Measurement and Research Department Reports 2000-2

IDENTIFIABILITY OF NON-LINEAR LOGISTIC TEST MODELS

T.M. Bechger
N.D. Verhelst
H.H.F.M. Verstralen

Cito

Measurement and Research Department Reports 2000-2

IDENTIFIABILITY OF NON-LINEAR LOGISTIC TEST MODELS

T.M. Bechger
N.D. Verhelst
H.H.F.M. Verstralen

Cito
Arnhem, 2000

Identifiability of Non-Linear Logistic Test Models

Timo M. Bechger[1], Norman D. Verhelst & Huub H. F. M. Verstralen
National Institute for Educational Measurement
Cito, Arnhem
The Netherlands

March 17, 2000

[1] Timo Bechger was supported by NWO, project nr. 30002

Abstract

The linear logistic test model (LLTM) specifies the item parameters as a weighted sum of basic parameters. The LLTM is a special case of a more general non-linear logistic test model (NLTM) where the weights are partially unknown. This paper is about the identifiability of the NLTM. Sufficient and necessary conditions for global identifiability are presented for a NLTM where the weights are linear functions, while conditions for local identifiability are shown to require less assumptions. It is also discussed how these conditions are checked using an algorithm due to Bekker, Merckens, and Wansbeek (1994). Several illustrations are given.

1 Introduction

For a set of k dichotomous items the *Rasch Model* (RM) is defined as

$$P(X_i = 1; \xi) = \frac{\exp(\xi - \beta_i)}{1 + \exp(\xi - \beta_i)} \quad (i \in \{1, ..., k\}) \tag{1}$$

where X_i indicates whether a response was correct (Rasch, 1960, 1966). The notation $P(X_i = 1; \xi)$ is used to denote the probability of a correct response to the ith item as a function of ξ. This probability is decreasing in β_i, which is recognized to be a *difficulty parameter* associated with the ith item. Similarly, ξ denotes a *person parameter.* The present paper is about models for the item parameters; the person parameters act as nuisance parameters.

The parameters of the RM are not unique in the sense that $\xi^* = \xi + c$ and $\beta_i^* = \beta_i + c$ for any constant c, give the same value of $P(X_i = 1; \xi)$ for all $i \in \{1, ..., k\}$. This uncertainty regarding the value of the parameters is inherent to the model and can not be diminished by taking very large samples. We can remove this arbitrariness in the parameterization by imposing a linear restriction on the item parameters, i.e.,

$$a_0 + \sum_{i=1}^{k} a_i \beta_i = 0. \tag{2}$$

The coefficients a_i are arbitrary but $\sum_{i=1}^{k} a_i$ should not be zero because otherwise (2) holds for β_i and $\beta_i + c$. A linear constraint with this property is called a *normalization.* The normalizations that are most frequently used in applications of the RM have a_0 set to zero. This facilitates the interpretation of the normalized item parameters because the absolute value of β_i will then be equal to its distance from a reference, which is formally defined by $\sum_{i=1}^{k} a_i \beta_i$. The normalized item parameters are no longer arbitrary and are said to be *identified* (or identifiable). In this report we consider the value of one of the item parameters as a reference. This amounts to setting $a_g = 1$ (for some $g \in \{1, ..., k\}$) and $a_i = 0$ $(i \neq g)$.

The *Linear Logistic Test Model* (LLTM) is defined as a RM subject to linear restrictions on the item parameters (e.g., see Fischer, 1995 and earlier references contained therein). Defining $\boldsymbol{\beta} = (\beta_1, ..., \beta_k)'$, these restrictions can be written in matrix notation as

$$\boldsymbol{\beta} = \mathbf{Q}\boldsymbol{\eta}, \tag{3}$$

where $\boldsymbol{\eta} = (\eta_1, ..., \eta_m)'$ with $m < k - 1$ is a vector of so-called *basic parameters*, and $\mathbf{Q} = (q_{ij})$ is a $k \times m$ matrix of constants which reflects some theory or hypothesis on the structure of the item parameters. Most applications of the LLTM aim at explaining the difficulty of items in terms of the underlying cognitive operations. In applications of this kind the basic parameters represent the difficulty of the certain cognitive operations.

The LLTM is a restrictive model because it requires the full specification of the $\mathbf{Q}$-matrix. Misspecification of this matrix may lead to systematic errors in the estimates of the basic parameters (Baker, 1993), and be responsible for the misfit that is frequently observed in applications of the LLTM (Fischer, 1995). To detect such misspecifications, Glas and Verhelst (1995) suggest that the score (Rao, 1947), or Lagrange multiplier test statistic (Silvey, 1959) be used to evaluate the appropriateness of an LLTM against that of a more general model where some elements of the $\mathbf{Q}$-matrix are random parameters (see Section 5.1. herein). The presence of variable elements in the $\mathbf{Q}$-matrix defines a class of *Non-Linear Logistic Test Models* (NLTM). Loosely, a NLTM is a model which has the same overall structure as the LLTM, but the weights in the matrix $\mathbf{Q}$ are only partially known.

Butter (1994), and Butter, De Boeck, and Verhelst (1998) discuss a particular NLTM (see Section 2 herein). They demonstrate that the method of conditional maximum likelihood can be used to estimate the parameters but offer no procedure to establish the identifiability of these parameters. It is essential that the parameters are identified. Unless the parameters of the NLTM are identified, there is no meaning to estimation of such parameters as several combinations of different values may lead to the same distribution of item responses. In particular, a parameter can not be estimated consistently if it is not identified (Gabrielsen, 1978).

In the present report we pursue the investigations by Butter, et al. (1998) by providing means to establish the identifiability of the parameters in a general class of models that includes their model as a special case. To illustrate our findings we discuss two applications. We focus on a NLTM where the random elements of the $\mathbf{Q}$-matrix are *linear* functions of a set of *σ-parameters.* A formal definition of this kind of NLTM is provided in Section 2. In Section 3, we present the conditions that are necessary and/or sufficient for the identification of the NLTM, given certain assumptions on the model. In Section 4, we describe a simple algorithm to determine the identifiability of a NLTM. This algorithm is due to Bekker, Merckens, and Wansbeek (1994). In Section 5, we apply the algorithm to investigate the identifiability of two

NLTMs which were encountered in practical applications. In Section 6 we conclude the article and formulate some open problems.

2 The Non-Linear Logistic Test Model

Preceeding a formal definition of the NLTM, we set the stage by discussing a small example that we have adapted from Butter, et al. (1998).

Subjects were administered three items that required them to solve mathematical problems. The first item required an addition, the second required a subtraction, and the third item an addition and a subtraction. Together these items are said to constitute a *family*. The items that refer to a single mental operation (or component) are called *subtasks*, and the item involving a combination of the components is called the *composite task*. The items may be presented in any order.

It is assumed that for a collection of $k/3$ item families (containing k items jointly) the RM is valid with the following restrictions

$$\beta_{jc} = \sigma_1\beta_{j1} + \sigma_2\beta_{j2} + \tau \qquad (j = 1, ..., k/3). \tag{4}$$

The parameter β_{jc} refers to the item difficulty of the composite item, and β_{j1} and β_{j2} to the subtask difficulties in the jth family. We assume that the difficulty of the composite task is an increasing function of the difficulties of the subtasks. As a consequence, the parameters σ_1 and σ_2 must be non-negative. When the item parameters are rescaled we get

$$\beta_{jc} + c = \sigma_1(\beta_{j1} + c) + \sigma_2(\beta_{j2} + c) + \tau \qquad (j = 1, ..., k/3), \tag{5}$$

for some c, which implies that $\sigma_1 + \sigma_2 = 1$. Under this restriction, and provided that they are non-negative, the parameters σ_1 and σ_2 give the relative weights of the components. The intercept τ represents the difficulty of a composite task over the difficulty of the subtasks. Butter (1994), and Butter, et al. (1998) present this model, without restrictions on the values of σ_1 and σ_2. Instead, the intercept τ is made dependent on the normalization (see Butter, et al., 1998, Equation 10) which complicates its interpretation.

The following equation shows the model in the form of the LLTM for

$k/3 = 2$ item families:

$$\begin{pmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{21} \\ \beta_{22} \\ \beta_{1c} \\ \beta_{2c} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1-\sigma_2 & \sigma_2 & 0 & 0 & 1 \\ 0 & 0 & 1-\sigma_2 & \sigma_2 & 1 \end{pmatrix} \begin{pmatrix} \eta_{11} \\ \eta_{12} \\ \eta_{21} \\ \eta_{22} \\ \tau \end{pmatrix} \tag{6}$$

If the values of σ_2 (or σ_1) were known, (4) would be a LLTM. In the present situation however, they are considered as model parameters. Because of the multiplication of parameters in (4) the logit of the response probabilities is non-linear in the parameters which is why we call this kind of model a non-linear logistic test model. This contrasts with Butter, et al. (1998), who refer to the model as "an item response Model with Internal Restrictions on Item Difficulty (MIRID)."

Now we normalize the item parameters and arbitrarily choose β_{11} as a reference. This implies that $\eta_{11} = 0$ so that the first column of the $\mathbf{Q}$-matrix can be deleted, giving

$$\begin{pmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{21} \\ \beta_{22} \\ \beta_{1c} \\ \beta_{2c} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sigma_2 & 0 & 0 & 1 \\ 0 & 1-\sigma_2 & \sigma_2 & 1 \end{pmatrix} \begin{pmatrix} \eta_{12} \\ \eta_{21} \\ \eta_{22} \\ \tau \end{pmatrix}. \tag{7}$$

With this normalization, the remaining parameters (η_{12}, η_{21} and η_{22}) must be interpreted relative to η_{11}.[1]

We define the *differentiable NLTM* (dNLTM) as a LLTM where elements of the $\mathbf{Q}$-matrix are differentiable functions of a set of p *σ-parameters* arranged in the vector $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_p)'$, i.e.,

$$\boldsymbol{\beta}(\boldsymbol{\eta}, \boldsymbol{\sigma}) = \mathbf{Q}(\boldsymbol{\sigma})\boldsymbol{\eta}. \tag{8}$$

The notation is chosen to indicate that the vector of item parameters and the $\mathbf{Q}$-matrix are a function (or mapping) of $(\boldsymbol{\eta}, \boldsymbol{\sigma})$ and $\boldsymbol{\sigma}$, respectively.

[1] The notion of "normalizing an LLTM" is presented by Bechger, Verstralen and Verhelst (2000). An example is given in Section 5.1., herein. We do not use a standard procedure to normalize in the dNLTM but treat each case separately.

Structural assumptions regarding the difficulty of the items determine the properties of this function. Note that the name NLTM is reserved for the larger class of models where this function is not required to be differentiable.

In this paper we focus on the class of dNLTMs where elements of the $\mathbf{Q}$-matrix are linear functions of the σ-parameters. A typical element of the matrix $\mathbf{Q}(\boldsymbol{\sigma})$ is thus given by $\mathbf{a}_{ij}^{'}\boldsymbol{\sigma}+b_{ij}$, where $\mathbf{a}_{ij} = (a_{ijh})$ is a p-dimensional vector of constants, and b_{ij} a real number. The index i is used for items, j for basic parameters, and h for elements of $\boldsymbol{\sigma}$. Using this notation,

$$\boldsymbol{\beta}(\boldsymbol{\eta},\boldsymbol{\sigma}) = \sum_{j=1}^{m}\sum_{h=1}^{p}\mathbf{a}_{.jh}\sigma_h\eta_j + \sum_{j=1}^{m}\mathbf{b}_{.j}\eta_j, \tag{9}$$

where $\mathbf{b}_{.j}$ denotes the vector $(b_{1j},...,b_{kj})^{'}$, and $\mathbf{a}_{.jh} = (a_{1jh}, a_{2jh},...,a_{kjh})^{'}$. Hence, each item parameter is a polynomial function of $\boldsymbol{\eta}$ and $\boldsymbol{\sigma}$. This polynomial has the following properties: an item parameter is never equal to a constant, and there are no powers over 1. In the absence of standard terminology for polynomials of this kind, we call this type of dNLTM the *Simple Polynomial Logistic Test Model* (SPLTM).

Changing the order of summation in (9) shows that the SPLTM may also be defined as

$$\boldsymbol{\beta}(\boldsymbol{\eta},\boldsymbol{\sigma}) = \mathbf{P}(\boldsymbol{\eta})\boldsymbol{\sigma} + \boldsymbol{\gamma}(\boldsymbol{\eta}). \tag{10}$$

The elements of the $k \times p$ matrix $\mathbf{P}(\boldsymbol{\eta})$ and the k-dimensional vector $\boldsymbol{\gamma}(\boldsymbol{\eta})$ are equal to $\mathbf{a}_{i.h}^{'}\boldsymbol{\eta}$ and $\mathbf{b}_{.i}^{'}\boldsymbol{\eta}$, respectively, where $\mathbf{a}_{i.h} = (a_{i1h},...,a_{imh})^{'}$ and $\mathbf{b}_{.i} = (b_{i1},...,b_{im})^{'}$ are vectors of constants. The model (7), for instance, can be written as

$$\begin{pmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{21} \\ \beta_{22} \\ \beta_{1c} \\ \beta_{2c} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \eta_{12} \\ \eta_{22}-\eta_{21} \end{pmatrix} (\sigma_2) + \begin{pmatrix} 0 \\ \eta_{12} \\ \eta_{21} \\ \eta_{22} \\ \tau \\ \eta_{21}+\tau \end{pmatrix} \tag{11}$$

Note that Equation 8 is a linear mapping if we consider the σ-parameters as constants, while Equation 10 defines an affine linear mapping if the basic parameters $\boldsymbol{\eta}$ are kept constant. This property of the SPLTM is used in the proof of Theorem 5 in the next section.

3 Identifiability

We assume that the item parameters are normalized, so that the parameter vector $\boldsymbol{\beta}$ is identified by the probability distribution of the item responses, and concentrate on the model for the item parameters. This implies that $\mathbf{Q}$ has been adapted, as in the previous section, so that the item parameters are normalized.

Let $\boldsymbol{\theta} = (\boldsymbol{\eta}', \boldsymbol{\sigma}')'$ denote the parameter vector of the dNLTM. Let $\boldsymbol{\beta}_0$ be the population value of the normalized item parameters. The dNLTM under investigation is assumed to hold and there exists a vector $\boldsymbol{\theta}_0 = (\boldsymbol{\eta}_0', \boldsymbol{\sigma}_0')'$ from the parameter space Ω such that

$$\boldsymbol{\beta}_0 = \boldsymbol{\beta}(\boldsymbol{\theta}_0) = \mathbf{Q}(\boldsymbol{\sigma}_0)\boldsymbol{\eta}_0. \tag{12}$$

The hypothetical "true" parameter point $\boldsymbol{\theta}_0$ being, of course, unknown. It will be supposed throughout this paper that Ω is an open and connected subset of $\mathbb{R}^{m+p}$, $(m+p)$-dimensional Euclidean space. We say that a parameter $\theta_w \in \boldsymbol{\theta}$ is *globally identified* (*locally identified*) at θ_{w0} if the value θ_{w0} is unique (locally unique)—that is, if $\boldsymbol{\beta}(\boldsymbol{\theta}^*) = \boldsymbol{\beta}(\boldsymbol{\theta}_0)$ and $\boldsymbol{\theta}^* \in \Omega$ ($\boldsymbol{\theta}^*$ is in an open neighborhood of $\boldsymbol{\theta}_0$) implies that $\theta_w^* = \theta_{w0}$ (Shapiro, 1986). We say that a dNLTM is (locally or globally) identified if all its parameters are (locally or globally) identified.

The rest of this section is structured as follows: We starts with a result concerning the local identifiability of the dNLTM. We then continue by making stronger assumptions about the model and the true parameter point. These additional assumptions allow us to derive necessary and sufficient conditions for global identifiability. An appendix is provided to show that all these assumptions hold in the SPLTM. This way the reader can see what results can be proven for the dNLTM and what results require the SPLTM. For ease of presentation we will first consider the identifiability of $\boldsymbol{\theta}$ as a whole before we consider the identifiability of separate parameters in Corollary 7.

Consider the dNLTM so that each item parameter is a differentiable function of $\boldsymbol{\theta}$. By definition, this implies that

$$\boldsymbol{\beta}(\boldsymbol{\theta}_2) - \boldsymbol{\beta}(\boldsymbol{\theta}_1) = \mathbf{J}(\boldsymbol{\theta}_1)(\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1) + ||\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1||G(\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1) \tag{13}$$

for some mapping G, defined for values of $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_1$ that are sufficiently close such that

$$\lim_{||\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1|| \to 0} G(\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1) = \mathbf{0}, \tag{14}$$

where $\mathbf{J}(\boldsymbol{\theta}_1) = \left(\frac{\partial \beta_i}{\partial \theta_w} |_{\theta_w = \theta_{w1}} \right)$ denotes the Jacobian matrix of the NLTM at $\boldsymbol{\theta}_1$, and $||.||$ denotes the Euclidean vector norm (e.g., Lang, 1996, XVI, §2). Now assume that the model is locally not identifiable at $\boldsymbol{\theta}_0$; that is, in every open neighborhood of $\boldsymbol{\theta}_0$ there are two distinct points $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ such that $\boldsymbol{\beta}(\boldsymbol{\theta}_2) - \boldsymbol{\beta}(\boldsymbol{\theta}_1) = \mathbf{0}$. Equation 13 implies that

$$G'(\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1)G(\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1) = \left(\frac{(\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1)}{||\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1||} \right)' \mathbf{J}'(\boldsymbol{\theta}_1)\mathbf{J}(\boldsymbol{\theta}_1) \frac{(\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1)}{||\boldsymbol{\theta}_2 - \boldsymbol{\theta}_1||}. \tag{15}$$

Equation 14 implies that the product on the left of (15) can be made as small as we wish by taking $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_1$ sufficiently close to $\boldsymbol{\theta}_0$. The right side of (15) is never smaller than the smallest eigenvalue of the matrix $\mathbf{J}'(\boldsymbol{\theta}_1)\mathbf{J}(\boldsymbol{\theta}_1)$ (e.g., Lang, 1987, Theorem 3.3), and since $\det(\mathbf{J}'(\boldsymbol{\theta}_1)\mathbf{J}(\boldsymbol{\theta}_1))$ is continuous (e.g., Borden, 1998, p. 257)

$$\det(\mathbf{J}'(\boldsymbol{\theta}_0)\mathbf{J}(\boldsymbol{\theta}_0)) = \lim_{\boldsymbol{\theta}_1 \to \boldsymbol{\theta}_0} \det(\mathbf{J}'(\boldsymbol{\theta}_1)\mathbf{J}(\boldsymbol{\theta}_1)) = 0, \tag{16}$$

which implies that the Jacobian matrix does not have full column rank at $\boldsymbol{\theta}_0$. The Axiom of Contraposition implies that we have proven the following proposition:

Proposition 1 *If the Jacobian matrix has full column rank at $\boldsymbol{\theta}_0$, the corresponding dNLTM is locally identifiable at $\boldsymbol{\theta}_0$.*

The converse is true under the following regularity condition (used by Wald 1950; see Fisher 1966; Rothenberg 1971).

Definition 2 *A point $\boldsymbol{\theta}_0 \in \Omega$ is* regular *if the rank of the Jacobian matrix is constant for every $\boldsymbol{\theta}$ in an open neighborhood of $\boldsymbol{\theta}_0$.*

If $\boldsymbol{\theta}_0$ is a regular point and the Jacobian matrix has deficient column rank at $\boldsymbol{\theta}_0$, the columns of the Jacobian matrix are linearly dependent for any specific $\boldsymbol{\theta}$ in an open neighborhood of $\boldsymbol{\theta}_0$. As a consequence, there exists a vector $\boldsymbol{\alpha}(\boldsymbol{\theta}) \neq \mathbf{0}$ such that

$$\mathbf{J}(\boldsymbol{\theta})\boldsymbol{\alpha}(\boldsymbol{\theta}) = \mathbf{0} \tag{17}$$

for all $\boldsymbol{\theta}$ in an open neighborhood of $\boldsymbol{\theta}_0$. Since $\mathbf{J}(\boldsymbol{\theta})$ is continuous and of constant rank, $\boldsymbol{\alpha}(\boldsymbol{\theta})$ is continuous in an open neighborhood of $\boldsymbol{\theta}_0$. This property allows us to consider a trajectory or curve $\boldsymbol{\theta}(t)$, which solves for $0 \leq t \leq t_*$ the differential equation

$$\frac{d\boldsymbol{\theta}(t)}{dt} = \boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{18}$$

where $\boldsymbol{\theta}(0) = \boldsymbol{\theta}_0$ so that the curve passes through the unknown true parameter point. It follows from the chain rule (e.g., Lang, 1996, XVI, §3) that

$$\frac{d\boldsymbol{\beta}(\boldsymbol{\theta}(t))}{dt} = \mathbf{J}(\boldsymbol{\theta})\frac{d\boldsymbol{\theta}(t)}{dt} = \mathbf{J}(\boldsymbol{\theta})\boldsymbol{\alpha}(\boldsymbol{\theta}) = \mathbf{0}, \tag{19}$$

where $\boldsymbol{\beta}(\boldsymbol{\theta}(t))$ denotes the value of the item parameters corresponding to parameter values on the curve. Equation 19 implies that $\boldsymbol{\beta}$ is constant along the curve for $0 \leq t \leq t_*$ so that $\boldsymbol{\theta}$ is not identified at $\boldsymbol{\theta}_0$.

Summarizing, we have now proven the following theorem, which is well known in Econometrics, albeit for a different kind of models (Fisher, 1966; Bekker, et al., 1994).

Theorem 3 *If $\boldsymbol{\theta}_0$ is a regular point, the parameters of the dNLTM are locally identifiable at $\boldsymbol{\theta}_0$ if and only if the Jacobian matrix of the model has full column rank at $\boldsymbol{\theta}_0$.*

Even if we restrict Ω to regular points in $\mathbb{R}^{m+p}$, global identifiability may not prevail, although the Jacobian matrix has full rank for all $\boldsymbol{\theta} \in \Omega$ (Parthasarathy, 1983). An example is given by the following dNLTM

$$\begin{pmatrix} \cos\sigma_1 & 0 \\ 0 & e^{\sigma_2} \\ e^{\sigma_2} & 0 \\ 0 & \sin\sigma_1 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} (\cos\sigma_1)\,\eta_1 \\ e^{\sigma_2}\eta_2 \\ e^{\sigma_2}\eta_1 \\ (\sin\sigma_1)\,\eta_2 \end{pmatrix}. \tag{20}$$

The determinant of the Jacobian matrix of this model is

$$\begin{vmatrix} \cos\sigma_1 & 0 & -(\sin\sigma_1)\,\eta_1 & 0 \\ 0 & e^{\sigma_2} & 0 & e^{\sigma_2}\eta_2 \\ e^{\sigma_2} & 0 & 0 & e^{\sigma_2}\eta_1 \\ 0 & \sin\sigma_1 & (\cos\sigma_1)\,\eta_2 & 0 \end{vmatrix} = -e^{2\sigma_2}\eta_1\eta_2. \tag{21}$$

The parameter σ_1 is not globally identified at any point in $\mathbb{R}$ because any two values that differ by a multiple of 2π have a common image, yet the

Jacobian matrix has full column rank except when η_1 or η_2 is zero. We will now demonstrate that (in contrast to the dNLTM) the parameters of the SPLTM are globally identified if and only if they are locally identified at any $\boldsymbol{\theta} \in \Omega$. To this aim, we need the following elementary lemma, which is easy to prove (e.g., Lang, 1996, XIV, §3).

Lemma 4 *An affine linear mapping of real vector spaces maps a straight line into a straight line.*

This lemma is now used to prove the following theorem.

Theorem 5 *The SPLTM is globally identified if and only if it is locally identifiable everywhere in Ω.* ***Proof.*** *Suppose that the model is not globally identified and there exist distinct parameter values $\boldsymbol{\theta}_3 = (\boldsymbol{\eta}_3^{'}, \boldsymbol{\sigma}_3^{'})'$ and $\boldsymbol{\theta}_1 = (\boldsymbol{\eta}_1^{'}, \boldsymbol{\sigma}_1^{'})'$ such that $\boldsymbol{\beta}(\boldsymbol{\theta}_1) = \boldsymbol{\beta}(\boldsymbol{\theta}_3)$. Let $\boldsymbol{\theta}_2 = (\boldsymbol{\eta}_3^{'}, \boldsymbol{\sigma}_1^{'})'$. Lemma 3 implies that the image of the line segment between $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ is a line segment between $\boldsymbol{\beta}(\boldsymbol{\theta}_1)$ and $\boldsymbol{\beta}(\boldsymbol{\theta}_2)$. Similarly, the image of the line segment between $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_3$ is a line segment between $\boldsymbol{\beta}(\boldsymbol{\theta}_2)$ and $\boldsymbol{\beta}(\boldsymbol{\theta}_3)$. Since $\boldsymbol{\beta}(\boldsymbol{\theta}_3) = \boldsymbol{\beta}(\boldsymbol{\theta}_1)$, and since there is only one straight line between two points, the two line segments must be the same. Hence, in every open neighborhood around $\boldsymbol{\theta}_2$, there are two points that map onto the same item parameters so that the model cannot be locally identifiable at $\boldsymbol{\theta}_2$ (Figure 1). Hence, if the model is not identifiable there must be a point where it is not identified locally. The reverse implication also holds as global identifiability implies local identifiability.* ■

The main result of this paper is given now in the following (new) theorem.

Theorem 6 *If $\boldsymbol{\theta}_0$ is a regular point, the SPLTM is globally identified at $\boldsymbol{\theta}_0$ if and only if the Jacobian matrix has full column rank at $\boldsymbol{\theta}_0$.* ***Proof.*** *Theorem 9 in the next section implies that the rank is equal at all regular points. Thus, if the Jacobian matrix has full column rank at $\boldsymbol{\theta}_0$, Theorem 3 implies that the model is locally identified at any regular point. Finally, Theorem 5 implies that the SPLTM is globally identified at any regular point in $\mathbb{R}^{m+p}$.* ■

If $\boldsymbol{\theta}$ as a whole is not identified it may still be the case that some separate parameter is identified. Equation 19 implies that, if a parameter is identified at $\boldsymbol{\theta}_0$, the corresponding element of the vector $\boldsymbol{\alpha}(\boldsymbol{\theta})$ must be zero. Since $\boldsymbol{\alpha}(\boldsymbol{\theta})$ is in the null space of the Jacobian matrix (see Equation 17), a zero row in a basis of this null space would imply that the corresponding parameter

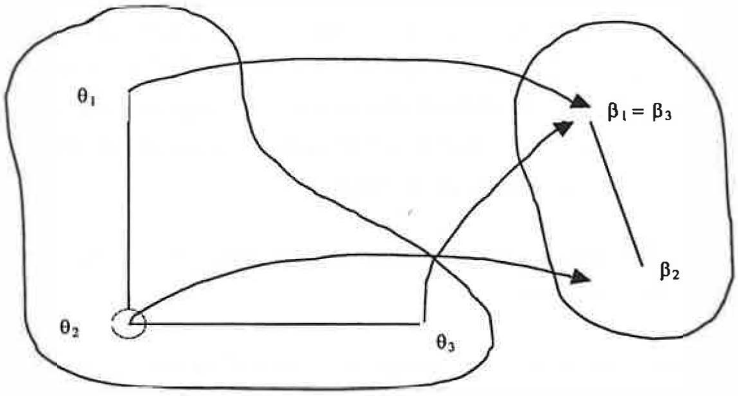

Figure 1: Relation between the parameter space and the space that contains the item parameters. The curved lines points from particular values of the parameters to their images. We have also drawn the line segments between the parameter values and their images. The circle around $\boldsymbol{\theta}_2$ represents an open ball around that point

is constant in an open neighborhood of any regular parameter point and therefore identified. These observations bring us to the following result (see also Bekker, et al., 1994, Corollary, 2.7.1).

Corollary 7 *Assume that $\boldsymbol{\theta}_0$ is a regular point. Let $\mathbf{N}(\boldsymbol{\theta})$ be a basis for the null space of $\mathbf{J}(\boldsymbol{\theta})$, i.e., $\mathbf{J}(\boldsymbol{\theta})\mathbf{N}(\boldsymbol{\theta}) = \mathbf{0}$. Let $\mathbf{e}_w$ be the wth unity vector. Then the wth parameter is globally identified if and only if $\mathbf{e}_w^{'}\mathbf{N}(\boldsymbol{\theta}) = \mathbf{0}$.*

A natural question that arises from the preceding discussion is: How frequent are the regular points ? The following important result is due to Fisher (1966, p. 167; also Andres, 1990, Section 3.3).

Theorem 8 *Let β_i, $i \in \{1, ..., k\}$, be real analytic functions on Ω.*[2] *Then the set of irregular points is of Lebesgue measure zero, i.e., almost all $\boldsymbol{\theta} \in \Omega$ are regular.*

[2]A function is real analytic on an open set $U \subseteq \mathbb{R}$ if it may be represented by a convergent power series in a neighborhood of every point of this set. See Dieudonné (1969, IX) or Krantz and Parks (1992; Definition 1.1.3). In particular, polynomial functions are analytic on $\mathbb{R}$.

In other words, if we consider $\boldsymbol{\theta}$ to be random with some non-degenerate, continuous distribution, the irregular values constitute a set of zero area and the probability that $\boldsymbol{\theta}_0$ is not a regular point is zero. In view of Theorem 8, a parameter is said to be identified (or not identified) *almost everywhere* in $\mathbb{R}^{m+p}$ (abbreviated to a.e.$[\mathbb{R}^{m+p}]$) to indicate that this statement is true under the assumption that $\boldsymbol{\theta}_0$ is a regular point.[3]

4 An Algorithm to Determine the Identifiability of an SPLTM

In the SPLTM, all item parameters are real analytic functions on $\mathbb{R}$. The Jacobian matrix has the following generic form

$$\mathbf{J}(\boldsymbol{\theta}) = (\mathbf{Q}(\boldsymbol{\sigma}), \mathbf{P}(\boldsymbol{\eta})). \tag{22}$$

Since $\mathbf{Q}(\boldsymbol{\sigma})$ is known, we can assemble the Jacobian matrix using the formulae presented in Section 2. Note that the Jacobian matrix may be much more complicated in the dNLTM (see e.g., 21).

To determine the rank of $\mathbf{J}(\boldsymbol{\theta}_0)$ and at the same time find a basis of the null space of $\mathbf{J}(\boldsymbol{\theta})$, Bekker, et al. (1994) propose a simple procedure, which can be performed by ubiquitous computer packages for symbolic math such as Maple.[4] When applied in the present situation, this procedure consists in the use of elementary row operations (Gaussian elimination) to transform the augmented matrix $[\mathbf{J}'(\boldsymbol{\theta}), \mathbf{I}_{m+p}]$ into reduced row echelon form (RREF), where $\mathbf{I}_{m+p}$ denotes an $(m+p)$-dimensional unit matrix. Let $\mathbf{E}$ be a non-singular matrix representing the elementary row operations. Let $\mathbf{E}$ be partitioned as $\mathbf{E} = (\mathbf{E}_1, \mathbf{E}_2)$, where $\mathbf{E}_1$ has d columns, and $\mathbf{E}_2$ has $k-d$ columns. The matrix that results from Gaussian elimination can be partitioned as follows.

$$\mathbf{E}'[\mathbf{J}'(\boldsymbol{\theta}), \mathbf{I}_{m+p}] = \begin{pmatrix} \mathbf{E}_1' \\ \mathbf{E}_2' \end{pmatrix} [\mathbf{J}'(\boldsymbol{\theta}), \mathbf{I}_{m+p}] = \begin{pmatrix} \mathbf{E}_1'\mathbf{J}'(\boldsymbol{\theta}) & \mathbf{E}_1' \\ \mathbf{E}_2'\mathbf{J}'(\boldsymbol{\theta}) & \mathbf{E}_2' \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix}, \tag{23}$$

[3]Note that a model that is identifiable a.e.$[\Omega]$ is sometimes called identified (e.g., Luijben, 1991, p. 656). We prefer to distinguish explicitly between local identifiability, global identifiability and identifiability a.e.$[\Omega]$.

[4]An alternative procedure is discussed by Bekker (1989), and a program called ERA is provided with the Bekker, et al. (1994) book.

where $\mathbf{R}_{21}$ is a zero matrix with k columns and as many rows as possible. Since $\mathbf{J}(\boldsymbol{\theta})\mathbf{E}_2 = \mathbf{0}$, the matrix $\mathbf{E}_2 = \mathbf{R}_{22}'$ gives a basis for the null space of the Jacobian matrix. According to Corollary 7, a zero-row in $\mathbf{E}_2$ indicates that the corresponding parameter is identifiable a.e.$[\mathbb{R}^{m+p}]$. The rank of the Jacobian matrix at any regular point is given by d, which is the number of columns minus the dimensionality of the null space, i.e., $k - (k - d)$ (cf. Bekker, 1989, prop. 2). The following theorem implies that d is also the maximum rank of the Jacobian matrix when the maximum is taken over real values of the parameters.

Theorem 9 *Let β_i, $i \in \{1, ..., k\}$, be real analytic functions on Ω. Then $\boldsymbol{\theta}_0$ is a regular point if and only if*

$$\mathrm{rank}(\mathbf{J}(\boldsymbol{\theta}_0)) = \max_{\boldsymbol{\theta} \in \mathbb{R}^{m+p}} \{\mathrm{rank}(\mathbf{J}(\boldsymbol{\theta}))\}. \tag{24}$$

Proof. *The proof is given by Shapiro (1983, p. 9), or Bekker, et al. (1994, Theorem 2.6.1). In the appendix we provide a relatively simple but less general proof.* ■

As an illustration we apply the algorithm to the following SPLTM:

$$\begin{pmatrix} 2\sigma_1 & \sigma_2 \\ \sigma_1 & 0 \\ 1 & \sigma_2 \\ \sigma_1 & 0 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} 2\sigma_1\eta_1 + \sigma_2\eta_2 \\ \sigma_1\eta_1 \\ \eta_1 + \sigma_2\eta_2 \\ \sigma_1\eta_1 \end{pmatrix}. \tag{25}$$

The augmented matrix is

$$\begin{pmatrix} 2\sigma_1 & \sigma_1 & 1 & \sigma_1 & 1 & 0 & 0 & 0 \\ \sigma_2 & 0 & \sigma_2 & 0 & 0 & 1 & 0 & 0 \\ 2\eta_1 & \eta_1 & 0 & \eta_1 & 0 & 0 & 1 & 0 \\ \eta_2 & 0 & \eta_2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{26}$$

If the augmented matrix is transformed into RREF we have

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -1 & 0 & \frac{1}{\eta_1}\sigma_1 & \frac{1}{\eta_2} \\ 0 & 1 & 0 & 1 & 2 & 0 & -\frac{-1+2\sigma_1}{\eta_1} & \frac{2}{\eta_2} \\ 0 & 0 & 1 & 0 & 1 & 0 & -\frac{1}{\eta_1}\sigma_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -\frac{1}{\eta_2}\sigma_2 \end{pmatrix}. \tag{27}$$

Consequently, for all regular points the rank of $\mathbf{J}(\boldsymbol{\theta})$ equals 3. The vector $\begin{bmatrix} 0 & 1 & 0 & -\frac{1}{\eta_2}\sigma_2 \end{bmatrix}'$ furnishes a basis for the null space of the Jacobian matrix, which indicates that σ_1 and η_1 are identifiable a.e.$[\mathbb{R}^{m+p}]$, while σ_2 and η_2 are *not* identifiable a.e.$[\mathbb{R}^{m+p}]$. If we set $\sigma_2 = 1$, the resulting model is identified a.e.$[\mathbb{R}^{m+p}]$.

Cramer's rule (e.g., Lang, 1987, p. 157) states that the elements of vectors in the null space are ratio's of subdeterminants of $\mathbf{J}(\boldsymbol{\theta})$. In the SPLTM, the elements of $\mathbf{J}(\boldsymbol{\theta})$ are polynomials and any subdeterminant of $\mathbf{J}(\boldsymbol{\theta})$ is the sum of products of polynomials (see Appendix). Since the ring of polynomials is closed under addition and multiplication, subdeterminants of $\mathbf{J}(\boldsymbol{\theta})$ are also polynomials. Hence the elements of $\boldsymbol{\alpha}(\boldsymbol{\theta})$ are rational functions in $\boldsymbol{\theta}$, as in the example.

5 Illustrations

5.1 The Identifiability of Single Elements of the Q-Matrix

Suppose that we follow the suggestion by Glas and Verhelst (1995), and use the score test to investigate the specification of an LLTM. First we adapt the $\mathbf{Q}$-matrix so that the item parameters are normalized. If the rth item is taken as a reference, we must pre-multiply $\mathbf{Q}$ with a matrix $\mathbf{L}_r$, that is obtained from $\mathbf{I}_k$ by subtracting the rth row from any other row (including the rth row) of $\mathbf{I}_k$. Thus, we obtain *a normalized LLTM* :

$$\boldsymbol{\beta}_r = \mathbf{L}_r\mathbf{Q}\boldsymbol{\eta} \equiv \mathbf{Q}_r\boldsymbol{\eta}. \tag{28}$$

Let this LLTM represent the null-hypothesis. The alternative hypothesis is a SPLTM with parameter vector $\boldsymbol{\theta} = (\boldsymbol{\eta}', \boldsymbol{\sigma}')'$, where $\boldsymbol{\sigma}$ contains parameters that represent the unknown values of those entries in the $\mathbf{Q}$-matrix whose specification we wish to investigate. Let $\mathbf{S}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_L)$ be the vector of first order partial derivatives of the loglikelihood of the SPLTM evaluated at $\boldsymbol{\theta}_L$, which denotes the Maximum Likelihood (ML) estimates under the LLTM. That is, $\boldsymbol{\theta}_L = (\hat{\boldsymbol{\eta}}', \mathbf{c}')'$, where $\mathbf{c}$ is a vector of constants, and $\hat{\boldsymbol{\eta}}$ denotes the ML estimates of the basic parameters. Similarly, $\mathbf{I}_{\boldsymbol{\theta},\boldsymbol{\theta}}(\boldsymbol{\theta}_L)$ denotes the information matrix evaluated at $\boldsymbol{\theta}_L$. The *Score Test* (ST) is defined as

$$\mathbf{S}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_L)'\mathbf{I}_{\boldsymbol{\theta},\boldsymbol{\theta}}(\boldsymbol{\theta}_L)^{-1}\mathbf{S}_{\boldsymbol{\theta}}(\boldsymbol{\theta}_L). \tag{29}$$

If the postulated LLTM holds, this statistic follows a central chi-square distribution with degrees of freedom equal to the number of parameters in $\boldsymbol{\sigma}$. The partial derivatives of the loglikelihood function with respect to the basic parameters are zero since they are evaluated at the ML estimates. Hence, the corresponding rows in $\mathbf{I}_{\theta,\theta}(\boldsymbol{\theta}_L)^{-1}$ can be deleted. Since $\boldsymbol{\theta} = (\boldsymbol{\eta}', \boldsymbol{\sigma}')'$ is partitioned, the score vector and the information matrix may likewise be partitioned:

$$\mathbf{S}_{\theta}(\boldsymbol{\theta}_L) = \begin{pmatrix} \mathbf{S}_{\eta}(\boldsymbol{\theta}_L) \\ \mathbf{S}_{\sigma}(\boldsymbol{\theta}_L) \end{pmatrix}, \mathbf{I}_{\theta,\theta}(\boldsymbol{\theta}_L) = \begin{pmatrix} \mathbf{I}_{\eta,\eta}(\boldsymbol{\theta}_L) & \mathbf{I}_{\eta,\sigma}(\boldsymbol{\theta}_L) \\ \mathbf{I}_{\sigma,\eta}(\boldsymbol{\theta}_L) & \mathbf{I}_{\sigma,\sigma}(\boldsymbol{\theta}_L) \end{pmatrix}. \tag{30}$$

The inverse of the information matrix can be partitioned using standard formulae for the inverse of a partitioned matrix. After simplification, we obtain the following expression for the ST:

$$\mathbf{S}_{\sigma}(\boldsymbol{\theta}_L)'\mathbf{D}^{-1}\mathbf{S}_{\sigma}(\boldsymbol{\theta}_L), \tag{31}$$

where

$$\mathbf{D} = \mathbf{I}_{\sigma,\sigma}(\boldsymbol{\theta}_L) - \mathbf{I}_{\sigma,\eta}(\boldsymbol{\theta}_L)\mathbf{I}_{\eta,\eta}(\boldsymbol{\theta}_L)^{-1}\mathbf{I}_{\eta,\sigma}(\boldsymbol{\theta}_L) \tag{32}$$

(cf. Glas & Verhelst, 1995).

Suppose that the ST is used to investigate the specification of the (g, l) entry in the $\mathbf{Q}$-matrix. An application of this kind is reported in detail by Bechger, Verstralen and Verhelst (2000). If the ST turns out to be significant, we may consider this entry as a parameter and estimate its value. Before doing so we must find out whether the resulting SPLTM is identified. This turns out to be relatively easy. The Jacobian matrix is

$$\mathbf{J} = [\mathbf{Q}_r(\sigma_{gl}), \mathbf{P}(\eta_l)], \tag{33}$$

where, for convenience, $r \neq g$ so that $\mathbf{P}(\eta_l)$ denotes a column vector with all elements equal to zero except the gth entry which is equal to η_l. Without loss of generality we first re-order the rows of the Jacobian matrix such that the gth row is the last row. Then, it follows from basic properties of partitioned matrices (e.g., Basilevsky, 1983, Equation 4.9.2) that

$$\det(\mathbf{J}'\mathbf{J}) = \eta_l^2 \det(\mathbf{Q}'_{-g}\mathbf{Q}_{-g}), \tag{34}$$

where $\mathbf{Q}_{-g}$ denotes the $\mathbf{Q}$-matrix with the gth row deleted. The results in the third section of this paper imply that the model is identified a.e.$[\mathbb{R}^{m+p}]$ if and only if $\mathbf{Q}_{-g}$ has full column rank. The point where $\eta_l = 0$ is irregular. It is, however, easy to see that the parameter σ_{gl} is not identified at $\eta_l = 0$. At all other irregular points, the Jacobian matrix has full column rank so that the model is locally identified there by Proposition 1. Theorem 5 implies that σ_{gl} is globally identifiable at $\boldsymbol{\theta}_0$ if and only if $\mathbf{Q}_{-g}$ has full column rank, and $\eta_l \neq 0$.

5.2 A Simple Polynomial Logistic Test Model for Facet Designs

Suppose that children that have participated in a music course are expected to know n_M melodies by name and be able to recognize each of n_T times. To test their knowledge as well as their musical abilities, the children are required to listen to small pieces of music, sing along with the melody, and simultaneously beat time with their hands. The exam also includes sessions where the children have to sing either a melody or beat time. At each occasion their performance is judged as sufficient or insufficient by their music teacher.

This example can be placed in a more general context by noting that the composite tasks are composed as elements of the Cartesian product of two *facets*: melodies and times. The use of this model for facet designs was suggested by Butter (1994). It is, however, beyond the scope of this paper to develop a SPLTM for facet designs in its full generality. We merely use the example to illustrate that the SPLTM can be used to analyze data of this kind.

As in Section 2, we assume that the difficulty of this composite task is an additive function of the difficulty of the subtasks: singing and beating time. To be more specific, the difficulty of the composite task is given by

$$\beta_{av} = \sigma_M \beta_{a0} + \sigma_T \beta_{0v} + \tau \quad (a \in \{1, ..., n_M\} \text{ and } v \in \{1, ..., n_T\}), \tag{35}$$

where β_{a0} denotes the difficulty of singing the ath melody, β_{0v} denotes the difficulty of beating the vth time, and β_{av} denotes the difficulty of doing both things simultaneously. To improve the interpretability of the parameters and to assure that the structure of the model is invariant under normalization, we make the following assumptions:

1. Monotonicity: For all melodies a, b and all times v,w,

$$\beta_{a0} \geq \beta_{b0} \Leftrightarrow \beta_{av} \geq \beta_{bv} \tag{36}$$

$$\beta_{0v} \geq \beta_{0w} \Leftrightarrow \beta_{aw} \geq \beta_{aw} \tag{37}$$

2. Invariance under normalization: For any $c \in \mathbb{R}$, $a \in \{1, ..., n_M\}$ and $v \in \{1, ..., n_T\}$

$$\beta_{av} + c = \sigma_M(\beta_{a0} + c) + \sigma_T(\beta_{0v} + c) + \tau. \tag{38}$$

The monotonicity assumption asserts that weak order relations among the difficulties of singing separate melodies, or beating separate times also hold when singing and beating time are performed simultaneously. This assumption implies that the weights σ_M and σ_T are non-negative.[5] The invariance under normalization assumption implies that $\sigma_M + \sigma_T = 1$. We can therefore interpret the parameters σ_M and σ_T as the relative difficulty of singing and beating time, respectively, as perceived by the music teacher.

The composite task requires the coordination of the hands and the voice, which is not necessary to perform any of the subtasks separately. The intercept τ represents the difficulty of this coordination as perceived by the music teacher. To see this, consider that if $\tau = 0$ the difficulty of the composite task is restricted to lie somewhere between the difficulties of each subtask. If $\tau \neq 0$, the difficulties of the composite tasks are all shifted an amount τ. Hence, the intercept τ represents the difficulty of a composite task over a subtask.

When there are two melodies and two times we have the following model for difficulty parameters:

$$\begin{pmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{21} \\ \beta_{22} \\ \beta_{10} \\ \beta_{20} \\ \beta_{01} \\ \beta_{02} \end{pmatrix} = \begin{pmatrix} 1-\sigma_T & 0 & \sigma_T & 0 & 1 \\ 1-\sigma_T & 0 & 0 & \sigma_T & 1 \\ 0 & 1-\sigma_T & \sigma_T & 0 & 1 \\ 0 & 1-\sigma_T & 0 & \sigma_T & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \eta_{10} \\ \eta_{20} \\ \eta_{01} \\ \eta_{02} \\ \tau \end{pmatrix} \tag{39}$$

We take β_{10} as a reference. This implies that $\eta_{10} = 0$ so that first column of

[5] Alternatively, we might have assumed that the weights are non-negative and this would, together with the invariance under normalization, have implied monotonicity.

$\mathbf{Q}$ can be deleted giving

$$\begin{pmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{21} \\ \beta_{22} \\ \beta_{10} \\ \beta_{20} \\ \beta_{01} \\ \beta_{02} \end{pmatrix} = \begin{pmatrix} 0 & \sigma_T & 0 & 1 \\ 0 & 0 & \sigma_T & 1 \\ 1-\sigma_T & \sigma_T & 0 & 1 \\ 1-\sigma_T & 0 & \sigma_T & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \eta_{20} \\ \eta_{01} \\ \eta_{02} \\ \tau \end{pmatrix} \tag{40}$$

Ignoring the zero row in $\mathbf{Q}$, the Jacobian matrix of this model is

$$[\mathbf{Q}(\boldsymbol{\sigma}), \mathbf{P}(\boldsymbol{\eta})] = \begin{pmatrix} 0 & \sigma_T & 0 & 1 & \eta_{01} \\ 0 & 0 & \sigma_T & 1 & \eta_{02} \\ 1-\sigma_T & \sigma_T & 0 & 1 & \eta_{20}-\eta_{01} \\ 1-\sigma_T & 0 & \sigma_T & 1 & \eta_{20}-\eta_{02} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

The null space of this matrix is empty, which means that the model is identifiable a.e.$[\mathbb{R}^{m+p}]$.

6 Concluding Remarks

The LLTM models the item difficulty parameters of the RM as a weighted combination of m basic parameters. In some applications, this model is considered too restrictive because the weights must be specified in advance. We have discussed two such applications. The first application concerned the estimation or respecification of a single entry in the $\mathbf{Q}$-matrix, and the second application was a model for measurements obtained from a crossed two-facet design. In each of these applications, the weights were specified as linear functions of additional random parameters. This model was called the SPLTM. Butter, et al. (1998) considered a special case of the SPLTM and demonstrated that the method of conditional maximum likelihood can be used to estimate the parameters. The aim of the present paper was to discuss conditions for the identification of the SPLTM, and the more general dNLTM, and provide ways to check these conditions.

We have found that the SPLTM is identified at almost every real parameter point if and only if the maximum rank of the symbolic matrix $[\mathbf{Q}(\boldsymbol{\sigma}), \mathbf{P}(\boldsymbol{\eta})]$ is equal to the number of its columns. We have also found that the null space of this matrix provides information on the identification of individual parameters. To arrive at these results we have profited from work done in Econometrics (Fisher, 1966; Rothenberg, 1971), and structural equation modeling (Andres, 1990; Bekker, et al., 1994; Shapiro, 1983;1986). These publications have almost exclusively dealt with local identification (as in the dNLTM) and our results about the global identifiability of the SPLTM are quite unique. Bekker, et al. (1994) have demonstrated that Gaussian elimination can be used to obtain all the required information. Most of the recently developed programs for computer algebra incorporate Gaussian elimination and can be used for this purpose. Without this algorithm it may be very difficult to determine whether a particular parameter is identifiable, even when the number of items is small. The reader is challenged to try, for instance, to determine whether (25) is identifiable.

Although a model that is not identified would occur in practice with negligible probability, one might find instances where a solution is close to an unidentifiable model causing numerical instabilities and convergence difficulties. Given a model that is identifiable a.e.$[\mathbb{R}^{m+p}]$, it is therefore of interest to know which irregular values of the parameters would render the model unidentified. This knowledge would also serve to avoid the specification of a restriction that would render a unidentified model. In general, the irregular values are real roots of the determinant of the symbolic matrix $\mathbf{J}'(\boldsymbol{\theta})\mathbf{J}(\boldsymbol{\theta}) = [\mathbf{Q}(\boldsymbol{\sigma}), \mathbf{P}(\boldsymbol{\eta})]'[\mathbf{Q}(\boldsymbol{\sigma}), \mathbf{P}(\boldsymbol{\eta})]$. In general, this determinant is a complicated multivariate polynomial. The applications in this paper show that in some cases these roots may be found on the back of an envelope but to the best of our knowledge there is no general algorithm which enables us to solve such complicated multivariate polynomials automatically. A notable exception is when there is only one parameter in the Jacobian matrix so that the elements are univariate polynomials (e.g., Henrion & Sebel, 1998).

There is no SPLTM that is globally identified at any (regular or irregular) true parameter point since we can always find an *irregular* point where the model is not identifiable, e.g., $\boldsymbol{\eta} = \mathbf{0}$. That is, we may always find a restriction that will make a model unidentified. Through our experience with the SPLTM we have come to believe that there are no irregular points where the SPLTM is locally identified. This conjecture has been proven wrong for factor analysis (Shapiro & Browne, 1983; Shapiro, 1985) but we did not find any counterexample for the SPLTM. The proof of this conjecture is a topic for future research. To find out whether the true parameter is regular, one could use a sample estimate of the Jacobian matrix to test whether this matrix has full column rank at the population value $\boldsymbol{\theta}_0$ (Gill & Lewbel, 1992). This same procedure can be used to test the local identifiability of a dNLTM by Proposition 1.

Finally, we hope that the present article stimulates the application of the SPLTM but we realize that this is contingent on the availability of user-friendly software. We therefore consider the development of such software to be an important topic for future work. A final topic for future research is the generalization of the ideas presented here to models for polytomous items (e.g., by extending the Linear Partial Credit Model (LPCM); Fischer & Ponocny, 1994; 1995) and/or multidimensional models (e.g., Andersen, 1995). We believe that the conditions for identifiability generalize to such models.

7 Appendix

A function is a real polynomial if it consists of the sum of products of real parameters. Hence, in the SPLTM each item parameter is a polynomial function of $\boldsymbol{\sigma}$ and $\boldsymbol{\eta}$. This polynomial is characterized by the fact that there are no powers over 1 and that no item parameter is equal to a constant. There appears to be no conventional way to refer to such polynomials and we have somewhat arbitrarily chosen to call it "a simple polynomial." A polynomial is also a power series on $\mathbb{R}^{m+p}$. To see this, look at the definition of a power series given by Dieudonné (1969, p. 199), and add an infinite number of zeroes to the polynomial. By definition then, a polynomial is a real analytic function on $\mathbb{R}^{m+p}$. Theorem 9.3.6 in Dieudonné (1969) implies that it is indefinitely differentiable.

