mastery scores. Psychometrika, 41, 65-79.

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

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

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.

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

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.

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 (Volume 1, pp. 313-331). Norwood, New Yersey: Ablex Publishing Corporation.

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 (Volume 4, pp. 235-244). Lisse: Swets & Zeitlinger.

Vos, H.J. (1994a). An intelligent tutoring system for classifying students into instructional treatments with mastery scores. In F.J. Maarse et al. (Eds.), Computers in psychology: Applications, methods, and instrumentation (Volume 5, pp. 8-18). Lisse, The Netherlands: Swets & Zeitlinger.

Vos, H.J. (1994b). Simultaneous optimization of test-based decisions in education. Ph.D.thesis, University of Twente, The Netherlands.

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.

Acknowledgment

The author is indebted to Jan Gulmans for providing the data for the empirical example.

Table 1

Statistics for the Selection, Placement, and Mastery Tests (X, Y, and Z)

	X	Y		Z Treatment	
				0	1
Mean	57.324	61.551	64.931	64.241	65.621
Standard Deviation	10.971	11.261	10.759	10.246	11.208
Reliability			0.802	0.798	0.806
Correlation	$\hat{\rho}_{XY}=0.782$	$\hat{\rho}_{XZ}=0.801$	$\hat{\rho}_{XZ0}=0.813$ $\hat{\rho}_{XZ1}=0.794$	$\hat{\rho}_{YZ0}=0.826$	$\hat{\rho}_{YZ1}=0.819$

BEST COPY AVAILABLE

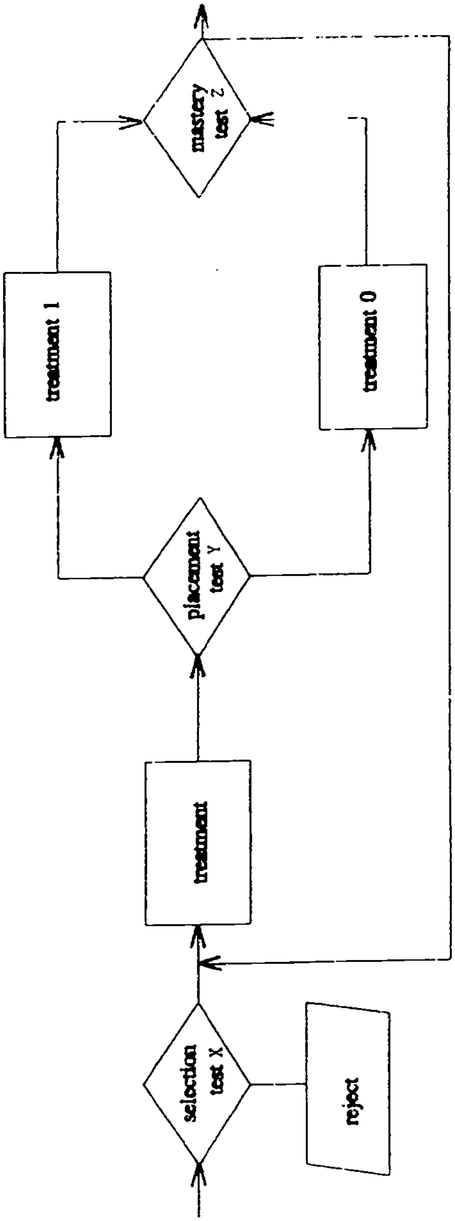
[illegible]

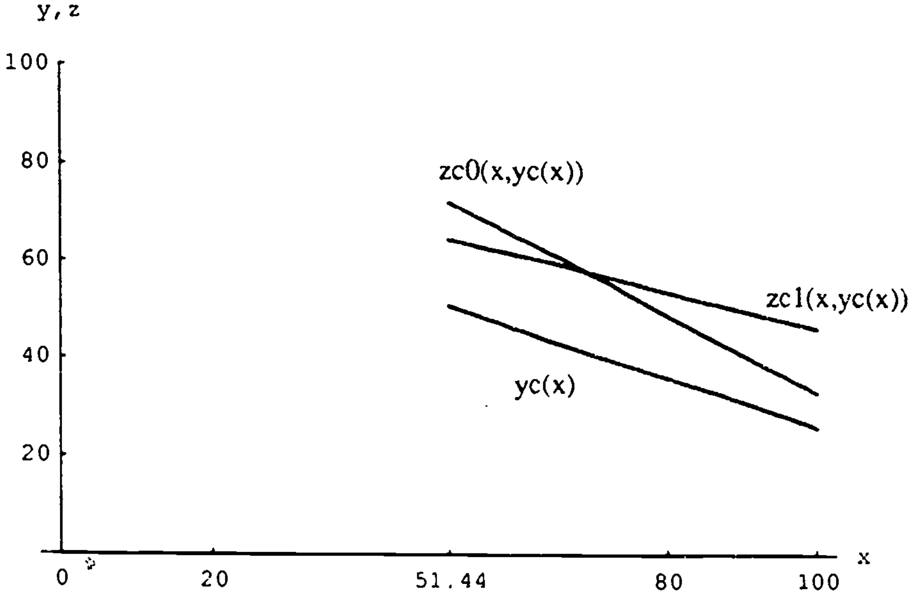
[illegible]	Simultaneous Weak	Simultaneous CMMEU	Separate	Simultaneous Weak	Simultaneous CMMEU	Separate
[illegible]	[illegible] [illegible] [illegible] [illegible]	[illegible] y*=46.80 [illegible] [illegible]	X=51.[illegible] Y=52.[illegible] [illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible] [illegible] [illegible] [illegible]	[illegible] y*=48.[illegible] [illegible] z*=5[illegible].45	X=51.[illegible] Y=5[illegible] [illegible]	11.[illegible]	14.[illegible]	[illegible]
[illegible]	[illegible] [illegible] [illegible] [illegible]	x*=51.5[illegible] y*=49.4[illegible] z*=[illegible].32 z*=5[illegible].32	X=51.[illegible] Y=5[illegible].84 Z=5[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible] [illegible] [illegible] [illegible]	x*=50.[illegible] y*=4[illegible].41 z*=[illegible].82 z*=51.[illegible]	X=51.[illegible]6 Y=5[illegible].54 Z=5[illegible].10	1[illegible]	15.[illegible]	14.[illegible]
[illegible]	[illegible] [illegible] [illegible].62 [illegible]	x*=50.56 y*=49.06 z*=65.48 z*=50.50	X=51.28 Y=52.84 Z=56.10	15.44	15.[illegible]	13.14
[illegible]	X=51.[illegible] [illegible]=52.[illegible] [illegible] [illegible]	x*=52.42 y*=49.45 z*=64.39 z*=49.84	X=51.28 Y=52.84 Z=56.10	21.96	[illegible].89	1[illegible]
[illegible]	X=51.17 [illegible]=49.61 [illegible]=74.06 [illegible]=65.74	x*=50.59 y*=47.94 z*=70.03 z*=53.35	X=51.26 Y=53.[illegible] Z=5[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible] [illegible] [illegible] [illegible]	x*=50.5[illegible] y*=49.34 z*=68.02 [illegible]	X=51.[illegible] Y=53.[illegible] Z=5[illegible].8[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible] [illegible] [illegible] [illegible]	[illegible] [illegible].85 [illegible].23 [illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Figure Caption

