PSYCHOMETRIKA—VOL. 52, NO. 3, 393–408
SEPTEMBER 1987

ON THE RELATIONSHIP BETWEEN ITEM RESPONSE THEORY AND FACTOR ANALYSIS OF DISCRETIZED VARIABLES

YOSHIO TAKANE

MCGILL UNIVERSITY

JAN DE LEEUW

UNIVERSITY OF LEIDEN

Equivalence of marginal likelihood of the two-parameter normal ogive model in item response theory (IRT) and factor analysis of dichotomized variables (FA) was formally proved. The basic result on the dichotomous variables was extended to multicategory cases, both ordered and unordered categorical data. Pair comparison data arising from multiple-judgment sampling were discussed as a special case of the unordered categorical data. A taxonomy of data for the IRT and FA models was also attempted.

Key words: marginal maximum likelihood estimation, dichotomous data, ordered and unordered categorical data, pair comparison data.

1. Introduction

Repeated measures designs, broadly construed, are frequently employed in psychological investigations. In these designs each of a group of subjects is repeatedly measured under a set of different conditions, thereby contributing more than one observation per data set. The conditions may represent different experimental manipulations, different occasions of measurement or different test items to which the subjects respond. There are various reasons for the popularity of the repeated measures designs in psychological research. In experimental-manipulative contexts complete randomization is often difficult to realize (particularly with human subjects), and an alternative approach based on matched samples is getting increasingly popular. In more observational settings, an interest may be in how the different measurement conditions (or variables) relate with each other in the population of subjects. When this latter interest is emphasized, the repeated measures data are simply called multivariate (profile) data in which each subject is characterized by a set of measurements taken under different conditions.

Whatever the reason may be for their employment, however, the repeated measures designs present some methodological problem. The repeated measures data typically contain both within-subject and between-subject variations. Since these two kinds of variations behave differently, they should be separated and treated differently. The between-subject variation, in particular, gives rise to dependencies among observations. Subject parameters are often introduced in order to account for the dependencies. However, this causes another problem. Since the number of parameters to be estimated in-

The work reported in this paper has been supported by Grant A6394 to the first author from the Natural Sciences and Engineering Research Council of Canada.

Requests for reprints should be sent to Yoshio Takane, Department of Psychology, McGill University 1205 Dr. Penfield Ave., Montreal, PQ H3A 1B1 CANADA.

0033-3123/87/0900-8075$00.75/0

creases linearly with the number of observations in this case, the usual asymptotic properties (BAN) of maximum likelihood (ML) or generalized least squares (GLS) estimators never hold (e.g., Andersen, 1980). The subject effect is therefore introduced as a random effect with prescribed distributional properties, and is subsequently marginalized out to obtain the marginal distribution of observations. This distribution is then used to estimate other parameters such as those related to the measurement conditions, the subject distribution, and so forth. The variables on which the subject distribution is defined are often called latent variables, or collectively, latent space.

The marginalization of the subject parameters is straightforward, so far as both within-subject and between-subject effects are assumed to follow the multivariate normal distribution, and the observed data are continuous and multivariate normal. However, a complication arises when the data are categorical, presumed to have been obtained by discretizing continuous multivariate normal processes. For example, in mental test situations subjects' responses may be recorded in ones (pass) and zeroes (fail), where the responses are considered functions of item difficulties and subject abilities assumed to follow the normal distribution. There have been two marginalization techniques in use for such situations. One technique is used in the marginal maximum likelihood (MML) estimation of item parameters in item response theory (IRT), originally proposed by Bock & Lieberman (1970), and subsequently generalized and improved by Bock (1972), Bock and Aitkin (1981) and Thissen (1982). The other technique is used in the factor analysis of discretized variables (FA), an approach initiated by Christoffersson (1975) and later expanded by Muthén (1978, 1983, 1984) and Muthén and Christoffersson (1981). Although they differ in their tradition, IRT and FA cover similar types of categorical data, and thus one may suspect that there is a special relationship between the two approaches. Indeed they are formally equivalent, as has been alluded to recently by several authors (Bartholomew, 1983, 1985; Bock, 1984; Muthén, 1983). In this paper we present a formal proof of the equivalence between the IRT and FA models for a variety of categorical data.

We first discuss the dichotomous case (section 2). The basic result on the dichotomous variables will be extended to the general ordered categorical case in section 3, and to the case of multiple-choice (unordered categorical) data in section 3, and to the case of multiple-choice (unordered categorical) data in section 4. In section 5, IRT formulations of the individual differences pair comparison models (Takane, 1985) are derived, which are equivalent to their original ACOVS (Analysis of Covariance Structures; Jöreskog, 1970) formulations of the same models. A taxonomy of the IRT models and the FA models are also attempted and presented in the final section. The two equivalent approaches, IRT and FA, closely parallel Thurstone's two alternative formulations of his pair comparison model and the model of first choice. This is shown in the appendix.

Throughout this paper a random variable is denoted by a symbol with a tilde on top, and a particular realization of the random variable by the same symbol without a tilde. Scalars are indicated by lowercase italics, vectors by boldface and matrices by uppercase italics. An uppercase letter will also be used for a region of integration, but it will be clear from the context when it is used for this purpose.

2. The Dichotomous Case

We first prove the equivalence for the dichotomous case. Although this is a special case of the general ordered categorical case, and also of the multiple-choice (unordered categorical) case, it deserves special attention because of its predominance in the item response theory.

Let $\tilde{\mathbf{x}}' = (\tilde{x}_1, \ldots, \tilde{x}_n)$ be a random vector of response patterns to n dichotomous test items, where each $\tilde{x}_i$ is defined as

$$\tilde{x}_i = \begin{cases} 1, & \text{if item } i \text{ is successfully passed} \\ 0, & \text{otherwise,} \end{cases}$$

for $i = 1, \ldots, n$. Let $\tilde{\mathbf{u}}$ be an m-component random vector of subject abilities ($m \leq n$) with its density function denoted by $g(\mathbf{u})$. $\tilde{\mathbf{u}}$ is unobservable directly, but is assumed to follow the multivariate normal distribution with mean $\mathbf{0}$ and covariance I (identity matrix); that is, $\tilde{\mathbf{u}} \sim N(\mathbf{0}, (I)$. The domain of $\tilde{\mathbf{u}}$ (denoted by U) is the multidimensional region defined by the direct product of $(-\infty, \infty)$.

The two-parameter normal ogive model in IRT specifies the marginal probability of $\tilde{\mathbf{x}} = \mathbf{x}$ (Bock & Aitkin, 1981; Bock & Lieberman, 1970) as

$$\Pr(\tilde{\mathbf{x}} = \mathbf{x}) = \int_U \Pr(\tilde{\mathbf{x}} = \mathbf{x} \mid \mathbf{u}) g(\mathbf{u})\, d\mathbf{u}, \tag{1}$$

where $\Pr(\tilde{\mathbf{x}} = \mathbf{x} \mid \mathbf{u})$ is the conditional probability of observing response pattern $\mathbf{x}$ given $\tilde{\mathbf{u}} = \mathbf{u}$. $\Pr(\tilde{\mathbf{x}} = \mathbf{x} \mid \mathbf{u})$ is further assumed to be

$$\Pr(\tilde{\mathbf{x}} = \mathbf{x} \mid \mathbf{u}) = \prod_i^n (p_i(\mathbf{u}))^{x_i}(1 - p_i(\mathbf{u}))^{1-x_i} \tag{2}$$

(local independence) with

$$p_i(\mathbf{u}) = \int_{-\infty}^{\mathbf{a}'\mathbf{u}+b} \phi(z)\, dz = \Phi(\mathbf{a}'\mathbf{u} + b), \tag{3}$$

where ϕ is the density function of the standard normal distribution and Φ the normal ogive function (i.e., the cumulative distribution function of the standard normal distribution).

In factor analysis of dichotomized variables (Christoffersson, 1975), on the other hand, the marginal probability of response pattern $\mathbf{x}$ is specified as

$$\Pr(\tilde{\mathbf{x}} = \mathbf{x}) = \int_R h(\mathbf{y})\, d\mathbf{y}, \tag{4}$$

where R is the multidimensional region of integration (to be more explicitly specified below) and

$$\tilde{\mathbf{y}} = C\tilde{\mathbf{u}} + \tilde{\mathbf{e}}. \tag{5}$$

Model (5) is the usual common factor analysis model with C being the matrix of factor loadings, $\tilde{\mathbf{u}}$ the vector of factor scores (which in the present case are the subject abilities) and $\tilde{\mathbf{e}}$ the random vector of uniqueness components. It is assumed that $\tilde{\mathbf{u}} \sim N(\mathbf{0}, I)$ as before, $\tilde{\mathbf{e}} \sim N(\mathbf{0}, Q^2)$ where Q^2 is further assumed to be diagonal (linear local independence), and $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{e}}$ are independent of each other. It follows that

$$\tilde{\mathbf{y}} \sim N(\mathbf{0}, CC' + Q^2), \tag{6}$$

(marginal distribution of $\tilde{\mathbf{y}}$) and

$$\tilde{\mathbf{y}} \mid \mathbf{u} \sim N(C\mathbf{u}, Q^2), \tag{7}$$

(conditional distribution of $\tilde{\mathbf{y}}$ given $\tilde{\mathbf{u}} = \mathbf{u}$). The continuous random variables, $\tilde{\mathbf{y}}$, are dichotomized by

$$\tilde{x}_i = \begin{cases} 1, & \text{if } \tilde{y}_i \geq r_i, \\ 0, & \text{if } \tilde{y}_i < r_i, \end{cases}$$

for $i = 1, \ldots, n$, where r_i is the threshold parameter for variable i. Thus, R, the region of integration above, is the multidimensional parallelopiped defined by the direct product of intervals, R_i for variable i, such that $R_i = (r_i, \infty)$ if $\tilde{x}_i = 1$ and $R_i = (-\infty, r_i)$ if $\tilde{x}_i = 0$.

Now (1) including (2) and (3) is equivalent to (4) with $\tilde{\mathbf{v}}$ defined in (5). We first prove (4)$\rightarrow$(1). From (4) we have

$$\begin{aligned}\Pr(\tilde{\mathbf{x}} = \mathbf{x}) &= \int_R h(\mathbf{y})\, d\mathbf{y} \\ &= \int_R \left(\int_U f(\mathbf{y} \mid \mathbf{u}) g(\mathbf{u})\, d\mathbf{u}\right) d\mathbf{y} \\ &= \int_U g(\mathbf{u}) \left(\int_R f(\tilde{\mathbf{y}} \mid \mathbf{u})\, d\mathbf{y}\right) d\mathbf{u}, \end{aligned} \tag{8}$$

where $f(\tilde{\mathbf{y}} \mid \mathbf{u})$ is the conditional density of $\tilde{\mathbf{y}}$ given $\tilde{\mathbf{u}} = \mathbf{u}$. (Note that (8) is "completely" general in that no distributional assumptions are involved. Even the local independence assumption, so characteristic of the latent variable methods, is not required.) But because of (7) we have

$$\begin{aligned}\int_R f(\tilde{\mathbf{y}} \mid \mathbf{u})\, d\mathbf{y} &= \prod_i \int_{R_i} f_i(y_i \mid \mathbf{u})\, dy_i \\ &= \prod_i \left(\int_{r_i}^{\infty} f_i(y_i \mid \mathbf{u})\, dy_i\right)^{x_i} \left(1 - \int_{r_i}^{\infty} f_i(y_i \mid \mathbf{u})\, dy_i\right)^{1-x_i}, \end{aligned} \tag{9}$$

where

$$\int_{r_i}^{\infty} f_i(y_i \mid \mathbf{u})\, dy_i = \Phi\left(\frac{\mathbf{c}_i'\mathbf{u} - r_i}{q_i}\right), \tag{10}$$

for $i = 1, \ldots, n$. Here q_i^2 is the i-th diagonal element of Q^2. Equation (9) is equivalent to (3) by setting

$$\mathbf{a}_i = \frac{\mathbf{c}_i}{q_i} \tag{11}$$

and

$$b_i = -\frac{r_i}{q_i} \tag{12}$$

for $i = 1, \ldots, n$.

The reverse ((1)$\rightarrow$(4)) can be easily proved by simply tracing back the above process. It looks as if FA with $\mathbf{c}_i$, r_i and $q_i (i = 1, \ldots, n)$ had more parameters than IRT with only $\mathbf{a}_i$ and $b_i (i = 1, \ldots, n)$. However, when the data are dichotomous, the variance of $\tilde{y}_i$ cannot be estimated due to the lack of relevant information in the data, and consequently q_i can be set to an arbitrary value. Thus, the effective number of parameters is identical in the two models.

Lord and Novick (1968, Theorem 16.8.1, p. 374) state a sufficient condition for the two-parameter normal ogive model for unidimensional ability, which may be interpreted as a special case of our general result presented above. More recently Bartholomew (1985) noted the relationship, (11) and (12). See also Muthén (1979, Appendix) and Muthén & Christoffersson (1979, p. 411).