We will now prove that Theorem 8 holds when the model is a SPLTM. To this aim we need the following Lemma (see also Shapiro, 1983, Lemma 2).

Lemma 10 *A simple multivariate polynomial is either identical to zero or zero for a subset of Lebesgue measure zero.* ***Proof.*** *Suppose that the polynomial is unequal to zero but vanishes on an open and connected subset of Ω. Hence, the polynomial is not a non-zero constant. It must then be possible to change only one arbitrary parameter without changing the value of the polynomial. However, a simple polynomial consists of sums of products and it is necessary to change at least one other parameter to compensate for the change in the former parameter. This means that there cannot be an open subset of positive Lebesgue measure where the polynomial vanishes. If a polynomial consists of a single parameter the set where it is zero is a set of Lebesgue measure zero.* ■

The determinant of $\mathbf{J}'(\boldsymbol{\theta})\mathbf{J}(\boldsymbol{\theta})$ is calculated as the sum of products of parameters in $\boldsymbol{\theta}$. Hence, it is a polynomial and so is each minor of $\mathbf{J}(\boldsymbol{\theta})$. This will be used to demonstrate that (24) holds.

Suppose that $\boldsymbol{\theta}_0$ is a regular point and $\operatorname{rank}(\mathbf{J}(\boldsymbol{\theta})) = d \leq (m+p)$, which means that $\mathbf{J}(\boldsymbol{\theta})$ has a non-zero $d \times d$ minor. Since this minor is continuous there exist an open neighborhood around $\boldsymbol{\theta}_0$ where it is non-zero. Hence this minor is non-zero for almost every $\boldsymbol{\theta} \in \mathbb{R}^{m+p}$. If $d < (m+p)$ there is a $s \times s$ minor, $s > d$, which includes the $d \times d$ non-zero minor, that is zero in a neighborhood of $\boldsymbol{\theta}_0$. Since the minor is a real polynomial, it follows from

Lemma 10 that any $s \times s$ minor is identically zero on $\mathbb{R}$. It follows that the rank of $\mathbf{J}$ is d or less than any $\boldsymbol{\theta} \in \mathbb{R}^{m+p}$. This suffices to prove Theorem 9 for the case of the SPLTM.

8 References

Andersen, E. B. (1995). Polytomous Rasch models and their estimation. In G. H. Fisher & I.W. Molenaar (Eds.). *Rasch models: Foundations, recent developments and applications.* New-York : Springer Verlag.

Andres, J. (1990). *Grundlagen Linearer Structurgleichungsmodelle.*[Foundations of linear structural equation models] New-York: Peter Lang.

Baker, F. B. (1993). Sensitivity of the linear logistic test model to misspecification of the weight matrix. *Applied Psychological Measurement*, *17*, 201-211.

Basilevsky, A. (1983). *Applied matrix algebra in the statistical sciences.* New-York: North-Holland.

Bechger, T. M. , Verstralen, H. H. F. M. , & Verhelst, N. D. (2000). A Modification index for the linear logistic test model. Under editorial review.

Bekker, P. A. (1989). Identification in restricted factor models. *Journal of Econometrics*, *41*, 5-16.

Bekker, P. A. , Merckens, A., & Wansbeek, T. J. (1994). *Identification, equivalent models, and computer algebra.* Boston: Academic Press.

Bollen, K. A. (1989). *Structural equations with latent variables.* New-York: Wiley

Borden, R. S. (1998). *A course in advanced calculus.* New-York: Dover Publ., Inc.

Buse, A.(1982). The likelihood ratio, Wald and Lagrange multiplier test: An expository note. *The American Statistician*, *36* (3), 153-157.

Butter, R. P. (1994). *Item response models with internal restrictions on item difficulty.* Unpublished Doctoral Dissertation, Catholic University of Leuven, Belgium.

Butter, R. P. , De Boeck, P. , & Verhelst, N. D. (1998). An item response model with internal restrictions on item difficulty. *Psychometrika*, *63*, 47-63.

Dieudonné, J. (1969). *Foundations of modern analysis.* New-York: Academic Press.

Fischer, G. H. (1983). Logistic latent trait models with linear constraints. *Psychometrika*, *48*, 3-26.

Fischer, G. H. (1995). The linear logistic test model. In G. H. Fisher & I.W. Molenaar (Eds.). *Rasch models: Foundations, recent developments and applications.* New-York : Springer Verlag.

Fischer, G. H. , & Ponocny, I. (1994). An extension of the partial credit model with an application to the measurement of change. *Psychometrika*, *59*, 177-192.

Fischer, G. H. , & Ponocny, I. (1995). Extended rating scale and partial credit models for assessing change. Chapter 19 In G. H. Fischer & I.W. Molenaar (Eds.). *Rasch models: Foundations, recent developments and applications.* New-York : Springer Verlag.

Fisher, F. M. (1966). *The identification problem in Econometrics.* New-York: McGraw-Hill.

Gabrielsen, A. (1978). Consistency and identifiability. *Journal of Econometrics*, *8*, 261-263.

Gill, L., & Lewbel, A. (1992). Testing the rank and definiteness of estimated matrices with application to factor, state-space and ARMA models. *Journal of the American Statistical Association*, *87*, 766-776.

Glas, C. A. W. , & Verhelst, N. D. (1995). Testing the Rasch model. In G. H. Fischer & I. W. Molenaar (Eds.), *Rasch models: Their foundations, recent developments and applications.* New-York: Springer Verlag.

Henrion, D. , & Sebel, M. (1998). Numerical methods for polynomial rank evaluation. *Proceedings of the IFAC conference on systems structure and control. Nantes.* France, July, 1998.

Krantz, S. G. , & Parks, H. R. (1992). A primer on real analytic functions. Basel: Birkhauser Verlag.

Lang, S. (1987). *Linear Algebra.* (Third Edition). New-York: Springer Verlag.

Lang, S. (1996). *Calculus of several variables.* (Third Edition). New-York: Springer Verlag.

Luijben, Th. C. W. (1991). Equivalent models in Covariance structure analysis. *Psychometrika*, *56*, 653-665.

Parthasarathy, T. (1983). *On global univalence theorems.* Lecture Notes in Mathematics. New-York: Springer Verlag.

Rao, C. R. (1947). Large sample tests of statistical hypothesis concerning several parameters with applications to the problems of estimation. *Proceedings Cambridge Philosophical Society*, *44*, 50-57.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Copenhagen: Danish Institute for Educational Research. [Expanded edition, University of Chigago Press, 1980.]

Rasch, G. (1966). An individualistic approach to item analysis. In P. F. Lazersfeld & N. W. Henry (Eds.) *Readings in mathematical social science.* Cambridge: The MIT Press.

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica*, *39*, 577-591.

Shapiro, A. (1983). *On local identifiability in structural models.* Unpublished research report. University of South Africa, Department of Mathematics and Applied Mathematics.

Shapiro, A. (1985). Identifiability of factor analysis: Some results and open problems. *Linear Algebra and Its Applications*, *70*, 1-7.

Shapiro, A. (1986). Asymptotic theory of overparameterized structural models. *Journal of the American Statistical Association*, *81*,142-149.

Shapiro, A. , & Browne, M. W. (1983). On the investigation of local identifiability: A counterexample. *Psychometrika*, *48*, 303-304.

Silvey, S. D. (1959). The Lagrange multiplier test. *Annals of Mathematical Statistics*, *30*, 398-407.

Wald, A. (1950). Note on the identifiability of economic relations. In T. C. Koopmans (Ed.), *Statistical Inference in Dynamic Economic Models.* Cowles Commission Monograph No. 10. New-York: Wiley.