DOCUMENT RESUME

ED 419 009 TM 028 306

AUTHOR	Vos, Hans J.
TITLE	A Minimax Sequential Procedure in the Context of Computerized Adaptive Mastery Testing.
INSTITUTION	Twente Univ., Enschede (Netherlands). Faculty of Educational Science and Technology.
REPORT NO	RR-97-07
PUB DATE	1997-00-00
NOTE	37p.
AVAILABLE FROM	Faculty of Educational Science and Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Adaptive Testing; Classification; *Computer Assisted Testing; Concept Formation; *Cutting Scores; Foreign Countries; Higher Education; *Mastery Tests; Medical Education; *Test Length
IDENTIFIERS	Decision Theory; *Minimax Procedure

ABSTRACT

The purpose of this paper is to derive optimal rules for variable-length mastery tests in case three mastery classification decisions (nonmastery, partial mastery, and mastery) are distinguished. In a variable-length or adaptive mastery test, the decision is to classify a subject as a master, a partial master, a nonmaster, or continuing sampling and administering another test item. The framework of minimax sequential decision theory is used; that is, optimal sequential rules minimizing the maximum expected losses associated with all possible decision rules. The binomial model is assumed for the conditional probability of a correct response given the true level of functioning, whereas the threshold loss is adopted for the loss function involved. Monotonicity conditions are derived, that is, conditions sufficient for optimal sequential rules to be in the form of cutting scores. The paper concludes with an empirical example of a computerized adaptive mastery test for concept-learning in medicine. (Contains 4 tables and 63 references.) (Author)

TM

A Minimax Sequential Procedure in the Context of Computerized Adaptive Mastery Testing

Research Report 97-07

Hans J. Vos

TM028306

A Minimax Sequential Procedure in the Context of
Computerized Adaptive Mastery Testing

Hans J. Vos

Abstract

The purpose of this paper is to derive optimal rules for variable-length mastery tests in case three mastery classification decisions (nonmastery, partial mastery, and mastery) are distinguished. In a variable-length or adaptive mastery test, the decision is to classify a subject as a master, a partial master, a nonmaster, or continuing sampling and administering another test item. The framework of minimax sequential decision theory is used; that is, optimal sequential rules minimizing the maximum expected losses associated with all possible decision rules. The binomial model is assumed for the conditional probability of a correct response given the true level of functioning, whereas threshold loss is adopted for the loss function involved. Monotonicity conditions are derived, that is, conditions sufficient for optimal sequential rules to be in the form of cutting scores. The paper concludes with an empirical example of a computerized adaptive mastery test for concept-learning in medicine.

Key words: adaptive mastery testing, minimax sequential rules, monotonicity conditions, least favorable prior, binomial distribution, threshold loss.

Submitted to *Applied Psychological Measurement*

Introduction

Over the past few decades, several adaptive instructional programs have been implemented on computers (e.g., Bejar & Braun, 1994; Bork, 1984; De Diana & Vos, 1988; Gegg-Harrison, 1992; Hambleton, 1974; Hansen, Ross, & Rakow, 1977; Kontos, 1985; Koper 1995; Vos, 1988, 1994a, 1994b, 1995a, 1995b, 1998a). Although these programs come in many forms, all of these forms share the same basic design. All systems are basically series of comparatively small instructional units or modules through which students are routed by means of a few achievement test items administered right after a module. An important challenge to their instructors is to adapt instruction to individual learner differences (aptitudes, prior knowledge) and learning needs (amount and sequence of instruction).

A situation that often arises in such systems is the following: It is desired to classify students as either a master or a nonmaster, i.e., those students who pass or fail the mastery test at the end of the instructional unit. Students who pass the mastery test may proceed with the next, presumably more complex module. Students who fail, however, are retained at the same instructional unit.

The situation described above refers to a fixed-length mastery test, where the performance on a fixed number of test items is used for deciding on either mastery or nonmastery. The fixed-length mastery problem has been studied extensively in the literature within the framework of (empirical) Bayesian decision theory (e.g., Cronbach & Gleser, 1965; Davis et al., 1973; De Gruijter & Hambleton, 1984; Hambleton & Novick, 1973; Huynh, 1976, 1977; Swaminathan et al., 1975; van der Linden, 1980, 1990; van der Linden & Mellenbergh, 1977; Wilcox, 1977). In addition, optimal rules for the fixed-length mastery problem have also been derived within the framework of the minimax (Wald) strategy (e.g., Huynh, 1980; Veldhuijzen, 1982).

In both approaches, the following two basic elements are distinguished: A psychometric model relating observed test scores and student's true level of functioning to each other, and a loss structure evaluating the total costs and benefits of all possible decision outcomes. Within the framework of Bayesian decision theory, optimal rules (i.e., the Bayes rules) are obtained by minimizing the posterior expected loss. The Bayes principle assumes that prior knowledge about student's true level of functioning is available and can be characterized by a probability distribution called the prior.

Using minimax decision theory, optimal rules (i.e., the minimax rules) are obtained by minimizing the maximum expected losses associated with all possible decision rules. Decision rules are hereby prescriptions specifying for each possible observed test score what action has to be taken. In fact, the minimax principle assumes that it is best to prepare for the worst and to establish the maximum expected loss for each possible decision rule (e.g., van der Linden, 1980). In other words, the minimax decision rule is a bit conservative and pessimistic (Coombs, Dawes, & Tversky, 1970).

Although the minimax principle assumes that no prior knowledge about true level of functioning is available, a minimax rule can be conceived as a rule that is based on minimization of posterior expected loss as well, but under the restriction that the prior is the least favorable element of the class of priors (e.g., Ferguson, 1967, Sect. 1.6). In other words, there exists a (least favorable) prior distribution on the true level of functioning such that the corresponding Bayes solution is exactly the same as the minimax decision rule (Huynh, 1980).

The test at the end of the instructional unit does not necessarily have to be a fixed-length mastery test but might also be a variable-length mastery test. In this case, in addition to the decisions declaring mastery or nonmastery, also the decision of continuing sampling and administering another item is available. In fact, variable-length mastery tests are more common in individualized instructional programs than fixed-length mastery tests. This is because variable-length mastery tests offer the possibility to provide shorter tests for those students who have clearly attained a certain level of mastery (or clearly nonmastery) and longer tests for whom the mastery decision is not as clear-cut (Lewis & Sheehan, 1990). The variable-length mastery test is also known as an adaptive mastery test (AMT).

The purpose of this paper is to derive optimal decision rules for adaptive mastery testing. Doing so, in addition to the decision of continuing sampling and administering another item, not two but three possible classification decisions are distinguished, namely declaring mastery, partial mastery, or nonmastery.

A useful application of the derived optimal rules is decision making in computerized adaptive instructional systems. The successful implementation of a computerized adaptive instructional system depends, at least in part, upon the availability of appropriate testing and decision making procedures to guide the student through the system (Hambleton, 1974). For instance, if a student is not directed to an appropriate module, his/her motivation may decrease due to the instruction not being matched to his/her specific learning characteristics.

Several approaches have been proposed for deriving optimal decision rules for the variable-length mastery problem. One of the earliest approaches dates back to Ferguson (1969a, 1969b) using Wald's sequential probabily ratio test (SPRT). In Ferguson's approach, the condional probability of a correct response given the true level of functioning (i.e., the psychometric model) is modeled as a binomial distribution. Alternative adaptive mastery testing procedures within an SPRT-framework, in combination with item response theory for modelling the conditional probability of a correct response given the true level of functioning, have been proposed by Kingsbury and Weiss (1983), Reckase (1983), and Spray and Reckase (1996).

Recently, Lewis and Sheehan (1990) and Sheehan and Lewis (1992) have proposed Bayesian sequential decision theory (e.g., DeGroot, 1970; Ferguson, 1967; Lehmann, 1959; Lindgren, 1976) to derive optimal rules for the variable-length mastery problem. Like in the fixed-length mastery problem, optimal sequential rules are obtained by choosing the action (i.e., mastery, nonmastery, or continuing sampling) that minimizes posterior expected loss at each stage of sampling using techniques of dynamic programming (i.e., backward induction). Doing so, the posterior expected loss associated with continuing sampling is determined by averaging the posterior expected loss associated with each of the possible future decision outcomes relative to the probability of observing those outcomes. For known psychometric model, prior, and observed item response vector, this probability can be calculated. As pointed out by Lewis and Sheehan (1990), the action chosen at each stage of sampling is optimal with respect to the entire adaptive mastery testing procedure.

In the present paper, the framework of minimax sequential decision theory (e.g., DeGroot, 1970; Ferguson, 1967; Lehmann, 1959; Lindgren, 1976) is proposed to derive optimal rules for the variable-length mastery problem; that is, minimizing the maximum expected losses associated with each of the possible decision outcomes. The maximum expected loss associated with continuing sampling is determined again by averaging the maximum expected losses associated with each of the possible future decision outcomes relative to the probability of observing those outcomes. For the prior needed to compute this probability, the least favorable prior will be taken.