Figure 1. A system of a selection, a placement, and a mastery decision (Case of two treatments).

Figure 2. Graphical displays of $y_c(x)$, $z_{c0}(x,y_c(x))$, and $z_{c1}(x,y_c(x))$ for $x \geq x_c$ in utility structure (2).


selection test X
reject
treatment
placement test Y
treatment 1
treatment 0
mastery test Z


y, z
100
80
60
40
20
zc0(x,yc(x))
zc1(x,yc(x))
yc(x)
0
20
51.44
80
100
x

Titles of recent Research Reports from the Department of
Educational Measurement and Data Analysis.
University of Twente, Enschede,
The Netherlands.

RR-94-17 H.J. Vos, *A compensatory model for simultaneously setting cutting scores for selection-placement-mastery decisions*
RR-94-16 H.J. Vos, *Applications of Bayesian decision theory to intelligent tutoring systems*
RR-94-15 H.J. Vos, *An intelligent tutoring system for classifying students into Instructional treatments with mastery scores*
RR-94-14 R.J.H. Engelen, W.J. van der Linden & S.J. Oosterloo, *Fisher's information in the Rasch model*
RR-94-13 W.J.J. Veerkamp & M.P.F. Berger, *A note on the attenuation paradox in item response theory*
RR-94-12 R.R. Meijer, *Nonparametric and group-based person-fit statistics: A validity study and an empirical example*
RR-94-11 M.P.F. Berger, *Optimal test designs for polytomously scored items*
RR-94-10 W.J. van der Linden & M.A. Zwarts, *Robustness of judgments in evaluation research*
RR-94-9 L.M.W. Akkermans, *Monte Carlo estimation of the conditional Rasch model*
RR-94-8 R.R. Meijer & K. Sijtsma, *Detection of aberrant item score patterns: A review of recent developments*
RR-94-7 W.J. van der Linden & R.M. Luecht, *An optimization model for test assembly to match observed-score distributions*
RR-94-6 W.J.J. Veerkamp & M.P.F. Berger, *Some new item selection criteria for adaptive testing*
RR-94-5 R.R. Meijer, K. Sijtsma & I.W. Molenaar, *Reliability estimation for single dichotomous items*
RR-94-4 M.P.F. Berger & W.J.J. Veerkamp, *A review of selection methods for optimal design*
RR-94-3 W.J. van der Linden, *A conceptual analysis of standard setting in large-scale assessments*
RR-94-2 W.J. van der Linden & H.J. Vos, *A compensatory approach to optimal selection with mastery scores*
RR-94-1 R.R. Meijer, *The influence of the presence of deviant item score patterns on the power of a person-fit statistic*
RR-93-1 P. Westers & H. Kelderman, *Generalizations of the Solution-Error Response-Error Model*
RR-91-1 H. Kelderman, *Computing Maximum Likelihood Estimates of Loglinear Models from Marginal Sums with Special Attention to Loglinear Item Response Theory*

ERRATUM

Title Research Report 94-13:

RR-94-13 W.J.J. Veerkamp & M.P.F. Berger, *A simple and fast item selection procedure for adaptive testing*

RR-90-8	M.P.F. Berger & D.L. Knol, *On the Assessment of Dimensionality in Multidimensional Item Response Theory Models*
RR-90-7	E. Boekkooi-Timminga, *A Method for Designing IRT-based Item Banks*
RR-90-6	J.J. Adema, *The Construction of Weakly Parallel Tests by Mathematical Programming*
RR-90-5	J.J. Adema, *A Revised Simplex Method for Test Construction Problems*
RR-90-4	J.J. Adema, *Methods and Models for the Construction of Weakly Parallel Tests*
RR-90-2	H. Tobi, *Item Response Theory at subject- and group-level*
RR-90-1	P. Westers & H. Kelderman, *Differential item functioning in multiple choice items*

Research Reports can be obtained at costs from Bibliotheek, Faculty of Educational Science and Technology, University of Twente, P.O. Box 217, 7500 AE Enschede,
The Netherlands.

faculty of
EDUCATIONAL SCIENCE
AND TECHNOLOGY

A publication by
The Faculty of Educational Science and Technology of the University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands