DOCUMENT RESUME

ED 424 268		TM 029 141
AUTHOR	Veldkamp, Bernard P.	
TITLE	Multidimensional Test Assembly Based on Lagrangian Relaxation Techniques. Research Report 98-08.	
INSTITUTION	Twente Univ., Enschede (Netherlands). Faculty of Educational Science and Technology.	
PUB DATE	1998-00-00	
NOTE	28p.	
AVAILABLE FROM	Faculty of Educational Science and Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.	
PUB TYPE	Reports - Research (143)	
EDRS PRICE	MF01/PC02 Plus Postage.	
DESCRIPTORS	*Ability; Estimation (Mathematics); Foreign Countries; Item Banks; Item Response Theory; Measurement Techniques; Models; Psychometrics; Selection; *Test Construction; Test Length	
IDENTIFIERS	Calibration; *Lagrangian Relaxation; *Multidimensionality (Tests); Test Specifications	

ABSTRACT

In this paper, a mathematical programming approach is presented for the assembly of ability tests measuring multiple traits. The values of the variance functions of the estimators of the traits are minimized, while test specifications are met. The approach is based on Lagrangian relaxation techniques and provides good results for the two dimensional case in a small amount of time. Empirical examples of a test assembly problem from a two dimensional mathematics item pool illustrate the method. In the area of ability measurement, Item Response Theory (IRT) is usually used as a psychometric theory underlying the test assembly process. In this process, three steps can be distinguished. First, an IRT model has to be chosen and the items in the item bank have to be calibrated. From this item bank, many different tests can be assembled. Therefore, the second step consists of specifying the properties of the desired test. One could specify, for example, the test length, the desired amount of information, or the administration time of the test. The third step of the test assembly process is to formulate a model that selects items from the item bank so that test specifications are met. A mathematical programming approach is often used for this step. (Contains five tables, two figures, and nine references.) (Author/SLD)

TM

ED 424 268

Multidimensional Test Assembly Based on Lagrangian Relaxation Techniques

Research Report 98-08

Bernard P. Veldkamp

faculty of
EDUCATIONAL SCIENCE AND TECHNOLOGY

Department of
Educational Measurement and Data Analysis

University of Twente

Multidimensional Test Assembly
Based on Lagrangian Relaxation Techniques

Bernard P. Veldkamp

Abstract

In this paper a mathematical programming approach is presented for the assembly of ability tests measuring multiple traits. The values of the variance functions of the estimators of the traits are minimized, while test specifications are met. The approach is based on Lagrangian relaxation techniques and provides good results for the two dimensional case in a small amount of time. Empirical examples of a test assembly problem from a two dimensional mathematics item pool illustrate the method. Keywords: Lagrangian relaxation, mathematical programming, multidimensional IRT, test assembly, test design. In the area of ability measurement Item Response Theory (IRT) is generally used as a psychometric theory underlying the test assembly process. In this process three steps can be distinguished. First an IRT-model has to be chosen and the items in the itembank have to be calibrated. From this itembank many different tests can be assembled. Therefore, the second step consists of specifying the properties of the desired test. One could specify for example the testlength, the desired amount of information or the administration time of the test. The third step of the test assembly process is to formulate a model that selects items from the itembank so that the test specifications are met. A mathematical programming approach is often used for this step.

Multidimensional Test Assembly based on Lagrangian Relaxation Techniques

In the area of ability measurement Item Response Theory (IRT) is generally used as a psychometric theory underlying the test assembly process. In this process three steps can be distinguished. First an IRT-model has to be chosen and the items in the itembank have to be calibrated. Form this itembank many different tests can be assembled. Therefore, the second step consists of specifying the properties of the desired test. One could specify for example the testlength, the desired amount of information or the administration time of the test. The third step of the test assembly process is to formulate a model that selects items from the itembank so that the test specifications are met. A mathematical programming approach is often used for this step.

The idea of using a mathematical programming approach was first suggested by Yen (1983). Ever since, a number of researchers have shown various LP algorithms and heuristics to solve test assembly problems. Still much research is done in this field of interest. Recently van der Linden (1996) published a paper on the subject of assembling tests measuring multiple traits. Although an algorithm to solve the assembling problem was provided, manual correction was needed to find the optimal solution. Therefore, the purpose of this study was to find a heuristic to solve the problem of assembling tests measuring multiple traits. An approach based on Lagrangian relaxation was used.

A Linear Logistic Multidimensional IRT Model

The IRT model

The model considered in this paper is a generalization of the two-parameter logistic model Lord (1980) to the multidimensional case and can be formulated in the following manner:

$$\begin{aligned} P_i(\boldsymbol{\theta}_j) &\equiv P(U_{ij}=1 \mid (\mathbf{a}_i, d_i, \boldsymbol{\theta}_j)) \\ &\equiv \frac{e^{(\mathbf{a}_i\cdot\boldsymbol{\theta}_j+d_i)}}{1+e^{(\mathbf{a}_i\cdot\boldsymbol{\theta}_j+d_i)}}, \end{aligned} \tag{1}$$

where U_{ij} represents the response of person $j = 1 \dots N$ to item $i = 1 \dots n$, $\mathbf{a}_i$ is the vector of discrimination parameters of item i along the abilities $\theta_{j1} \dots \theta_{jm}$, m is the dimensionality of the ability space, and d_i is the parameter representing the difficulty of the item. $P(U_{ij} = 1|(\mathbf{a}_i, d_i, \boldsymbol{\theta}_j))$ is the probability of a correct response (score of 1) for person j on test item i, and

$\mathbf{a}_i \cdot \boldsymbol{\theta}_j$ is the inner product of the vector of discrimination parameters of item i and the ability vector $\boldsymbol{\theta}_j$ for person j. The item parameters are supposed to be known and the model is used to estimate the ability vectors $\boldsymbol{\theta}_j$ from a realization of the response variables $U_{ij} = u_{ij}$ for $i = 1 \ldots n$ and $j = 1 \ldots N$.

This multidimensional IRT model was also described by Reckase (1997). In this article several assumptions underlying this model were presented. First of all, the monotonicity assumption implies that, with an increase in the ability of the examinee, the probability of obtaining a correct response to a test item is non-decreasing. Second, the probability of obtaining a vector of responses is equal to the product of the probabilities of obtaining these responses. This assumption is called the local independence assumption. The third and last assumption deals with the derivatives of the function describing the model. These derivatives have to be properly defined.

Variance functions

Fisher's information matrix is defined as:

$$I(\boldsymbol{\theta}) \equiv -E \begin{bmatrix} \frac{\partial^2 \ln L}{\partial \theta_1^2} & \cdots & \frac{\partial^2 \ln L}{\partial \theta_1 \partial \theta_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \ln L}{\partial \theta_m \partial \theta_1} & \cdots & \frac{\partial^2 \ln L}{\partial \theta_m^2} \end{bmatrix}, \tag{2}$$

were the likelihood equation L for the model in Equation 1 is given by:

$$L = \prod_{i=1}^{n} P(u_{ij} | (\mathbf{a}_i, d_i, \boldsymbol{\theta})). \tag{3}$$

According to van der Linden (1996) and using the notation $P_i \equiv P_i(\boldsymbol{\theta})$ and $Q_i \equiv 1 - P_i$, Fisher's information matrix for this model can be formulated as:

$$I(\boldsymbol{\theta}) = \begin{bmatrix} \sum_{i=1}^{n} a_{1i}^2 P_i Q_i & \cdots & \sum_{i=1}^{n} a_{1i} a_{mi} P_i Q_i \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} a_{1i} a_{mi} P_i Q_i & \cdots & \sum_{i=1}^{n} a_{mi}^2 P_i Q_i \end{bmatrix}. \tag{4}$$

Because Fisher's information function is a matrix instead of a scalar in the multiparameter case, the (asymptotic) variances of the MLEs of the ability parameters $\theta_1, .., \theta_m$ are not given by the reciprocals of the diagonal elements of the information matrix, but by the diagonal elements of the variance-covariance matrix. This matrix is the inverse of Fisher's information matrix. For notational simplicity we consider the case that $m = 2$. In this case the inverse matrix is of the following form:

$$V(\hat{\boldsymbol{\theta}}|\boldsymbol{\theta}) = \frac{1}{|I(\boldsymbol{\theta})|} \begin{bmatrix} \sum_{i=1}^{n} a_{2i}^2 P_i Q_i & -\sum_{i=1}^{n} a_{1i} a_{2i} P_i Q_i \\ -\sum_{i=1}^{n} a_{1i} a_{2i} P_i Q_i & \sum_{i=1}^{n} a_{1i}^2 P_i Q_i \end{bmatrix}, \tag{5}$$

where $|I(\boldsymbol{\theta})|$ is the determinant of the matrix in Equation 4. It should be noted that the variance of $\hat{\theta}_1$ not only depends on θ_1 but also on the values of θ_2. The variance functions $Var(\hat{\theta}_k|\boldsymbol{\theta}) = \sum_{i=1}^{n} a_{ki}^2 P_i Q_i / |I(\boldsymbol{\theta})|$ simplify to:

$$Var(\hat{\theta}_1|\boldsymbol{\theta}) = (\sum_{i=1}^{n} a_{2i}^2 P_i Q_i) \left[\left(\sum_{i=1}^{n} a_{1i}^2 P_i Q_i \right) \left(\sum_{i=1}^{n} a_{2i}^2 P_i Q_i \right) - \left(\sum_{i=1}^{n} a_{1i} a_{2i} P_i Q_i \right)^2 \right]^{-1}, \tag{6}$$

and

$$Var(\hat{\theta}_2|\boldsymbol{\theta}) = (\sum_{i=1}^{n} a_{1i}^2 P_i Q_i) \left[\left(\sum_{i=1}^{n} a_{1i}^2 P_i Q_i \right) \left(\sum_{i=1}^{n} a_{2i}^2 P_i Q_i \right) - \left(\sum_{i=1}^{n} a_{1i} a_{2i} P_i Q_i \right)^2 \right]^{-1}. \tag{7}$$

The objective of test assembly problems is to minimize the values of these variance functions. Therefore a multidimensional maximin model was formulated.

A Multidimensional Maximin Model

Introducing a decision variable x_i, $i = 1, \ldots, I$,where $x_i = 1$ if item i is in the test, $x_i = 0$ if item i is not in the test and I is the number of items in the itembank, Equation 6 and Equation 7 can be rewritten such that:

$$Var(\hat{\theta}_1|\boldsymbol{\theta}) = \left[\sum_{i=1}^{I} a_{2i}^2 P_i Q_i x_i - \left(\sum_{i=1}^{I} a_{1i} a_{2i} P_i Q_i x_i \right)^2 / \left(\sum_{i=1}^{I} a_{1i}^2 P_i Q_i x_i \right) \right]^{-1}, \tag{8}$$

and

$$Var(\hat{\theta}_2|\boldsymbol{\theta}) = \left[\sum_{i=1}^{I} a_{1i}^2 P_i Q_i x_i - \left(\sum_{i=1}^{I} a_{1i} a_{2i} P_i Q_i x_i \right)^2 / \left(\sum_{i=1}^{I} a_{2i}^2 P_i Q_i x_i \right) \right]^{-1}. \tag{9}$$

Objective function

The objective of the test assembling problem is to minimize the values of both variance functions. This problem is in the class of multi objective decision problems. Problems in this class can be solved in several ways, see also Hwang and Masud (1979). In this article an approach based on combining both functions into one objective function was chosen. To find a good combination of both variance functions the effect of the sums on the variance functions was investigated. About the variance functions the following observations can be made. From Equation 8 it can be derived that the value of $Var(\hat{\theta}_1|\boldsymbol{\theta})$ for a given $\boldsymbol{\theta}$ decreases when

$$\sum_{i=1}^{I} a_{2i}^2 P_i Q_i x_i \tag{10}$$

increases in value,

$$\left(\sum_{i=1}^{I} a_{1i} a_{2i} P_i Q_i x_i\right)^2 \tag{11}$$

decreases in value and

$$\sum_{i=1}^{I} a_{1i}^2 P_i Q_i x_i \tag{12}$$

increases in value.

It should also be noticed that the value of $Var(\hat{\theta}_2|\boldsymbol{\theta})$ for a given $\boldsymbol{\theta}$ decreases when

$$\sum_{i=1}^{I} a_{1i}^2 P_i Q_i x_i \tag{13}$$

increases in value,

$$\left(\sum_{i=1}^{I} a_{1i} a_{2i} P_i Q_i x_i\right)^2 \tag{14}$$

decreases in value and

$$\sum_{i=1}^{I} a_{2i}^2 P_i Q_i x_i \tag{15}$$

increases in value. So, both variance functions show the same behaviour for these sums.

Therefore, it seems reasonable to choose the function

$$\frac{\left(\sum_{i=1}^{I} a_{1i}a_{2i}P_iQ_ix_i\right)^2}{\sum_{i=1}^{I} a_{1i}^2 P_iQ_ix_i + \sum_{i=1}^{I} a_{2i}^2 P_iQ_ix_i}, \tag{16}$$

which shows the same behaviour, as an objective function.

The model

The objective function defined in Equation 16 is a continuous function of the variables (θ_1, θ_2). However, in the test assembling process it suffices to optimize this objective function for a grid of points. Therefore, the following mathematical programming problem has been chosen, which minimizes the maximum value of a combination of the variance functions for a set of θ-points $(\theta_{1p}, \theta_{2q})$, $p = 1, \ldots, P$ and $q = 1, \ldots, Q$:

$$\min \max_{\substack{s=1\ldots S\\ t=1\ldots T}} \frac{\left(\sum_{i=1}^{I} a_{1i}a_{2i}P_i(\theta_{1s},\theta_{2t})Q_i(\theta_{1s},\theta_{2t})x_i\right)^2}{\left(\sum_{i=1}^{I} a_{1i}^2 P_i(\theta_{1s},\theta_{2t})Q_i(\theta_{1s},\theta_{2t})x_i\right) + \frac{1}{c}\left(\sum_{i=1}^{I} a_{2i}^2 P_i(\theta_{1s},\theta_{2t})Q_i(\theta_{1s},\theta_{2t})x_i\right)}, \tag{17}$$

subject to

$$\sum_{i=1}^{I} x_i = n, \tag{18}$$

$$x_i \in \{0,1\}, \qquad i = 1 \ldots I, \tag{19}$$

where c in Equation 17 is a weighting factor. By varying this weighting factor one can influence the relative importance of the two sums in the denominator. Equation 18 specifies the testlength, but, of course, other constraints can be added to this problem. Unfortunately, the objective function, Equation 17, is a complicated function of the variables x_i. By creating three new variables y, K_1 and K_2 the problem can be restated as:

$$\min \frac{y^2}{K_1 + K_2}, \tag{20}$$

subject to

$$\sum_{i=1}^{I} a_{1i}a_{2i}P_i(\theta_{1s},\theta_{2t})Q_i(\theta_{1s},\theta_{2t})x_i - y \leq 0, \qquad s=1\ldots S, t=1\ldots T, \tag{21}$$

$$\sum_{i=1}^{I} a_{1i}^2 P_i(\theta_{1s},\theta_{2t})Q_i(\theta_{1s},\theta_{2t})x_i \geq K_1, \qquad s=1\ldots S, t=1\ldots T, \tag{22}$$

$$\sum_{i=1}^{I} a_{2i}^2 P_i(\theta_{1s},\theta_{2t})Q_i(\theta_{1s},\theta_{2t})x_i \geq cK_2, \qquad s=1\ldots S, t=1\ldots T, \tag{23}$$

$$\sum_{i=1}^{I} x_i = n, \tag{24}$$

$$x_i \in \{0,1\}, \qquad i=1\ldots n, \tag{25}$$

$$y, K_1, K_2 > 0. \tag{26}$$

The constraints, Equations 21, 22 and 23, must be satisfied as equalities at optimality.

Other models

The following multidimensional maximin model to solve this problem, was described in van der Linden (1996):

$$\min\ y \tag{27}$$

subject to

$$\sum_{i=1}^{I} a_{1i}a_{2i}P_i(\theta_{1s},\theta_{2t})Q_i(\theta_{1s},\theta_{2t})x_i - y \leq 0, \qquad s=1\ldots S, t=1\ldots T, \tag{28}$$

$$\sum_{i=1}^{I} a_{1i}^2 P_i(\theta_{1s}, \theta_{2t}) Q_i(\theta_{1s}, \theta_{2t}) x_i \geq K_1, \qquad s = 1 \ldots S, t = 1 \ldots T, \tag{29}$$

$$\sum_{i=1}^{I} a_{2i}^2 P_i(\theta_{1s}, \theta_{2t}) Q_i(\theta_{1s}, \theta_{2t}) x_i \geq K_2, \qquad s = 1 \ldots S, t = 1 \ldots T, \tag{30}$$

$$\sum_{i=1}^{I} x_i = n, \tag{31}$$

$$x_i \in \{0, 1\}, \qquad i = 1 \ldots I, \tag{32}$$

$$y \geq 0. \tag{33}$$

The basic idea is to run this model, while the variables K_1 and K_2 are systematically varied, until the optimal solution is found. The main difference between this model and the model described by Equations 20 until 26 is that the objective function, Equation 20, is linearized and that this problem has to be solved many times for many combinations of (K_1, K_2). Furthermore, the factor c is not necessary, because the ratio of K_1 and K_2 can be used as a weighting factor.

Several more models can be formulated that solve the same problem. However, the purpose of this study is to find a heuristic which provides good results in a small amount of time. Therefore it might be better not to eliminate the variables and to look for a different way to handle this problem. In this study an approach based on Lagrangian Relaxation was chosen.

Application of Lagrangian Relaxation to this Problem

The problem stated above is in the class of non-linear mixed integer programming problems. The problem is NP-hard, which means that it can not be solved in polynomial time. In Armstrong, Jones and Wang (1995) an approach, based on Lagrangian Relaxation techniques, is described to solve various test assembly problems.

Lagrangian Relaxation techniques are based on the idea that a good lower or upper bound to the solution of a mathematical programming problem can be obtained by relaxing some constraints. These constraints are not just thrown away, but are brought into the objective function by a correction term. This correction term is a linear combination of the constraints. The remaining relaxed problem is solved several times using various linear combinations of the constraints to obtain a good approximate solution to the original problem. The main difficulty in this technique is the calculation of the correction term.

Notational simplification

Before we try to use this approach for our problem we introduce the following definitions for notational simplicity:

$$A_{ij} \equiv \begin{bmatrix} -a_{11}a_{21}P_{1j}(\theta_{11j},\theta_{21j})Q_{1j}(\theta_{11j},\theta_{21j}) & \cdots & -a_{1n}a_{2n}P_{nj}(\theta_{11j},\theta_{21j})Q_{nj}(\theta_{11j},\theta_{21j}) \\ \vdots & \ddots & \vdots \\ -a_{11}a_{21}P_{1j}(\theta_{1Sj},\theta_{2Tj})Q_{1j}(\theta_{1Sj},\theta_{2Tj}) & \cdots & -a_{1n}a_{2n}P_{nj}(\theta_{1Sj},\theta_{2Tj})Q_{nj}(\theta_{1Sj},\theta_{2Tj}) \\ a_{11}^2P_{1j}(\theta_{11j},\theta_{21j})Q_{1j}(\theta_{11j},\theta_{21j}) & \cdots & a_{1n}^2P_{nj}(\theta_{11j},\theta_{21j})Q_{nj}(\theta_{11j},\theta_{21j}) \\ \vdots & \ddots & \vdots \\ a_{11}^2P_{1j}(\theta_{1Sj},\theta_{2Tj})Q_{1j}(\theta_{1Sj},\theta_{2Tj}) & \cdots & a_{1n}^2P_{nj}(\theta_{1Sj},\theta_{2Tj})Q_{nj}(\theta_{1Sj},\theta_{2Tj}) \\ a_{21}^2P_{1j}(\theta_{11j},\theta_{21j})Q_{1j}(\theta_{11j},\theta_{21j}) & \cdots & a_{2n}^2P_{nj}(\theta_{11j},\theta_{21j})Q_{nj}(\theta_{11j},\theta_{21j}) \\ \vdots & \ddots & \vdots \\ a_{21}^2P_{1j}(\theta_{1Sj},\theta_{2Tj})Q_{1j}(\theta_{1Sj},\theta_{2Tj}) & \cdots & a_{2n}^2P_{nj}(\theta_{1Sj},\theta_{2Tj})Q_{nj}(\theta_{1Sj},\theta_{2Tj}) \end{bmatrix}, \tag{34}$$

were $j = 1, .., 3ST$ and $i = 1, .., n$, and

$$\mathbf{b}_{y,K_1,K_2} \equiv [-y, \cdots, -y, K_1, \cdots, K_1, cK_2, \cdots, cK_2]. \tag{35}$$

So, the constraint $A\mathbf{x} \geq \mathbf{b}_{y,K_1,K_2}$ is equivalent to the constraints defined in Equations 28 until 30. Using this notation, the maximin model simplifies to:

$$\min \frac{y^2}{K_1 + K_2}, \tag{36}$$

subject to

$$A\mathbf{x} \geq \mathbf{b}_{y,K_1,K_2}, \tag{37}$$

$$\sum_{i=1}^{n} x_i = n, \tag{38}$$

$$x_i \in \{0,1\}, \qquad i = 1 \ldots n, \tag{39}$$

$$K_1, K_{2,}y > 0. \tag{40}$$

Lagrangian Relaxation

A Lagrangian Relaxation of this problem can be obtained by relaxing Equation 37, i.e. by relaxing the constraints due to a rewriting of the objective function. Other constraints, e.g. the constraint specifying the test length, may not be relaxed. Otherwise, the test would no longer meet its specifications. The resulting Lagrangian Relaxation problem can be stated as:

$$\min \frac{y^2}{K_1 + K_2} + \mathbf{u}^T(\mathbf{b}_{y,K_1,K_2} - A\mathbf{x}) \tag{41}$$

subject to

$$\sum_{i=1}^{I} x_i = n, \tag{42}$$

$$x_i \in \{0,1\}, \qquad i = 1 \ldots I, \tag{43}$$

$$u_j, K_1, K_{2,}y > 0, \qquad j = 1 \ldots 3ST. \tag{44}$$

The elements of **u** are called Lagrangian multipliers.

Some remarks can be made about this mathematical programming problem. Looking closer at Equation 41, it can be observed that the objective function can be split in two independent parts. Consider Equation 41:

$$\min \frac{y^2}{K_1 + K_2} + \mathbf{u}^T(\mathbf{b} - A\mathbf{x}) \tag{45}$$

where the values of vector $\mathbf{b}$ depends on y, K_1 and K_2. Equation 41 can be rewritten:

$$\min \frac{y^2}{K_1 + K_2} + \mathbf{u}^T\mathbf{b} - \mathbf{u}^T A\mathbf{x} \tag{46}$$

$\Longleftrightarrow$

$$\min \left(\frac{y^2}{K_1 + K_2} + \mathbf{u}^T\mathbf{b} \right) - \left(\mathbf{u}^T A\mathbf{x} \right) \tag{47}$$

$\Longleftrightarrow$

$$\min f_{\mathbf{u}}\left(y, K_1, K_2\right) - g_{\mathbf{u}}\left(\mathbf{x}\right) \tag{48}$$

where $y, K_1, K_2 > 0$, and the conditions defined by Equation 42 and Equation 43 have to be met. So, the functions $f_{\mathbf{u}}$ and $g_{\mathbf{u}}$ can be minimized separately. The result of this separation is that the problem is split in a continuous non-linear optimization part and a zero-one linear programming part. The remaining problem is to find a good $\mathbf{u}$ vector.

Subgradient optimization

A general approach to the problem of finding a good $\mathbf{u}$-vector is subgradient optimization. This technique is based on two observations concerning the Lagrangian relaxation of a mathematical programming problem. If the following problem has to be solved:

$$z_P = \min \mathbf{c}^T\mathbf{x} \tag{49}$$

subject to

$$\begin{aligned} D\mathbf{x} &\leq \mathbf{e}, \\ F\mathbf{x} &\leq \mathbf{g}, \end{aligned} \tag{50}$$

where $D\mathbf{x} \leq \mathbf{e}$ defines a set of constraints and $F\mathbf{x} \leq \mathbf{g}$ defines a set of constraints, a Lagrangian relaxation of this problem is:

$$z_D(\mathbf{u}) = \min\{\mathbf{c}^T\mathbf{x} + \mathbf{u}^T(D\mathbf{x} - \mathbf{e})\} \tag{51}$$

subject to

$$
\begin{aligned}
F\mathbf{x} &\leq \mathbf{g}, \\
u_j &\geq 0.
\end{aligned} \tag{52}
$$

If the equation $F\mathbf{x} \leq \mathbf{g}$ defines a finite set of points the Lagrangian relaxation can be written as:

$$z_D(\mathbf{u}) = \min_{t=1,\ldots T_{\max}} \{\mathbf{c}^T\mathbf{x}_t + \mathbf{u}^T(D\mathbf{x}_t - \mathbf{e})\} \tag{53}$$

and the following observations can be made. Notice that $z_D(\mathbf{u})$ minimizes a finite number of linear, therefore concave functions of $\mathbf{u}$. Because of this, there exists a subgradient of $z_D(\mathbf{u})$ for every $\mathbf{u}$, see also Definition 1. Further, there exists at least one $\mathbf{u}^*$ which maximizes $z_D(\mathbf{u})$. For a more detailed description of these observations see also Dirickx, Baas and Dorhout (1987).

Definition 1 *A vector* $\mathbf{r} \in R^m$ *is a subgradient of a concave function* $f(\mathbf{u})$ *in* $\overline{\mathbf{u}} \in R^m$, *if for every* $\mathbf{u} \in R^m$

$$f(\mathbf{u}) \leq f(\overline{\mathbf{u}}) + \mathbf{r}^T(\mathbf{u} - \overline{\mathbf{u}}).$$

Application of subgradient optimization

Applying this technique to the problem stated in Equation 36 until Equation 40, equation

$$A\mathbf{x} \geq \mathbf{b}_{y,K_1,K_2} \tag{54}$$

can be seen as $D\mathbf{x} \leq \mathbf{e}$ and the equations

$$\sum_{i=1}^{I} x_i = n, \tag{55}$$

$$x_i \in \{0,1\}, \qquad i = 1 \ldots I \tag{56}$$

$$u_j, K_1, K_{2,}y > 0, \qquad j = 1 \ldots 3ST \tag{57}$$

can be seen as $F\mathbf{x} \leq \mathbf{g}$.

To check if the subgradient optimization technique can be applied to the problem described in Equations 41 until 44 we have to ensure that all conditions are met. First, the objective function of the relaxation has to be a linear function of the variable $\mathbf{u}$. From Equation 41, it can be derived that the vector $\mathbf{u}$ is only involved in an inner product, so this condition is met. Second, the feasible region has to be a finite set of points. The feasible region is defined by Equations 42 until 44. Unfortunately, K_1, K_2 and y are continuous variables. However, as a result of splitting the optimization problem, the optimal values of y, $K_1$and K_2 can be computed separately and these variables can be fixed at their optimal values, so the feasible region is a finite set of points. Now all conditions are met and a description of the procedure can be given.

The following procedure outlines the search for a good $\mathbf{u}$-vector:

- Step 0. Set $k = 0$, let $\widehat{z}$ be an upper bound to $z_D(\mathbf{u})$, and create an initial vector $\mathbf{u}_0$.
- Step 1. Calculate the subgradient $\mathbf{r}_k$ by solving the mathematical programming problem

$$\min \left\{ \frac{y^2}{K_1 + K_2} + \mathbf{u}^T(\mathbf{b} - A\mathbf{x}) | \sum_{i=1}^{I} x_i = n, y \geq 0, K_1 \geq 0, K_2 \geq 0 \right\} \tag{58}$$

If $\overline{y}, \overline{K}_1, \overline{K}_2, \overline{\mathbf{x}}$ is the optimal solution to this problem, define $\mathbf{r}_k = A\overline{\mathbf{x}}_k - \mathbf{b}(\overline{y},\overline{K}_1,\overline{K}_2)$.

- Step 2. $t_k = \mu_k \frac{\widehat{z} - z_D(\mathbf{u}_k)}{\|A\mathbf{x}_k - \mathbf{b}\|^2}$, where $0 \leq \mu_k \leq 2$.
- Step 3. $\mathbf{u}_{k+1} = \mathbf{u}_k + t_k \cdot \mathbf{r}_k$.
- Step 4. $k := k + 1$.
- Step 5. If $k \leq 75$ go to step 1, else go to step 6.
- Step 6. End of the $\mathbf{u}$-search, set $\mathbf{u} = \mathbf{u}_{75}$.

The reason for terminating this procedure when $k > 75$ is that the values of the elements of the correction vector $\mathbf{u}$ strongly depend on the parameter μ_k, which, in turn, depends on the stopping criterium $k = 75$. The larger the stopping criterium the more elements of $\mathbf{u}$ equal to zero and the higher the values of the remaining elements. In general, the value of the stopping criterium should be in the interval [40..125], and $k = 75$ seems a good choice. For a detailed description and validation of this subgradient procedure see Dirickx, Baas and Dorhout (1987) and Held, Wolfe and Crowder (1974).

Empirical Examples

Two examples are presented. First, the LR-problem stated in Equation 41 till Equation 44 was solved for data from an ACT Assessment Program Mathematics Item Pool. Second, the same problem was solved for the same data, but with an additional set of constraints.

The item pool consisted of 176 items, to which an acceptable fit was shown by a two dimensional version of the model in Equation 1. The items were classified according to content specification and skill.

The above described method was used to solve the models on a PC with a Pentium processor (130 mHz). Heuristic seven of the program Contest was used to solve the zero-one linear programming parts of the second example.

Example 1

It was attempted to assemble a test for both ability estimators over the complete grid of points defined by $\theta_1, \theta_2 = -1, 0, 1$. The test should contain 25 items. The upper bound to the Lagrangian relaxation $z_D(\mathbf{u})$ was set equal to $\widehat{z} = 3$ and the weighting factor μ_k in Step 2. of the above described algorithm was set equal to 0.2 and halved every $\left(\frac{\text{stopping criterium}}{10}\right) \approx 7$ iterations. The values of the upper bound and the weighting factor depend on the problem and should be found empirically. The resulting variance functions are shown in Figure 1.

To test the robustness of the above described method attempts were made to assemble tests for several initial vectors $\mathbf{u}_0$. The mean values (μ_1, μ_2) and the standard deviations (σ_1, σ_2) of the values of both variance functions over nine points of the grid were used as a measure to compare the results. Table 1 shows the results and the CPU-time for several starting values $\mathbf{u}_0$.

Table 1

Results for starting values $\mathbf{u}_0$ and $k = 75$

$\mathbf{u}_0$	μ_1	μ_2	σ_1	σ_2	μ	σ	$\mu + \sigma$	time (s)
0.05	1.081	0.889	0.153	0.188	0.985	0.341	1.326	4.73
0.10	1.089	0.858	0.146	0.182	0.974	0.328	1.302	4.89
0.15	1.135	0.876	0.157	0.180	0.989	0.337	1.326	4.94
0.20	1.157	0.888	0.147	0.187	1.023	0.334	1.357	4.89
0.25	1.234	0.846	0.163	0.177	1.040	0.340	1.380	5.05
0.30	1.275	0.845	0.184	0.179	1.060	0.363	1.423	4.94

The CPU-time needed to find a solution is only five seconds, which shows that this method is rather quick. Because of small differences in the resulting $\mu + \sigma$, the method seems rather robust to small modifications of the starting values $\mathbf{u}_0$. The method performs best when $0.05 \leq \mathbf{u}_0 \leq 0.15$. To examine the stopping criterium $(k = 75)$ the mean values and standard deviations of the values of both variance functions were compared for several values of k. The results are shown in Table 2.

Table 2

Results for several values of k and starting values $u_0 = 0.15$

	μ	σ	$\mu+\sigma$
$k=50$	1.020	0.365	1.385
$k=75$	0.989	0.337	1.326
$k=75$	0.992	0.375	1.367

It can be concluded that the value of $\mu+\sigma$ does not change much for $k > 50$, although $k = 75$ performs best. On the other hand, if attention is paid to the values of the correction vector u, it can be seen that many violations of the relaxed constraints are corrected by a small term if $k = 50$. On the other hand, if $k = 100$, a few violations are corrected by a big term. Therefore it seems reasonable to choose $k = 75$, which is somewhere in between, as a stopping criterium.

Example 2

Both content and skill constraints were added to the LR-problem described in Equation 41 till Equation 44:

(1) The test should contain at least 2 plane geometry, 2 pre-algebra, 2 elementary algebra, 2 coordinate geometry, 2 trigonometry, and 2 intermediate algebra items.

(2) At least 7 basic skill items, 7 application items, and 2 analysis items should be included in the test.

This results in the following additional constraints:

$$\begin{aligned}
\sum_{i\in V_{PG}} x_i &\geq 2, \\
\sum_{i\in V_{PA}} x_i &\geq 2, \\
\sum_{i\in V_{EA}} x_i &\geq 2, \\
\sum_{i\in V_{CG}} x_i &\geq 2, \\
\sum_{i\in V_{TG}} x_i &\geq 2, \\
\sum_{i\in V_{IA}} x_i &\geq 2,
\end{aligned} \tag{59}$$

$$\sum_{i \in V_{BA}} x_i \geq 7,$$
$$\sum_{i \in V_{AP}} x_i \geq 7,$$
$$\sum_{i \in V_{AN}} x_i \geq 2,$$

where for example V_{PG} is the set indices of the items with content classification plane geometry (PG). The same specifications as in example 1 were used. The resulting variance functions are shown in Figure 2. To compare the results a description of the results of this example is given in Table 3.

Table 3

Results for starting values $\mathbf{u}_0$ and $k = 75$

$\mathbf{u}_0$	μ_1	μ_2	σ_1	σ_2	μ	σ	$\mu + \sigma$	time (s)
0.05	1.099	1.283	0.154	0.208	1.191	0.362	1.553	23.12
0.10	1.094	1.415	0.150	0.230	1.254	0.380	1.634	23.56
0.15	1.094	1.415	0.150	0.230	1.254	0.380	1.634	23.57
0.20	1.158	1.337	0.168	0.214	1.248	0.382	1.630	23.89
0.25	1.158	1.337	0.168	0.214	1.248	0.382	1.630	22.52
0.30	1.158	1.337	0.168	0.214	1.248	0.382	1.630	22.63

As can be seen, the values of $\mu + \sigma$ increase in comparison to the values of $\mu + \sigma$ in example 1. This is because of the additional constraints. Because only small differences in the values of $\mu + \sigma$ occur, the method is rather robust to a little modification of the initial values $\mathbf{u}_0$ even if there are additional constraints. The CPU-time to find a solution is about twenty three seconds. Most of this time is needed to switch to Contest for solving the linear programming part of the problem. However, twenty three seconds is still pretty fast for assembling a test measuring multiple traits. To examine the stopping criterium ($k = 75$), the results for several stopping criteria are showed in Table 4.

Table 4

Results for several values of k and starting values $\mathbf{u}_0 = 0.15$

	μ	σ	$\mu + \sigma$
$k = 50$	1.302	0.438	1.740
$k = 75$	1.255	0.373	1.628
$k = 100$	1.256	0.379	1.635

As can be seen, for this example the best results were obtained for $k = 75$. And choosing $k = 75$ as an stopping criterium seems a good choice.

In both examples no attention was paid to the differences in the shapes of the resulting variance functions. However, if it is important that both resulting variance functions are of the same shape one could use the parameter c in Equation 23 as a stabilizing factor. For the problem stated in Example 1, the results for various values of the factor c are shown in Table 5. The shapes of both variance functions are represented by the values of $\mu_1 + \sigma_1$ and $\mu_2 + \sigma_2$. The value of $\mu_1 + \sigma_1$, decreases when the value of c decreases and the value of $\mu_2 + \sigma_2$ increases when the value of c decreases. The difference between both shapes will be minimal for $c \approx 0.98$. However, it has to be emphasized that the optimal value of c is conditional on both the problem and the itembank.

Table 5

Effect of the stabilizing value c

c	μ_1	μ_2	σ_1	σ_2	$\mu_1 + \sigma_1$	$\mu_2 + \sigma_2$
1.25	1.623	0.742	0.227	0.167	1.850	0.909
1.10	1.218	0.799	0.160	0.196	1.378	0.995
1.05	1.135	0.876	0.156	0.220	1.291	1.096
1.00	1.135	0.876	0.156	0.220	1.291	1.096
0.95	0.977	1.045	0.129	0.276	1.106	1.321
0.90	0.918	1.181	0.116	0.337	1.034	1.518
0.75	0.918	1.181	0.116	0.337	1.034	1.518

Conclusion

Although it seems quite obvious to use the variance functions as a base for assembling tests measuring multiple traits, the remaining mathematical programming problems is hard to solve. In this paper a method is presented which offers the possibility of computing a good upper bound to this problem. Especially the opportunity of splitting the nonlinear programming problem in two problems which are easy to solve, simplifies the calculation process.

Therefore, it can be concluded that an approach based on Lagrangian relaxation is a good method for solving the problem of assembling tests measuring multiple traits. On the other hand a rather general approach to find a good u vector was used and a more specific algorithm might further reduce the CPU-time and improve the test assembling process.

References

Armstrong, R.D., Jones, D.H., & Wang, Z. (1995). *Network optimization in constrained standardized test construction* Applications of Management Science, 8, 189-212.

Dirickx, Y.M.I., Baas, S.M., & Dorhout, B. (1987), *Operationele Research.* Den Haag, The Netherlands: Academic Service.

Held, H., Wolfe, P., & Crowder, H.P. (1974), *Validation of subgradient optimization.* Mathematical Programming, 6, 62-88.

Hwang, C-L and Masud, A.S.Md. (1979), Multiple Objective Decision Making - Methods and Applications. Heidelberg: Springer Verlag.

van der Linden, W.J. (1996). *Assembling Tests for the Measurement of Multiple Abilities.* Applied Psychological Measurement, 20, 373-388.

Lord, F.M. (1980). *Applications of item response theory to practical testing problems* (13th ed.). Hillsdale, NJ: Erlbaum Associates.

Reckase, M.D. (1997). A linear logistic multidimensional model for dichotomous item response data. In W.J. van der Linden & R.K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 271-286). New York, NY: Springer-Verlag.

Wagner, H.M. (1969). *Principles of Operations Research.* Englewood Cliffs, NJ: Prentice-Hall Inc.

Yen, W.M. (1983). Use of the three-parameter model in the development of standardized achievement tests. In R. K. Hambleton (Ed.), *Applications of item response theory* (pp. 123-141). Vancouver: Educational Research Institute of British Columbia.

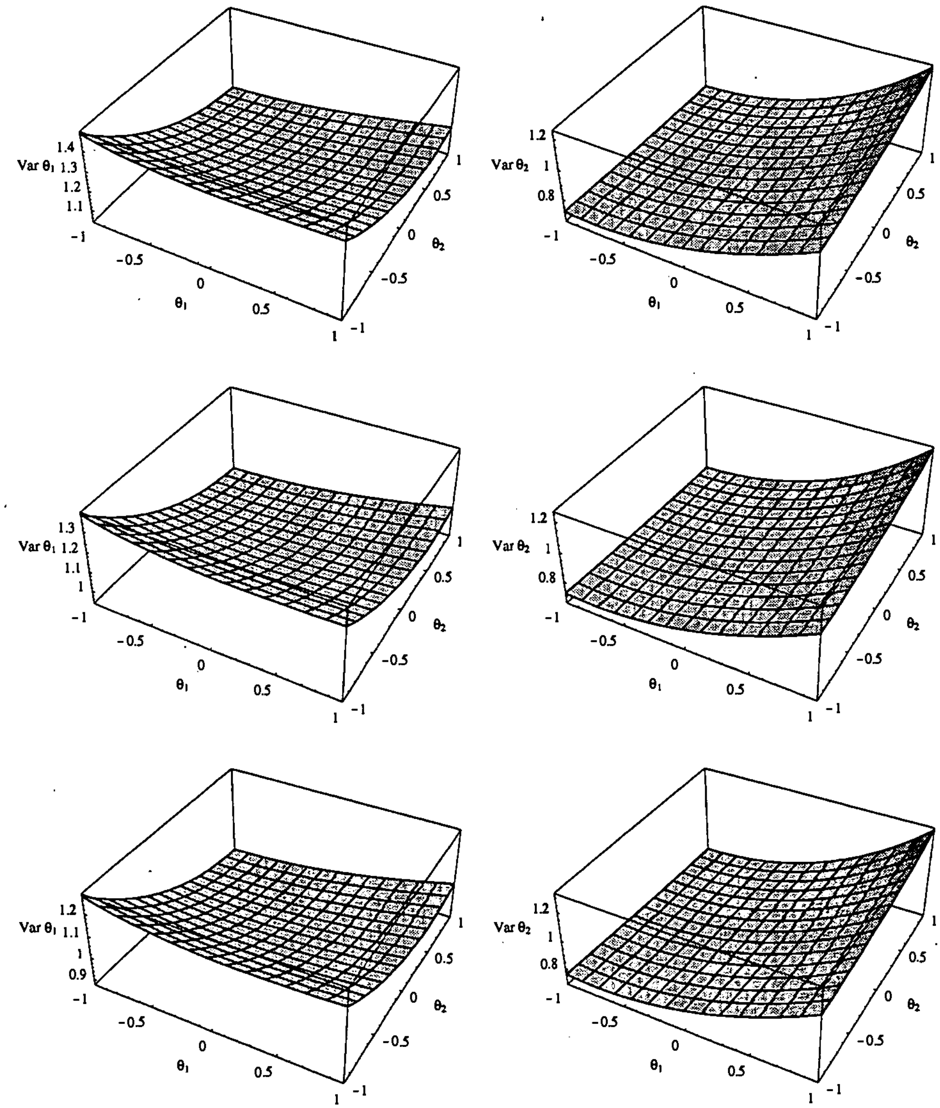
Acknowledgements

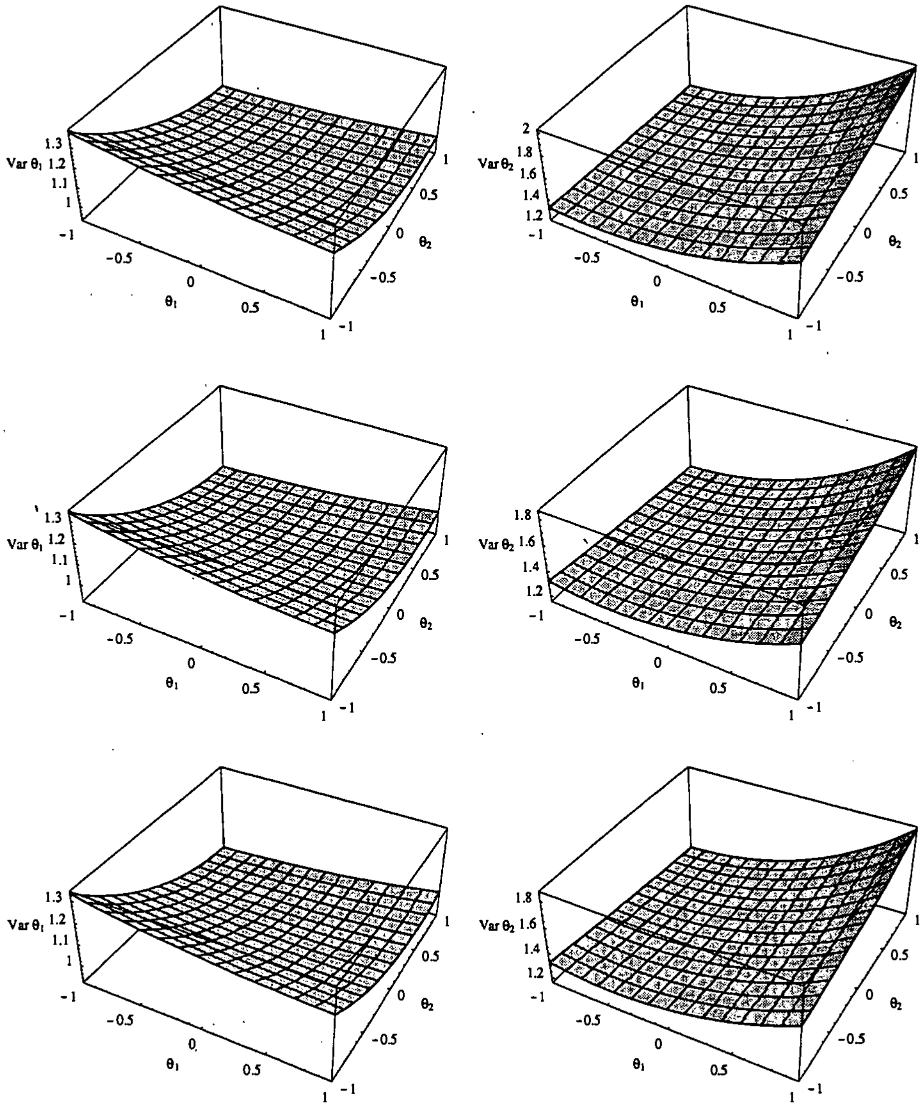
The author is indebted to Terry A. Ackerman for the dataset used in the empirical examples and to Wim J. van der Linden for his valuable comments.

Figure Captions

Figure 1. Variance functions of the MLEs of the ability parameters for the tests assembled in Example 1, when the stopping criterium was set equal to $k = 50$ (at the top), $k = 75$ (in the middle) and $k = 100$ (at the bottom) and the starting values of the correction vector were set equal to $u_0 = 0.15$.

Figure 2. Variance functions of the MLEs of the ability parameters for the tests assembled in Example 2, when the stopping criterium was set equal to $k = 50$ (at the top), $k = 75$ (in the middle) and $k = 100$ (at the bottom) and the starting values of the correction vector were set equal to $u_0 = 0.15$.


Var θ1
1.4
1.3
1.2
1.1
θ1
θ2
-1
-0.5
0
0.5
1
Var θ2
1.2
1
0.8
θ1
θ2
Var θ1
1.3
1.2
1.1
1
Var θ2
1.2
1
0.8
Var θ1
1.2
1.1
1
0.9
Var θ2
1.2
1
0.8


Var θ1
θ1
θ2
Var θ2

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

RR-98-08	B.P. Veldkamp, *Multidimensional Test Assembly Based on Lagrangian Relaxation Techniques*
RR-98-07	W.J. van der Linden & C.A.W. Glas, *Capitalization on Item Calibration Error in Adaptive Testing*
RR-98-06	W.J. van der Linden, D.J. Scrams & D.L.Schnipke, *Using Response-Time Constraints in Item Selection to Control for Differential Speededness in Computerized Adaptive Testing*
RR-98-05	W.J. van der Linden, *Optimal Assembly of Educational and Psychological Tests, with a Bibliography*
RR-98-04	C.A.W. Glas, *Modification Indices for the 2-PL and the Nominal Response Model*
RR-98-03	C.A.W. Glas, *Quality Control of On-line Calibration in Computerized Assessment*
RR-98-02	R.R. Meijer & E.M.L.A. van Krimpen-Stoop, *Simulating the Null Distribution of Person-Fit Statistics for Conventional and Adaptive Tests*
RR-98-01	C.A.W. Glas, R.R. Meijer, E.M.L.A. van Krimpen-Stoop, *Statistical Tests for Person Misfit in Computerized Adaptive Testing*
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 Testing 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*

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.

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

NOTICE

REPRODUCTION BASIS

- [x] 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").