Optimal sequential rules for the adaptive four-action mastery problem, following Ferguson (1969a, 1969b), will be derived under the assumption of a binomial model for the conditional probability of a correct response given the true level of functioning. The choice of this psychometric model assumes that, given the true level of functioning, each item has the same probability of being correctly answered, or that items are sampled at random. For the loss structure involved, the well-known threshold loss function is adopted. Furthermore, conditions sufficient for optimal sequential rules to be monotone, that is, in the form of cutting scores, are derived. The procedures for sequentially setting optimal cutting scores are demonstrated by an empirical example of a computerized adaptive mastery test for concept-learning in medicine.

Notations

Within the framework of both minimax and Bayesian sequential decision theory, optimal rules can be obtained without specifying a maximum test length. In the following, however, an adaptive four-action mastery test is supposed to have a maximum length of n ($n \geq 1$). As pointed out by Ferguson (1969a, 1969b), a maximum test length is needed in order to classify within a reasonable period of time those students for whom the decision of declaring mastery, partial mastery, or nonmastery is not as clear-cut.

In the following, the observed item response at each stage of sampling k ($1 \leq k \leq n$) for a randomly sampled student will be denoted by a discrete random variable X_k, with realization x_k. The observed response variables $X_1,...,X_k$ are assumed to be independent and identically distributed for each value of k ($1 \leq k \leq n$), and take the values 0 and 1 for respectively correct and incorrect responses to the kth item. Furthermore, let the observed number-correct score be denoted by a discrete random variable $S_k = X_1 +...+ X_k$ ($1 \leq k \leq n$), with realization $s_k = x_1 +...+ x_k$ ($0 \leq s_k \leq k$).

Student's true level of functioning is unknown due to measurement and sampling error. All that is known is his/her observed number- correct score from a small sample of test items. In other words, the mastery test is not a perfect indicator of student's true performance. Therefore, let student's true level of functioning be denoted by a continuous random variable T, with realization $t \in [0,1]$.

Finally, assuming $X_1 = x_1,...,X_k = x_k$ has been observed, the two basic elements of minimax sequential decision making discussed earlier can now be formulated as follows: A psychometric model relating observed number-correct score s_k to student's true level of functioning t at each stage of sampling k ($1 \leq k \leq n$), and a loss function describing the loss

$l(a_i(x_1,...,x_k),t)$ incurred when action $a_i(x_1,...,x_k)$ is taken for the student whose true level of functioning is t.

Statement of the Adaptive Four-Action Mastery Problem

In the following, linking up with common practice in criterion-referenced testing, the decision rules for the adaptive four-action mastery problem at each stage of sampling k ($1 \leq k \leq n$) are assumed to be in the form of sequential cutting scores $s_{c1}(k)$ and $s_{c2}(k)$ on the observed number-correct scale S_k, where $s_{c1}(k) < s_{c2}(k)$. Conditions sufficient for sequentially setting cutting scores (i.e., monotonicity conditions) are given later on.

Given the observed item response vector $(x_1,...,x_k)$, the following four actions can be taken at each stage of sampling k ($1 \leq k < n$): First, the action nonmastery $a_1(x_1,...,x_k)$ is taken if $s_k \leq s_{c1}(k)$. Second, the action partial mastery $a_2(x_1,...,x_k)$ is taken if $s_{c1}(k) < s_k < s_{c2}(k)$. Third, the action mastery $a_3(x_1,...,x_k)$ is taken if $s_k \geq s_{c2}(k)$. Fourth, the action continue sampling $a_4(x_1,...,x_k)$ is taken if the maximum expected loss associated with this rule is smallest among all possible decision rules. At the final stage of sampling, n, only one of the three classification actions nonmastery, partial mastery, or mastery can be taken.

In addition to the sequential cutting scores $s_{c1}(k)$ and $s_{c2}(k)$, criteria levels t_{c1} and t_{c2} ($0 \leq t_{c1} < t_{c2} \leq 1$) on the true level of functioning scale T can be identified. A student is considered a true nonmaster and true master if his/her true level of functioning t is smaller or larger than t_{c1} and t_{c2}, respectively. Furthermore, a student is considered a partial true master if $t_{c1} < t < t_{c2}$. Unlike the sequential cutting scores $s_{c1}(k)$ and $s_{c2}(k)$, the criteria levels t_{c1} and t_{c2} do not depend on the observed item response vector $(x_1,...,x_k)$ but must be specified in advance by the decision-maker using methods of standard setting (e.g., Angoff, 1971; Ebel, 1972; Nedelsky, 1954).

Given the values of the criteria levels t_{c1} and t_{c2} on T, the adaptive four-action mastery problem can now be stated at each stage of sampling k ($1 \leq k < n$) as choosing values of $s_{c1}(k)$ and $s_{c2}(k)$ or continue sampling such that the maximum expected loss associated with the preferred rule among all possible decision rules is minimal. At the final stage of sampling, n, the problem reduces to choosing values of $s_{c1}(n)$ and $s_{c2}(n)$ such that the maximum expected loss associated with action $a_1(x_1,...,x_n)$, $a_2(x_1,...,x_n)$, or $a_3(x_1,...,x_n)$ is the smallest.

Constant Losses

Generally speaking, a loss function evaluates for each possible action the consequences for a student whose true level of functioning is t. These consequences may reflect all moral, ethical, social, economic, psychological, etc. considerations deemed relevant.

As earlier indicated, here the well-known threshold loss function (e.g., Ben-Shakhar & Beller, 1983; Chuang et al., 1981; Davis et al. 1973; Hambleton & Novick, 1973; Huynh, 1976; Lewis & Sheehan, 1990; Novick & Lewis, 1974; Raju et al, 1991; Swaminathan et al., 1975) is adopted as the loss structure. The choice of this function implies that the "seriousness" of all possible consequences of the decisions can be summarized by possibly different constants, one for each of the possible decision outcomes. Other loss functions have also been frequently used in the literature, such as linear loss (e.g., Huynh, 1980; van der Linden & Vos, 1996; Vos, 1990, 1991, 1995c, 1997a, 1997b, 1998a, 1998b, 1998c) and normal-ogive loss (e.g., Novick & Lindley, 1979; van der Linden, 1981).

For our variable-length mastery problem, a threshold loss function can be formulated as a natural extension of the one for the standard fixed-length two-action mastery problem at each stage of sampling k ($1 \leq k \leq n$) as follows (see also Lewis & Sheehan, 1990):

Table 1. Table for Threshold Loss Function at Stage k ($1 \leq k \leq n$) of Sampling

True Level / Action	$T \leq t_{c1}$	$t_{c1} < T < t_{c1}$	$T \geq t_{c2}$
$a_1(x_1, \ldots, x_k)$	ke	$l_{12} + ke$	$l_{13} + ke$
$a_2(x_1, \ldots, x_k)$	$l_{12} + ke$	ke	$l_{23} + ke$
$a_3(x_1, \ldots, x_k)$	$l_{31} + ke$	$l_{32} + ke$	ke

The value e represents the costs of administering one random item. For the sake of simplicity, following Lewis and Sheehan (1990), these costs are assumed to be equal for each decision outcome as well as for each sampling occasion. Applying an admissable positive linear transformation (e.g., Luce & Raiffa, 1957), and assuming the losses l_{11}, l_{22}, and l_{33} associated with the correct decision outcomes are equal and take the smallest values, the threshold loss function in Table 1 was rescaled in such a way that l_{11}, l_{22}, and l_{33} were equal to zero.

Furthermore, the loss function associated with action $a_1(x_1,...,x_k)$ must be nondecreasing in t, since action $a_1(x_1,...,x_k)$ is most appropriate when t is small. Similarly, the loss function associated with action a_3 $(x_1,...,x_k)$ must be nonincreasing in t due to the fact that action $a_3(x_1,...,x_k)$ is most appropriate when t takes large values. Since it cannot be determined beforehand whether l_{21} is smaller than, equal to, or larger than l_{23}, the form of the loss function associated with action $a_2(x_1,...,x_k)$ is unknown. The loss associated with the correct partial mastery decision (i.e., l_{22}), however, must be the smallest.

The loss parameters l_{ij} $(i = 1,2,3; i \neq j)$ associated with the incorrect decisions have to be empirically assessed, for which several methods have been proposed in the literature. Most texts on decision theory, however, propose lottery methods (e.g., Luce & Raiffa, 1957) for assessing loss functions empirically. In general, the consequences of each pair of actions and true level of functioning are scaled in these methods by looking at the most and least preferred outcomes (see also Vos, 1988).

Psychometric Model

A psychometric model is needed to specify the statistical relation between the observed number-correct score and student's true level of functioning at each stage of sampling. As earlier remarked, here the well-known binomial model will be adopted.

As indicated by van den Brink (1982), when tests are criterion-referenced tests by means of sampling from item domains, such as in our adaptive four-action mastery problem, the well-known binomial model is a natural choice for estimating the distribution of student's number-correct score s_k and making decisions. Hence, the binomial density function is a convenient choice as the psychometric model involved (see also Millman, 1972). Its distribution relating the observed number-correct score s_k $(0 \leq s_k \leq k)$ to student's true level of functioning t, $f(s_k \mid t)$, at stage k of sampling $(1 \leq k \leq n)$ can be written as follows:

$$f(s_k | t) = \binom{k}{s_k} t^{s_k} (1-t)^{k-s_k}. \tag{1}$$

If each response is independent of the other, and if the examinee's probability of a correct answer remains constant, the probability function of s_k, given the true level of functioning t, is given by Equation 1 (Wilcox, 1981). The binomial model assumes that the test given to each student is a random sample of items drawn from a large item pool (van den Brink, 1982; Wilcox, 1981). Therefore, for each subject a new random sample of items must be drawn in practical applications of the adaptive four-action mastery problem, such as, for instance, in computerized adaptive instructional systems.

Sufficient Conditions for Optimal Sequential Rules to be Monotone

As noted earlier, the optimal sequential rules in this paper are assumed to have monotone forms. The restriction to monotone rules, however, 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 (Ferguson, 1967, p.55). Using a minimax rule, the minimum of the maximum expected losses associated with all possible decision rules is taken as the criterion of optimality.

The maximum expected loss for continuing sampling is determined by averaging the maximum expected loss associated with each of the possible future decision outcomes relative to the probability of observing those outcomes (Lewis & Sheehan, 1990). Therefore, it follows immediately that the conditions sufficient for setting cutting scores for the fixed-length three-action mastery problem, are also sufficient for the adaptive four-action mastery problem at each stage of sampling. Generally, conditions sufficient for setting cutting scores for the fixed-length multiple-decision problem are given in Ferguson (1967, p.286).

First, $f(s_k \mid t)$ must have a monotone likelihood ratio (MLR); that is, it is required that for any $t_1 > t_2$, the likelihood ratio $f(s_k \mid t_1)/f(s_k \mid t_2)$ is a nondecreasing function of s_k. MLR implies that the higher the observed number-correct score, the more likely it will be that the latent true level of functioning is high too. Second, the condition of monotonic loss must hold; that is, there must be an ordering of the actions such that for each pair of adjacent actions the loss functions possess at most one point of intersection.

The binomial density function belongs to the monotone likelihood ratio family (Ferguson, 1967, chap. 5). Furthermore, it can be verified from Table 1 that for threshold loss the condition of monotonic loss is satisfied if at each stage of sampling k ($1 \leq k \leq n$):

$$\begin{aligned} &(l_{13}+ke) - (l_{23}+ke) \geq (l_{12}+ke) - ke \geq ke - (l_{21}+ke) \\ &(l_{23}+ke) - ke \geq ke - (l_{32}+ke) \geq (l_{21}+ke) - (l_{31}+ke), \end{aligned} \qquad (2)$$

or, equivalently,

$$\begin{aligned} l_{13} - l_{23} \geq l_{12} \geq - l_{21} \\ l_{23} \geq - l_{32} \geq l_{21} - l_{31}. \end{aligned} \tag{3}$$

Optimizing Cutting Scores for the Adaptive Four-Action Mastery Problem

In this section, it will be shown how optimal cutting scores for the adaptive four-action mastery problem can be derived using the framework of minimax sequential decision theory. Doing so, first the minimax principle will be applied to the fixed-length three-action mastery problem, given an observed item response vector $(x_1,...,x_k)$ $(1 \leq k \leq n)$, by determining which of the maximum expected losses associated with the three actions $a_1(x_1,...,x_k)$, $a_2(x_1,...,x_k)$, or $a_3(x_1,...,x_k)$ is the smallest.

Next, applying the minimax sequential principle, decision rules for the adaptive four-action mastery problem are optimized at each stage of sampling k $(1 \leq k \leq n)$ by comparing this quantity with the maximum expected loss associated with action $a_4(x_1,...,x_k)$ (i.e., continuing sampling).

Applying the Minimax Principle to the Fixed-Length Mastery Problem

Given $X_1 = x_1,...,X_k = x_k$ $(1 \leq k \leq n)$, it follows that the minimax decision rule for the fixed-length three-action mastery problem can be found by minimizing the maximum expected losses associated with the actions $a_1(x_1,...,x_k)$, $a_2(x_1,...,x_k)$, and $a_3(x_1,...,x_k)$.

In searching for a minimax rule for the fixed-length three-action mastery problem, assuming the conditions of monotonicity are satisfied, we may confine ourselves to partitions of the range of the observed number-correct scores into three disjoint subsets $A_1 = \{s_k;\ s_k \leq s_{c1}(k)\}$, $A_2 = \{s_k;\ s_{c1}(k) < s_k < s_{c2}(k)\}$, and $A_3 = \{s_k;\ s_k \geq s_{c2}(k)\}$ for action $a_1(x_1,...,x_k)$, $a_2(x_1,...,x_k)$, and $a_3(x_1,...,x_k)$ each with a conditional probability of $P(A_1 \mid t)$, $P(A_2 \mid t)$, and $P(A_3 \mid t)$, respectively. It follows that the minimax decision rule can be obtained by minimizing the maximum expected lossess associated with the actions $a_1(x_1,...,x_k)$, $a_2(x_1,...,x_k)$, and $a_3(x_1,...,x_k)$, or, equivalently, by minimizing the following function:

$$M(s_{c1}(k),s_{c2}(k)) = \max\,\{L_1(s_{c1}(k)),L_2(s_{c1}(k),s_{c2}(k)),L_3(s_{c2}(k))\}, \tag{4}$$

where

$$L_1(s_{c1}(k)) = \sup_{t \le t_{c1}} l(a_1(x_1,...,x_k),t)P(A_1 \mid t)$$

$$L_2(s_{c1}(k),s_{c2}(k)) = \sup_{t_{c1}<t<t_{c2}} l(a_2(x_1,...,x_k),t)P(A_2 \mid t)$$

$$L_3(s_{c2}(k)) = \sup_{t \ge t_{c2}} l(a_3(x_1,...,x_k),t)P(A_3 \mid t). \tag{5}$$

The abovementioned procedure yields the optimal cutting scores $s'_{c1}(k)$ and $s'_{c2}(k)$ (i.e., the minimax cutting scores with $s'_{c1}(k) < s'_{c2}(k)$) for a fixed-length three-action mastery problem with test length k. These minimax cutting scores can be obtained by computing the values of $L_1(s_{c1}(k))$, $L_2(s_{c1}(k),s_{c2}(k))$, and $L_3(s_{c2}(k))$ for all possible values of $s_{c1}(k)$ and $s_{c2}(k)$, with $s_{c1}(k)$, $s_{c2}(k) = 0,1,2,...,k$, and then selecting those values $s'_{c1}(k)$ and $s'_{c2}(k)$ at which $M(s_{c1}(k),s_{c2}(k))$ is the smallest.

For our problem we are not primarily interested in determining the optimal cutting scores $s'_{c1}(k)$ and $s'_{c2}(k)$, however, but more in determining which of the three actions $a_1(x_1,...,x_k)$, $a_2(x_1,...,x_k)$, or $a_3(x_1,...,x_k)$ is optimal, given an observed item response vector $(x_1,...,x_k)$ with number correct-score s_k ($0 \le s_k \le k$). In this situation it is more convenient to apply the following sequential procedure: First, the maximum expected losses associated with actions $a_3(x_1,...,x_k)$ and $a_2(x_1,...,x_k)$ are compared with each other. If the maximum expected loss associated with action a_3 is the smallest of these two quantities, then, action $a_3(x_1,...,x_k)$ is taken. If the maximum expected loss associated with action $a_2(x_1,...,x_k)$, however, is the smallest, then, this quantity must be compared with the maximum expected loss associated with action $a_1(x_1,...,x_k)$. If the maximum expected loss associated with action $a_2(x_1,...,x_k)$ is the smallest of these two quantities, then, action $a_1(x_1,...,x_k)$ is chosen; otherwise action $a_1(x_1,...,x_k)$ is chosen.

Applying the abovementioned procedure to a given item response vector $(x_1,...,x_k)$ with observed number correct-score s_k ($0 \le s_k \le k$), it can easily be verified from Table 1 that mastery ($a_3(x_1,...,x_k)$) is declared when the number-correct score s_k is such that

$$\sup_{t \le t_{c2}} (l_{31} + ke) \sum_{y=s_k}^{k} \binom{k}{y} t^y (1-t)^{k-y} + \sup_{t_{c1} < t < t_{c2}} (l_{32} + ke) \sum_{y=s_k}^{k} \binom{k}{y} t^y (1-t)^{k-y} +$$

$$\sup_{t \ge t_{c2}} (ke) \sum_{y=0}^{s_k - 1} \binom{k}{y} t^y (1-t)^{k-y} \le \sup_{t \le t_{c1}} (l_{21} + ke) \sum_{y=s_k}^{k} \binom{k}{y} t^y (1-t)^{k-y} +$$

$$\sup_{t_{c1} < t < t_{c2}} (ke) \sum_{y=s_k}^{k} \binom{k}{y} t^y (1-t)^{k-y} + \sup_{t \ge t_{c2}} (l_{32} + ke) \sum_{y=0}^{s_k - 1} \binom{k}{y} t^y (1-t)^{k-y}, \tag{6}$$

where $y = 0,1,..k$ represents all possible values the number-correct score s_k can take after having observed k item responses ($1 \le k \le n$). Since the cumulative binomial density function is decreasing in t, it follows that the inequality in (6) can be written as:

$$(l_{31} + ke) \sum_{y=s_k}^{k} \binom{k}{y} t_{c1}^{t} (1-t_{c1})^{k-y} + (l_{32} + ke) \sum_{y=s_k}^{k} \binom{k}{y} t_{c2}^{y} (1-t_{c2})^{k-y} -$$

$$(l_{32} + ke) \sum_{y=s_k}^{k} \binom{k}{y} t_{c1}^{y} (1-t_{c1})^{k-y} + (ke) \sum_{y=0}^{s_k - 1} \binom{k}{y} t_{c2}^{y} (1-t_{c2})^{k-y} \le$$

$$(l_{21} + ke) \sum_{y=s_k}^{k} \binom{k}{y} t_{c1}^{y} (1-t_{c1})^{k-y} + (ke) \sum_{y=s_k}^{k} \binom{k}{y} t_{c2}^{y} (1-t_{c2})^{k-y} -$$

$$(ke) \sum_{y=s_k}^{k} \binom{k}{y} t_{c1}^{y} (1-t_{c1})^{k-y} + (l_{32} + ke) \sum_{y=0}^{s_k - 1} \binom{k}{y} t_{c2}^{y} (1-t_{c2})^{k-y}. \tag{7}$$

Rearranging terms, it follows that mastery is declared when the number-correct score s_k is such that:

$$l_{31} - l_{21} - l_{32}) \sum_{y=s_k}^{k} \binom{k}{y} t_{c1}^{y} (1-t_{c1})^{k-y} + (l_{23} + l_{32}) \sum_{y=s_k}^{k} \binom{k}{y} t_{c2}^{y} (1-t_{c2})^{k-y} \le l_{23} \tag{8}$$

If the inequality in (8) does not hold, it can easily be verified from Table 1 that partial mastery ($a_2(x_1,...,x_k)$) is declared if it holds for number-correct score s_k ($0 \le s_k \le k$) that

$$\sup_{t \le t_{c2}} (l_{21} + ke) \sum_{y=s_k}^{k} \binom{k}{y} t^y (1-t)^{k-y} + \sup_{t_{c1}<t<t_{c2}} (ke) \sum_{y=s_k}^{k} \binom{k}{y} t^y (1-t)^{k-y} +$$

$$\sup_{t \ge t_{c2}} (l_{23} + ke) \sum_{y=0}^{s_k-1} \binom{k}{y} t^y (1-t)^{k-y} \le \sup_{t \le t_{c1}} (ke) \sum_{y=s_k}^{k} \binom{k}{y} t^y (1-t)^{k-y} +$$

$$\sup_{t_{c1}<t<t_{c2}} (l_{12} + ke) \sum_{y=s_k}^{k} \binom{k}{y} t^y (1-t)^{k-y} + \sup_{t \ge t_{c2}} (l_{13} + ke) \sum_{y=0}^{s_k-1} \binom{k}{y} t^y (1-t)^{k-y}, \tag{9}$$

and that nonmastery ($a_1(x_1,...,x_k)$) is declared otherwise.

Rearranging terms, it follows that partial mastery is declared if the number-correct score is such that

$$(l_{12} + l_{21}) \sum_{y=s_k}^{k} \binom{k}{y} t_{c1}^y (1-t_{c1})^{k-y} + (l_{13} - l_{23} - k_{12}) \sum_{y=s_k}^{k} \binom{k}{y} t_{c2}^y (1-t_{c2})^{k-y} \le l_{13} - l_{23}, \tag{10}$$

and that nonmastery is declared otherwise.

Applying the Minimax Sequential Principle to Adaptive Mastery Testing

The optimal sequential rule (i.e., the minimax sequential rule) at the final stage of sampling n is given by the minimum of the maximum expected losses associated with the actions $a_1(x_1,...,x_n)$, $a_2(x_1,...,x_n)$, and $a_3(x_1,...,x_n)$, since the action $a_4(x_1,...,x_n)$ (i.e., continuing sampling) is not available at that stage of sampling. Minimax sequential rules at the other stages of sampling k (i.e., $1 \le k < n$) are found by first computing the minimum of the maximum expected losses associated with the actions $a_1(x_1,...,x_k)$, $a_2(x_1,...,x_k)$, and $a_3(x_1,...,x_k)$, and next comparing this quantity with the maximum expected loss associated with the action $a_4(x_1,...,x_k)$.

Let $d_k(x_1,...,x_k)$ denote the action $a_1(x_1,...,x_k)$, $a_2(x_1,...,x_k)$, or $a_3(x_1,...,x_k)$ ($1 \le k \le n$) yielding the minimum of the maximum expected losses associated with these three actions, and let the maximum expected loss associated with $d_k(x_1,...,x_k)$ be denoted as $V_k(x_1,...,x_k)$. These notations can also be generalized to the case that no observations has been taken yet; that is, $d_0(x_0)$ denotes the action $a_1(x_0)$, $a_2(x_0)$, or $a_3(x_0)$ which yields the smallest of the maximum expected losses associated with these three actions, and $V_0(x_0)$ denotes the smallest maximum expected loss associated with $d_0(x_0)$. From the foregoing it then follows that the minimax sequential rules for our adaptive four-action mastery problem can be found by using the following backward induction computational scheme:

First, the minimax sequential rule at the final stage of sampling n is computed which is given by $d_n(x_1,...,x_n)$; its associated maximum expected loss is given by $V_n(x_1,...,x_n)$. Next, the minimax sequential rule at stage (n-1) of sampling is computed by comparing $V_{n-1}(x_1,...,x_{n-1})$ with the maximum expected loss associated with action $a_4(x_1,...x_{n-1})$ (i.e., continuing sampling). As noted before, the maximum expected loss associated with taking one more observation, given $X_1 = x_1,...,X_{n-1} = x_{n-1}$, is computed by considering all possible future decision outcomes at stage n (i.e., $x_n = 0$ or 1) relative to the probability of observing those outcomes (i.e., backward induction).

Let $P(X_n \mid X_1 = x_1,...,X_{n-1} = x_{n-1})$ denote the conditional distribution of X_n, given the observed item response vector $(x_1,...,x_{n-1})$, then, the maximum expected loss associated with taking one more observation after (n-1) observations have been taken, $E[V_n(x_1,...,x_{n-1},X_n) \mid X_1 = x_1,...,X_{n-1} = x_{n-1}]$, is computed as follows:

$$\begin{aligned}
&E[V_n(x_1,...,x_{n-1},X_n) \mid X_1 = x_1,...,X_{n-1} = x_{n-1}] = \\
&V_n(x_1,...,x_n = 0)P(X_n = 0 \mid X_1 = x_1,...,X_{n-1} = x_{n-1}) + \\
&V_n(x_1,...,x_n = 1)P(X_n = 1 \mid X_1 = x_1,...,X_{n-1} = x_{n-1}). \qquad (11)
\end{aligned}$$

$P(X_k \mid X_1 = x_1,...,X_{k-1} = x_{k-1})$ is also called the posterior predictive distribution of X_k at stage (k-1) of sampling $(1 \leq k \leq n)$. It will be indicated in the next section how this posterior predictive distribution can be computed.

Hence, given $X_1 = x_1,...,X_{n-1} = x_{n-1}$, it follows that at stage (n-1) of sampling the minimax sequential rule is given by: Take one more observation (i.e., action $a_4(x_1,...,x_{n-1})$) if $E[V_n(x_1,...,x_{n-1},X_n) \mid X_1 = x_1,...,X_{n-1} = x_{n-1}]$ is smaller than $V_{n-1}(x_1,...,x_{n-1})$, and take action $d_{n-1}(x_1,...,x_{n-1})$ if $E[V_n(x_1,...,x_{n-1},X_n) \mid X_1 = x_1,...,X_{n-1} = x_{n-1}]$ is larger than $V_{n-1}(x_1,...,x_{n-1})$. If $E[V_n(x_1,...,x_{n-1},X_n) \mid X_1 = x_1,...,X_{n-1} = x_{n-1}]$ and $V_{n-1}(x_1,...,x_{n-1})$ are equal to each other, it does not matter whether or not the decision-maker takes one more observation.

Analogous to the computation at stage (n-1), the minimax sequential rule at stage (n-2) can be computed by comparing $V_{n-2}(x_1,...,x_{n-2})$ with the maximum expected loss associated with taking one more observation. In order to compute the maximum expected loss associated with taking one more observation at stage (n-2), first the minimum of the maximum expected losses associated with actions $a_1(x_1,...,x_{n-1})$, $a_2(x_1,...,x_{n-1})$, $a_3(x_1,...,x_{n-1})$, and $a_4(x_1,...,x_{n-1})$ at stage (n-1) is computed. This quantity is called the conditional minimax expected utility at stage (n-1), given $X_1 = x_1,...,X_{n-1} = x_{n-1}$, and will be denoted as $R_{n-1}(x_1,...,x_{n-1})$. It follows that $R_{n-1}(x_1,...,x_{n-1})$ is equal to the minimum of $V_{n-1}(x_1,...,x_{n-1})$ and $E[V_n(x_1,...,x_{n-1},X_n) \mid X_1 = x_1,...,X_{n-1} = x_{n-1}]$.

At stage (n-2) of sampling, the maximum expected loss associated with taking one more observation after (n-2) observations, $E[R_{n-1}(x_1,...,x_{n-2},X_{n-1}) \mid X_1 = x_1,...,X_{n-2} = x_{n-2}]$, can now be computed as follows:

$$
\begin{aligned}
&E[R_{n-1}(x_1,...,x_{n-2},X_{n-1}) \mid X_1 = x_1,...,X_{n-2} = x_{n-2}] = \\
&R_{n-1}(x_1,...,x_{n-1} = 0)P(X_{n-1} = 0 \mid X_1 = x_1,...,X_{n-2} = x_{n-2}) + \\
&R_{n-1}(x_1,...,x_{n-1} = 1)P(X_{n-1} = 1 \mid X_1 = x_1,...,X_{n-2} = x_{n-2}). \qquad (12)
\end{aligned}
$$

Given $X_1 = x_1,...,X_{n-2} = x_{n-2}$, the minimax sequential rule at stage (n-2) of sampling can now be found by comparing $E[R_{n-1}(x_1,...,x_{n-2},X_{n-1}) \mid X_1 = x_1,..., X_{n-2} = x_{n-2}]$ and $V_{n-2}(x_1,...,x_{n-2})$ with each other. Hence, it follows that one more observation is taken (i.e., action $a_4(x_1,...,x_{n-2})$) if $E[R_{n-1}(x_1,...,x_{n-2},X_{n-1}) \mid X_1 = x_1,...,X_{n-2} = x_{n-2}]$ is smaller than $V_{n-2}(x_1,...,x_{n-2})$, and action $d_{n-2}(x_1,...,x_{n-2})$ is taken if $E[R_{n-1}(x_1,...,x_{n-2},X_{n-1}) \mid X_1 = x_1,...,X_{n-2} = x_{n-2}]$ is larger than $V_{n-2}(x_1,...,x_{n-2})$. In the case of equality between $V_{n-2}(x_1,...,x_{n-2})$ and $E[R_{n-1}(x_1,...,x_{n-2},X_{n-1}) \mid X_1 = x_1,...,X_{n-2} = x_{n-2}]$, it does not matter again whether or not the decision-maker takes one more observation.

Following the same computational backward scheme as in determining the minimax sequential rules at stages (n-1) and (n-2), the minimax sequential rule at stage (n-3) can be obtained by comparing $V_{n-3}(x_1,...,x_{n-3})$ and $E[R_{n-2}(x_1,...,x_{n-3},X_{n-2}) \mid X_1 = x_1,...,X_{n-3} = x_{n-3}]$ with each other. The conditional minimax expected loss at stage (n-2), $R_{n-2}(x_1,...,x_{n-2})$, is thereby computed inductively as the minimum of $V_{n-2}(x_1,...,x_{n-2})$ and $E[R_{n-1}(x_1,...,x_{n-2},X_{n-1}) \mid X_1 = x_1,...,X_{n-2} = x_{n-2}]$.

Similarly, following the same computational backward scheme, the minimax sequential rules at stages (n-3),...,1,0 are computed. The minimax sequential rule at stage 0 denotes the decision whether or not to take at least one observation.

Determination of the Least Favorable Prior

As mentioned before, there exists a least favorable prior distribution on the true level of functioning such that the corresponding Bayes sequential rule is exactly the same as the minimax sequential rule. In computing the posterior predictive distribution $P(X_k \mid X_1 = x_1,..., X_{k-1} = x_{k-1})$, the least favorable prior on the true level of functioning is needed ($1 \leq k \leq n$). Therefore, in this section first the form of the least favorable prior will be investigated.

Let $I_P(r,s)$ denote the incomplete beta function with parameters r and s (r,s > 0) as tabulated in Pearson (1934) (see also Johnson & Kotz, 1970). It has been known for some time that

$$\sum_{x=m}^{n}\binom{n}{x}p^x(1-p)^{n-x} = I_p(m,n-m+1). \tag{13}$$

Hence, the inequalities in (8) and (10) can be written as:

$$(l_{31}-l_{21}-l_{32})I_{t_{c1}}(s_k,k-s_k+1) + (l_{23}+l_{32})I_{t_{c2}}(s_k,k-s_k+1) \leq l_{23}, \tag{14}$$

$$(l_{12}+l_{21})I_{t_{c1}}(s_k,k-s_k+1) + (l_{13}-l_{23}-l_{12})I_{t_{c2}}(s_k,k-s_k+1) \leq l_{13} - l_{23}. \tag{15}$$

Within the framework of Bayesian decision theory, given $X_1 = x_1,...,X_k = x_k$, it can be verified from Table 1 that mastery is declared for the fixed sample problem if number-correct score s_k ($0 \leq s_k \leq k$) is such that

$$\begin{aligned}&(l_{31}+ke)P(T \leq t_{c1} \mid s_k) + (l_{32}+ke)P(t_{c1} < T < t_{c2} \mid s_k) + (ke)P(T \geq t_{c2} \mid s_k) \leq \\ &(l_{21}+ke)P(T \leq t_{c1} \mid s_k) + (ke)P(T \geq t_{c2} \mid s_k) + (l_{23}+ke)P(T \geq t_{c2}).\end{aligned} \tag{16}$$

Rearranging terms, it can easily be verified from (16) that mastery is declared if

$$(l_{31}-l_{21}-l_{32})P(T \leq t_{c1} \mid s_k) + (l_{23}+l_{32})P(T \leq t_{c2} \mid s_k) \leq l_{23}. \tag{17}$$

Assuming an incomplete beta prior, it follows from an application of Bayes' theorem that under the assumed binomial model from (1), the posterior distribution of T will again be a member of the incomplete beta family (the conjugacy property, see e.g., Lehmann, 1959). In fact, if the incomplete beta function with parameters α and β (α, $\beta > 0$) is chosen as the prior distribution and student's observed number-correct score is s_k from a test of length k ($1 \leq k \leq n$), then the posterior distribution of T is $I_t(\alpha+s_k,\beta+k-s_k)$.

Hence, assuming an incomplete beta prior, it follows from (17) that mastery is declared if:

$$(l_{31}-l_{21}-l_{32})\, I_{t_{c1}}(\alpha+s_k,\beta+k-s_k) + (l_{23}+l_{32})\, I_{t_{c2}}(\alpha+s_k,\beta+k-s_k) \leq l_{23}. \qquad (18)$$

If the inequality in (18) is not satisfied, it can easily be verified from Table 1 that partial mastery is declared when number correct-score s_k is such that

$$\begin{aligned}&(l_{21}+ke)P(T \leq t_{c1} \mid s_k) + (ke)P(t_{c1} < T < t_{c2} \mid s_k) + (l_{23}+ke)P(T \geq t_{c2} \mid s_k) \leq \\ &(ke)P(T \leq t_{c1} \mid s_k) + (l_{12}+ke)P(t_{c1} < T < t_{c2} \mid s_k) + l_{13}P(T \geq t_{c2} \mid s_k),\end{aligned} \qquad (19)$$

and that nonmastery is declared otherwise. Assuming an incomplete beta prior, it follows that partial mastery is declared if

$$(l_{12}+l_{21})\, I_{t_{c2}}(\alpha+s_k,\beta+k-s_k) + (l_{13}-l_{23}-l_{12})\, I_{t_{c2}}(\alpha+s_k,\beta+k-s_k) \leq l_{13}-l_{23}, \qquad (20)$$

and that nonmastery is declared otherwise.

Comparing (14) and (15) with (18) and (20), respectively, it can be seen that the least favorable prior for the minimax solution is given by an incomplete beta prior $I_t(\alpha,\beta)$ with α sufficiently small and $\beta = 1$. It should be noted that the parameter $\alpha > 0$ can not be chosen equal to zero, because otherwise the prior distribution for T should be improper; that is, the prior does not integrate to 1 but to infinity. Hence, α must be chosen sufficiently small.

Computation of $P(X_k \mid X_1 = x_1,...,X_{k-1} = x_{k-1})$

To compute the maximum expected loss associated with taking one more observation at stage (k-1) of sampling ($1 \leq k \leq n$), $E[R_k(x_1,...,x_{k-1},X_k \mid X_1 = x_1,...,X_{k-1} = x_{k-1}]$, we need the posterior predictive distribution $P(X_k \mid X_1 = x_1,...,X_{k-1} = x_{k-1})$. From Bayes' theorem, it follows that:

$$P(X_k \mid X_1 = x_1,...,X_{k-1} = x_{k-1}) =$$
$$P(X_1 = x_1,...,X_k = x_k)/P(X_1 = x_1,...,X_{k-1} = x_{k-1}). \tag{21}$$

For the binomial model as the psychometric model involved and the incomplete beta function $I_t(\alpha,1)$ as least favorable prior (with α sufficiently small), it is known (e.g., Keats & Lord, 1962) that the unconditional distributions of $(X_1,...,X_k)$ and $(X_1,...,X_{k-1})$ are equal to:

$$P(X_1 = x_1,...,X_k = x_k) =$$
$$[\Gamma(\alpha+1)\Gamma(\alpha+s_k)\Gamma(1+k-s_k)]/[\Gamma(\alpha)\Gamma(1)\Gamma(\alpha+1+k)], \tag{22}$$

$$P(X_1 = x_1,...,X_{k-1} = x_{k-1}) =$$
$$[\Gamma(\alpha+1)\Gamma(\alpha+s_{k-1})\Gamma(k-s_{k-1})]/[\Gamma(\alpha)\Gamma(1)\Gamma(\alpha+k)], \tag{23}$$

where Γ is the usual gamma function. From (21)-(23) it then follows that the posterior predictive distribution of X_k, given $X_1 = x_1,..., X_k = x_k$, can be written as:

$$P(X_k \mid X_1 = x_1,...,X_{k-1} = x_{k-1}) =$$
$$[\Gamma(\alpha+s_k)\Gamma(1+k-s_k)\Gamma(\alpha+k)]/[\Gamma(\alpha+s_{k-1})\Gamma(k-s_{k-1})\Gamma(\alpha+1+k)]. \tag{24}$$

Using the well-known identity $\Gamma(j+1) = j\Gamma(j)$ and the fact that $s_k = s_{k-1}$ and $s_k = s_{k-1}+1$ for $x_k = 0$ and 1, respectively, it finally follows from (24) that:

$$P(X_k \mid X_1 = x_1,...,X_{k-1} = x_{k-1}) = \begin{cases} (k - s_{k-1})/(\alpha + k) & \text{if } x_k = 0 \\ (\alpha + s_{k-1})/(\alpha + k) & \text{if } x_k = 1. \end{cases} \tag{25}$$

An Empirical Example

The procedures for computing the minimax sequential decision rules were applied to a computerized adaptive four-action mastery test for concept-learning in medicine for freshmen. Concept-learning is the process in which subjects learn to categorize objects, processes or events, for instance, formation of diagnostic skills in medicine or psychology (see Tennyson and Cocchiarella, 1986, for a complete review of the theory of concept-learning).

In order to comply with the assumption of the binomial density as psychometric model involved, it was necessary to draw for each of the students a new random sample of items from a large pool of items.

The instructors of the program considered students as 'true masters' if they had mastered at least 60% of the total number of items covering the subject matter of that concept. Therefore, t_{c2} was fixed at 0.6. Furthermore, students were considered as 'true nonmasters' if they had mastered less than 50% of the total number of items covering the subject matter of the present concept. Therefore, t_{c1} was fixed at 0.5.

Finally, the constant cost for administering one random item was assumed to be rather small; hence, the value of e was set equal to 0.05.

<u>Results for the Minimax Sequential Rules</u>

Using the lottery method discussed in Vos (1988), assuming equal losses for the correct decisions l_{11}, l_{22}, and l_{33} and taking into account the requirements $l_{31} > l_{32}$ and $l_{13} > l_{12}$, the losses from Table 1 were empirically assessed by the instructors of the program; the results are presented in Table 2.

Table 2. <u>Threshold Loss Table at Stage k ($1 \leq k \leq n$) of Sampling for Empircal Example</u>

True Level / Action	$T \leq t_{c1}$	$t_{c1} < T < t_{c1}$	$T \geq t_{c2}$
$a_1(x_1, \ldots, x_k)$	ke	2 + ke	4 + ke
$a_2(x_1, \ldots, x_k)$	1 + ke	ke	2 + ke
$a_3(x_1, \ldots, x_k)$	3 + ke	1 + ke	ke

As noted before, the sign of $[l_{21}-l_{23}]$ could not be determined beforehand. In Table 2, it is assumed that $l_{23} > l_{21}$ (i.e., $l_{23} = 2$ and $l_{21} = 1$). In other words, for this specific empirical example it is assumed that the loss associated with taking action a_2 is twice as large for a student whose true level of functioning exceeds t_{c2} than for a student whose true level of functioning is below t_{c1}.

Using the numerical values for the loss parameters l_{ij} (i,j = 1,2,3) of Table 2, for a maximum of 30 items (i.e., n = 30), the appropriate action (i.e., nonmastery, partial mastery, mastery, or continue sampling) is depicted in Table 3 at each stage of sampling k ($0 \le k \le n$) for different number-correct score s_k ($0 \le s_k \le k$) as a closed interval.

Table 3. Appropriate Action Calculated by Stage of Sampling and Number-Correct

Stage of sampling	Appropriate Action by Number-Correct				
	Nonmastery	Continue	Partial Mastery	Continue	Mastery
0		0			
1	0	1			
2	0	[1,2]			
3	0	[1,3]			
4	0	[1,3]			4
5	[0,1]	[2,4]			5
6	[0,1]	[2,5]			6
7	[0,2]	[3,5]			[6,7]
8	[0,2]	[3,6]			[7,8]
9	[0,3]	[4,6]			[7,9]
10	[0,3]	[4,7]			[8,10]
11	[0,4]	[5,7]			[8,11]
12	[0,4]	[5,8]			[9,12]
13	[0,5]	[6,9]			[10,13]
14	[0,5]	[6,9]			[10,14]
15	[0,6]	[7,10]			[11,15]
16	[0,6]	[7,10]			[11,16]
17	[0,7]	[8,11]			[12,17]
18	[0,7]	[8,11]			[12,18]
19	[0,8]	9	10	[11,12]	[13,19]
20	[0,8]	[9,10]	11	12	[13,20]
21	[0,9]	10	11	[12,13]	[14,21]
22	[0,9]	[10,11]	12	[13,14]	[15,22]
23	[0,10]	11	12	[13,14]	[15,23]
24	[0,10]	[11,12]	13	[14,15]	[16,24]
25	[0,11]	12	[13,14]	15	[16,25]
26	[0,11]	[12,13]	14	[15,16]	[17,26]
27	[0,12]	13	[14,15]	16	[17,27]
28	[0,12]	13	[14,16]	17	[18,28]
29	[0,13]	14	[15,16]	17	[18,29]
30	[0,14]		[15,17]		[18,30]

Table 3 was constructed by using the following backward induction computational scheme. First, the appropriate action at each stage of sampling k ($0 \leq k \leq n$) for the fixed-length three-action mastery problem (i.e., $d_k(x_1,...,x_k)$) was determined by examining if the inequalities in (8) and (10) were satisfied. More specifically, mastery was declared for those values of s_k ($0 \leq s_k \leq k$) for which the inequality in (8) was satisfied; partial mastery was declared for those values of s_k for which the inequality in (8) was not satisfied but the inequality in (10) did hold; and mastery was declared for those values of s_k for which both the inequalities in (8) and (10) were not satisfied. Note that it can be inferred from Table 2 that $s_{c1}(30)$ and $s_{c2}(30)$ are equal to 14 and 18, respectively.

Next, the appropriate action nonmastery, partial mastery, or mastery and its associated maximum expected loss were computed at stage 29 of sampling for $s_{29} = 0,...,29$ (i.e., $d_{29}(x_1,...,x_{29})$ and $V_{29}(x_1,...,x_{29})$). Using (11), (25), and the maximum expected losses calculated at the final stage of sampling, the maximum expected loss associated with taking one more observation at stage 29 of sampling was computed for $s_{29} = 0,...,29$ (i.e., $E[V_{30}(x_1,...,x_{29},X_{30}) \mid X_1 = x_1,...,X_{29} = x_{29}]$). Comparing these values with $V_{29}(x_1,...,x_{29})$, the appropriate action nonmastery, partial mastery, mastery, or continue sampling was determined at stage 29 of sampling.

In order to compute the appropriate action nonmastery, partial mastery, mastery, or continue sampling at stage 28 of sampling, first the conditional minimax expected loss at stage 29 of sampling was computed (i.e., $R_{29}(x_1,...,x_{29})$) by taking the minimum of $V_{29}(x_1,...,x_{29})$ and $E[V_{30}(x_1,...,x_{29},X_{30}) \mid X_1 = x_1,...,X_{29} = x_{29}]$ for $s_{29} = 0,...,29$. Next, using (12) and (25), the maximum expected losses associated with taking one more observation at stage 28 of sampling, $E[R_{29}(x_1,...,x_{28},X_{29}) \mid X_1 = x_1,...,X_{28} = x_{28}]$ were computed for $s_{28} = 0,...,28$. Comparing these values with $V_{28}(x_1,...,x_{28})$, the appropriate action nonmastery, partial mastery, mastery, or continue sampling was determined at stage 28 of sampling.

Similarly, the appropriate action at stage 27 until stage 0 of sampling was determined. The conditional minimax expected loss at stage 28 of sampling, $R_{28}(x_1,...,x_{28})$, was thereby computed inductively as the minimum of $V_{28}(x_1,...,x_{28})$ and $E[R_{29}(x_1,...,x_{28},X_{29}) \mid X_1 = x_1,...,X_{28} = x_{28}]$ for $s_{28} = 0,...,28$.

Using the recurrent relation $\binom{k+1}{y+1} = \binom{k}{y} + \binom{k}{y+1}$, in combination with $\binom{n}{n} = \binom{n}{0} =$ 1, for computing the binomial coefficients in (8) and (10), a computer program called MINIMAX was developed to determine the appropriate action at each stage of sampling. Furthermore, the parameter α in (25) of the incomplete beta prior was set equal to 10^{-9}. A copy of the program MINIMAX is available from the author upon request.

As can be seen from Table 3, the decision-maker takes at least one observation. Table 3 also shows that at stage k of sampling, a student whose percentage number-correct (i.e., s_k/k) is in the region of t_{c1} or t_{c2} is hard to be classified as a nonmaster, partial master, or master. For such students longer tests are needed, since continue sampling is the appropriate action in this situation. Shorter tests can be provided, however, for students whose percentage number-correct is not in the region of t_{c1} or t_{c2}. The main advantage of variable-length mastery testing, as stated already in the Introduction, is demonstrated here clearly; that is, the test is adapted to the actual level of functioning.

Furthermore, Table 3 shows that continue sampling decisions are taken for the first time after 18 items have been administered in situations where neither clearly partial mastery nor mastery can be declared. Continue sampling decisions are taken at each stage of sampling, however, in situations where neither clearly nonmastery nor partial mastery can be declared.

This finding might be accounted for that the losses associated with taking false nonmastery decisions are twice as large as the losses associated with taking false partial mastery decisions (i.e., 2 and 4 relative to 1 and 2), whereas the differences between the losses associated with taking false partial mastery and mastery decisions (i.e., 1 and 2 relative to 3 and 1) are not that large. Consequently, it seems better to continue sampling if students neither have clearly attained a certain level of nonmastery nor partial mastery in order to avoid relatively large expected losses associated with taking false decisions.

Finally, it can be inferred from Table 3 that with increasing number of items being administered, less continue sampling decisions are taken; that is, the probability of a classification decision increases.

Since at least one observation is taken, the minimax sequential procedure starts with administering one ramdomly selected item and stops after a classification decision (i.e., declaring nonmastery, partial mastery, or mastery). Hence, the minimax sequential procedure proceeds only after the continue sampling decision. Then, it can easily be inferred from Table ? that the minimax sequential decision rule, in case of stopping sampling after a classification decision, can be depicted in Table 4 at each stage of sampling k ($1 \leq k \leq 30$) for different number-correct score s_k ($0 \leq s_k \leq k$) as follows:

Table 4. Minimax Sequential Rule in Case of Stopping Sampling after Classification

Stage of sampling	Minimax Sequential Rule by Number-Correct				
	Nonmastery	Continue	Partial Mastery	Continue	Mastery
1	0	1			
2		[1,2]			
3		[1,3]			
4		[1,3]			4
5	1	[2,4]			
6		[2,5]			
7	2	[3,5]			6
8		[3,6]			
9	3	[4,6]			7
10		[4,7]			
11	4	[5,7]			8
12		[5,8]			
13	5	[6,9]			
14		[6,9]			10
15	6	[7,10]			
16		[7,10]			11
17	7	[8,11]			
18		[8,11]			12
19	8	9	10	[11,12]	
20		[9,10]	11	12	13
21	9	10	11	[12,13]	
22		[10,11]	12	[13,14]	
23	10	11	12	[13,14]	15
24		[11,12]	13	[14,15]	
25	11	12	[13,14]	15	16
26		[12,13]		[15,16]	
27	12	13	[14,15]	16	17
28		13	14 or 16	17	
29	13	14 or 17			18
30	14		15 or 17		18

Note that not all possible number-correct scores s_k are necessarily present at each stage of sampling k, because it is assumed in Table 4 that the minimax sequential rule stops after classifying a student as a nonmaster, partial master, or master. For instance, the number-correct score s_5 can only take the values 1 until 4, and thus, not the values 0 and 5. This is because nonmastery and mastery was already declared for $s_1 = 0$ and $s_4 = 4$, respectively, implying the minimax sequential rule stops for $s_1 = 0$ or $s_4 = 4$. It follows that s_2, s_3, s_4, and s_5 can only take the values [1,2], [1,3], [1,4], and [1,4], respectively.

As can be seen from Table 4, at the final stage of sampling (i.e., stage 30) the number-correct score s_{30} can only take the values 14, 15, 17, or 18. In other words, values larger than 18 and smaller than 14 can not occur for s_{30}, whereas the value of 16 can not occur as well. This is because for all possible values of s_{30} equal to or smaller than 14, students have already been clearly classified as a master, partial master, or nonmaster before the final stage of sampling has been reached. For the possible value of $s_{30} = 16$, students could possibly have been classified already at stage 25 or 28 as a master or partial master, respectively. Possible values of s_{30} larger than 18 can not occur, since students would have clearly attained already a certain level of mastery, partial mastery, or nonmastery for values of s_k ($1 \leq k \leq 30$) equal to or smaller than 18.

Discussion

Optimal rules for the adaptive four-action mastery problem (nonmastery, partial mastery, mastery, and continuing sampling) were derived using the framework of minimax sequential decision theory. The binomial distribution was assumed for the psychometric model involved, whereas threshold loss was adopted for the loss function. The least favorable prior, needed for computing the maximum expected loss associated with continuing sampling, turned out to be the incomplete beta prior with parameter α sufficiently small and parameter β equal to 1. The minimax sequential rules were demonstrated by an empirical example for concept learning in medicine, with a maximum test length of 30.

The results indicated that the chances of being classified as a nonmaster, partial master, or master increased if the number of items administered increased. This result was in accordance with our expectations. Furthermore, the minimax sequential decision rules in case of stopping sampling after a classification decision were computed at each stage of sampling for different number-correct score. It turned out that, for instance, at the final stage of sampling (i.e., after 30 items have been administered) the number-correct score could only take the values of 14, 15, 17, or 18.

It is important to notice that, as pointed out by Veldhuijzen (1982) and Huynh (1980), the minimax principle is very attractive when the only information is the student's observed number-correct score; that is, no group data of 'comparable' students or prior information about the individual student is available. This situation is often the case in individualized instructional programs. If group data of 'comparable' students or prior information about the individual student is available, however, it is better to use this information in the minimax sequential decision procedure regarding declaring mastery, partial mastery, or nonmastery. Hence, in this situation it is better to use Bayesian instead of minimax sequential decision theory. Even if information in the form of group data of 'comparable' students or prior information about the individual student is available, it is sometimes too difficult a job to accomplish to express this information into a prior distribution. In these circumstances, the minimax sequential procedure might also be more appropriate.

Several procedures have been proposed which are simple variants of the minimax strategy (e.g., Coombs, Dawes, & Tversky, 1970), and may also be applied to the adaptive four-action mastery problem. The first is the minimin (complete optimism) strategy, where optimal rules are obtained by minimizing the minimum expected losses associated with all possible decision rules. This strategy is optimal if the best that could happen always happens. As pointed out by Veldhuijzen (1982), the minimin strategy seems rather useless in the context of individualized instructional programs.

The second is the pessismism-optimism (Hurwicz) strategy. This strategy is a combination of the minimax and minimin strategies, and decisions are taken on the basis of the smallest and largest expected losses associated with each possible decision outcome.

The third is the minimax-regret (Savage) strategy. This strategy is similar to the minimax strategy since the focus is on the worst possible decision outcome, but 'worst' is here defined by maximal regret; that is, the difference between the maximal expected loss that was actually obtained, and the maximal expected loss among all possible decision outcomes.

Two final notes are appropriate. First, it might be assumed that guessing has to be taken into account. Huynh (1980) has developed a model with corrections for guessing within a minimax decision-theoretic framework (see also van den Brink & Koele, 1980) for the fixed-length two-action mastery problem (i.e., mastery/nonmastery) in case of a linear loss function. This framework of correction for guessing (i.e., knowledge-or-random-guessing) can be used to derive minimax sequential rules for the adaptive four-action mastery problem.

Second, the minimax sequential decision procedures presented in this paper were applied to the problem of deciding on nonmastery, partial mastery, nonmastery, or continuing sampling in the context of computerized adaptive mastery testing. It should be emphasized, however, that the decision-making procedures advocated here have a larger scope. For instance, the adaptive four-action mastery problem may be important in clinical settings where therapies are followed by mastery tests, which decide on whether or not patients are dismissed from the therapies.

References

Angoff, W.H. (1971). Scales, norms and equivalent scores. In R.L. Thorndike (Ed.), *Educational Measurement*. Washington, D.C.: American Council on Education.

Bejar, I.I., & Braun, H.I. (1994). On the synergy between assessment and instruction: Early lessons from computer-based simulations. *Machine-Mediated Learning*, *4*, 5-25.

Ben-Shakhar, G., & Beller, M. (1983). An application of a decision-theoretic model to a quota-free selection problem. *Journal of Applied Psychology*, *68*, 137-146.

Bork, A. (1984). Producing computer-based learning materials at the Educational Technology Center. *Journal of Computer-Based Instruction*, *11*, 78-81.

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.

Coombs, C.H., Dawes, R.M., Tversky, A. (1970). *Mathematical psychology: An elementary introduction*. Englewood Cliffs, New Jersey: Prentice-Hall Inc.

Cronbach, L.J., & Gleser, G.L. (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: The American College Testing Program.

De Diana, I.P.F., & Vos, H.J. (1988). A lattice representational definition of a hierarchy of instructional processors usable in educational courseware. *Computers & Education*, *12*(3), 427-434.

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

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

Ebel, R.L. (1972). *Essentials of educational measurement*. New York: Prentice Hall.

Ferguson, R.L. (1969a). *Computer-assisted criterion-referenced measurement* (Working Paper No. 41). Pittsburgh PA: University of Pittsburgh Learning and Research Development Center. (Eric Documentation Reproduction No. ED 037 089).

Ferguson, R.L. (1969b). *The development, implementation, and evaluation of a computer-assisted branched test for a program of individually prescribed instruction*. Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh PA.

Nedelsky, L. (1954). Absolute grading standards for objective tests. *Educational and Psychological Measurement*, 14, 3-19.

Novick, M.R., & Lewis, C. (1974). Prescribing test length for criterion-referenced measurement. In C.W. Harris, M.C. Alkin, & W.J. Popham (Eds.), *Problems in criterion-referenced measurement*. CSE monograph series in evaluation, No. 3 Los Angeles: Center for the Study of Evaluation, University of California.

Novick, M.R., & Lindley, D.V. (1979). The use of more realistic utility functions in educational applications. *Journal of Educational Measurement, 15*, 181-191.

Pearson, K. (1930). *Tables for statisticians and biometricians*. London: Cambridge University Press.

Raju, N.S., Steinhaus, S.D., Edwards, J.E., & Lehessio, J. (1991). A logistic regression model for personnel selection. *Applied Psychological Measurement, 15*, 139-152.

Reckase, M.D. (1983). A procedure for decision making using tailored testing. In D.J. Weiss (Ed.), *New horizons in testing: Latent trait test theory and computerized adaptive testing* (pp. 237-257). New York: Academic Press.

Sheehan, K., & Lewis, C. (1992). Computerized mastery testing with nonequivalent testlets. *Applied Psychological Measurement, 16*, 65-76.

Spray, J.A., & Reckase, M.D. (1996). Comparison of SPRT and sequential Bayes procedures for classifying examinees into tow categories using a computerized test. *Journal of Educational and Behavioral Statistics, 21*, 405-414.

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

Tennyson, R.D., & Cocchiarella, M.J. (1986). An empirically based instructional design theory for teaching concepts. *Review of Educational Research, 56*, 40-71.

van den Brink, W.P. (1982). Binomial test models for domain-referenced testing. *Evaluation in Education, 5*, 166-176.

van den Brink, W.P., & Koele, P. (1980). Item sampling, guessing and decision-making in achievement testing. *British Journal of Mathematical and Statistical Psychology, 33*, 104-108.

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

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

Gegg-Harrison, T.S. (1992). Adapting instruction to the student's capabilities. *Journal of Artificial Intelligence in Education, 3*(2), 169-183.

Hambleton, R.K. (1974). Testing and decision-making procedures for selected individualized instructional programs. *Review of Educational Research, 10*, 371-400.

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.

Hansen, D.N., Ross, S.M., & Rakow, E. (1977). *Adaptive models for computer-based training systems* (Annual report to Naval Personnel Research and Development Center). Memphis, TN: Memphis State University.

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

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

Johnson, N.L., & Kotz, S. (1970). *Distributions in statistics: Continuous univariate distributions* -2. Boston: Houghton Miffin.

Kingsbury, G.G., & Weiss, D.J. (1983). A comparison of IRT-based adaptive mastery testing and a sequential mastery testing procedure. In D.J. Weiss (Ed.), *New horizons in testing: Latent trait test theory and computerized adaptive testing* (pp. 257-283). New York: Academic Press.

Kontos, G. (1985). Instructional computing: In search of better methods for the production of CAI lessons. *Journal of Educational Technological Systems, 13*, 121-133.

Koper, R. (1995). PROFIL: A method for the development of multimedia courseware. *British Journal of Educational Technology, 26*, 94-108.

Lehmann, E.L. (1959). Testing statistical theory (3rd ed.). New York: Macmillan.

Lewis, C., & Sheehan, K. (1990). Using Bayesian decision theory to design a computerized mastery test. *Applied Psychological Measurement, 14*, 367-386.

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

Luce, R.D., & Raiffa, H. (1957). *Games and decisions*. New York: John Wiley and Sons.

Millman, J. (1972). *Determining test length, passing scores and test lengths for objectives-based tests*. Los Angeles: Instructional Objectives Exchange.

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. (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*, 129-155. Boston: Kluwer.

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

van der Linden, W.J., & Vos, H.J. (1996). A compensatory approach to optimal selection with mastery scores. *Psychometrika*, *61*, 155-172.

Veldhuijzen, N.H. (1982). Setting cutting scores: A minimum information approach. In W.J. van der Linden (Ed.), Aspects of criterion-referenced measurement. *Evaluation in Education: An International Review Series*, *5*, 141-148.

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

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*, *1*, 313-331. Norwood, New Yersey: Ablex Publishing Corporation.

Vos, H.J. (1994a). Applications of general systems theory to the development of an adjustable tutorial software machine. *Computers & Education*, *22*, 265-276.

Vos, H.J. (1994b). An intelligent tutoring system for classifying students into instructional treatments with mastery scores. In F.J. Maarse, A.E. Akkerman, A.N. Brand, L.J.M. Mulder, & M.J. van der Stelt (Eds.), *Computers in psychology: Applications, methods, and instrumentation* (Vol. 5, pp. 8-18). Lisse, The Netherlands: Swets & Zeitlinger.

Vos, H.J. (1995a). Applications of Bayesian decision theory to intelligent tutoring systems. *Computers in Human Behavior*, *11*, 149-162.

Vos, H.J. (1995b). Adapting the appropriate amount of instruction to learning needs in computerized adaptive instructional systems. *Machine-Mediated Learning*, *5*, 255-268.

Vos, H.J. (1995c). A compensatory model for simultaneously setting cutting scores for selection-placement-mastery decisions. In I. Partchev (Ed.), *Multivariate analysis in the behavioral sciences: Philosophic to technical* (pp. 75-90). Sofia, Bulgaria: Academic Publishing House.

Vos, H.J. (1997a). Adapting the amount of instruction to individual student needs. *Educational Research and Evaluation, 3*, 79-97.

Vos, H.J. (1997b). Simultaneous optimization of quota-restricted selection decisions with mastery scores. *British Journal of Mathematical and Statistical Psychology, 50*, 105-125.

Vos, H.J. (1998a). Instructional decision making procedures for use in intelligent tutoring systems. *Journal of Structural Learning and Intelligent Systems*, to appear.

Vos, H.J. (1998b). A simultaneous approach to optimizing treatment assignments with mastery scores. *Multivariate Behavioral Research*, to appear.

Vos, H.J. (1998c). Optimal classification into two treatments each followed by a mastery decision. In *Proceedings of the Sixth Conference of the International Federation of Classification Societies*, to appear.

Wilcox, R.R. (1977). A note on the length and passing score of a mastery test. *Journal of Educational Statistics, 1*, 359-364.

Wilcox, R.R. (1981). A review of the beta-binomial model and its extensions. *Journal of Educational Statistics, 6*, 3-32.

Acknowledgments

The author is indebted to Wim J. van der Linden for his valuable.

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

RR-97-07	H.J. Vos, *A Minimax Sequential Procedure in the Context of Computerized Adaptive Mastery Testing*
RR-97-06	H.J. Vos, *Applications of Bayesian Decision Theory to Sequential Mastery Testing*
RR-97-05	W.J. van der Linden & Richard M. Luecht, *Observed-Score Equating as a Test Assembly Problem*
RR-97-04	W.J. van der Linden & J.J. Adema, *Simultaneous Assembly of Multiple Test Forms*
RR-97-03	W.J. van der Linden, *Multidimensional Adaptive Yesting with a Minimum Error-Variance Criterion*
RR-97-02	W.J. van der Linden, *A Procedure for Empirical Initialization of Adaptive Testing Algorithms*
RR-97-01	W.J. van der Linden & Lynda M. Reese, *A Model for Optimal Constrained Adaptive Testing*
RR-96-04	C.A.W. Glas & A.A. Béguin, *Appropriateness of IRT Observed Score Equating*
RR-96-03	C.A.W. Glas, *Testing the Generalized Partial Credit Model*
RR-96-02	C.A.W. Glas, *Detection of Differential Item Functioning using Lagrange Multiplier Tests*
RR-96-01	W.J. van der Linden, *Bayesian Item Selection Criteria for Adaptive Testing*
RR-95-03	W.J. van der Linden, *Assembling Tests for the Measurement of Multiple Abilities*
RR-95-02	W.J. van der Linden, *Stochastic Order in Dichotomous Item Response Models for Fixed Tests, Adaptive Tests, or Multiple Abilities*
RR-95-01	W.J. van der Linden, *Some decision theory for course placement*
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-94-13	W.J.J. Veerkamp & M.P.F. Berger, *A simple and fast item selection procedure for adaptive testing*
RR-94-12	R.R. Meijer, *Nonparametric and group-based person-fit statistics: A validity study and an empirical example*
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*
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, Faculty of Educational Science and Technology, University of Twente, Mr. J.M.J. Nelissen, P.O. Box 217, 7500 AE Enschede, The Netherlands.

TM028306

NOTICE

REPRODUCTION BASIS

☑ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").

(9/92)