It is clear from the above discussion that IRT and FA are two alternative formulations of a same model. Perhaps because of this Bock (1984) now calls his IRT approach item factor analysis. The only crucial difference is where the marginalization is performed. In the IRT formulation dichotomization of $\tilde{\mathbf{y}}$ is done conditionally on $\mathbf{u}$ and then the marginalization is performed. In the FA tradition, the marginalization is undertaken on continuous $\tilde{\mathbf{y}}$, followed by the dichotomization. An advantage of the IRT formulation is that the dichotomization is relatively straightforward (it can be done separately for each $\tilde{y}_i$ given $\mathbf{u}$ due to the local independence assumption). The probability of a full response pattern can be obtained by a multiple integral of dimensionality m, where m is the dimensionality of the latent space. However, this integration usually involves numerical integration, which may be quite time consuming. In the FA formulation the marginalization is rather trivial, but the dichotomization is extremely difficult. It always involves integration of n correlated multivariate normal variables over n dimensional parallelopiped, no matter what the dimensionality of the latent space is. Thus, most often only one-way and two-way marginal probabilities (i.e., $\Pr(\tilde{x}_i = x_i)$ and $\Pr(\tilde{x}_i = x_i \text{ and } \tilde{x}_j = x_j)$) can be evaluated. These considerations largely determine the choice of optimization criteria in the two approaches. Whereas the IRT formulation uses the maximum likelihood estimation based on the full joint probabilities of response patterns (Bock & Aitkin, 1981; Bock & Lieberman, 1970), the FA approach typically uses a generalized least squares (GLS) estimation based on the first and second order marginal probabilities (Christoffersson, 1975; Muthén, 1978).

In closing of this section it might be noted that the logistic model proposed by Birnbaum (Lord & Novick, 1968) is often used to approximate the normal ogive model (3). The equivalence of marginal probabilities in IRT and FA holds approximately with the logistic model as well, but only to the extent that the logistic distribution provides a good approximation to the normal distribution.

3. The Ordered Categorical Case

So far we have discussed the relationship between IRT and FA for dichotomous data. An analogous relationship is expected to hold for general ordered categorical data, and indeed it can be shown that the marginal likelihood of the normal ogive model for graded scores (Samejima, 1969) is formally equivalent to factor analysis of ordered categorical data recently proposed by Muthén (1984).

Let $\tilde{\mathbf{x}}' = (\tilde{\mathbf{x}}'_1, \ldots, \tilde{\mathbf{x}}'_n)$ be a random vector of response patterns, where $\tilde{\mathbf{x}}'_i$, i-th subvector of $\tilde{\mathbf{x}}$, is an n_i-component vector, $\tilde{\mathbf{x}}'_i = (\tilde{\mathbf{x}}_{i(1)}, \ldots, \tilde{\mathbf{x}}_{i(n_i)})$ with its j-th element defined by

$$\tilde{\mathbf{x}}_{i(j)} = \begin{cases} 1, & \text{if response to item } i \text{ falls in category } j, \\ 0, & \text{otherwise,} \end{cases}$$

for $j = 1, \ldots, n_i$ and $i = 1, \ldots, n$. We assume $\tilde{x}_{i(j)}\tilde{x}_{i(k)} = 0$, for $j \neq k$ and $\sum_j^{n_i} \tilde{x}_{i(j)} = 1$. Note that in the special case of dichotomous variables we have $\tilde{x}_{i(2)} = \tilde{x}_i$ and $\tilde{x}_{i(1)} = 1 - \tilde{x}_i$.

The proof is rather straightforward following the line presented in the previous section for the dichotomous case. The factor analysis model (5) remains the same. Let R be the multivariate region defined as the direct product of intervals $R_i (i = 1, \ldots, n)$, where $R_i = (r_{i(j-1)}, r_{i(j)})$ if $\tilde{x}_{i(j)} = 1$. (Note there is only one $\tilde{x}_{i(j)}$ equal to unity for each i.) Here $r_{i(j)}$ is the category boundary between the $(j-1)$-st and j-th successive categories. We define $r_{i(0)} = -\infty$ and $r_{i(n_i)} = \infty$. Then (8) is still valid with this new definition of R, and

$$\int_R f(\mathbf{y}\,|\,\mathbf{u})\,d\mathbf{y} = \prod_i^n \int_{R_i} f_i(y_i\,|\,\mathbf{u})\,dy_i$$

$$= \prod_i^n \prod_j^{n_i} \left\{\Phi\left(\frac{\mathbf{c}_i'\mathbf{u} - r_{(j-1)}}{q_i}\right) - \Phi\left(\frac{\mathbf{c}_i'\mathbf{u} - r_{i(j)}}{q_i}\right)\right\}^{x_{i(j)}}$$

$$= \prod_i^n \prod_j^{n_i} \left\{\Phi\left(\mathbf{a}_i'\mathbf{u} + b_{i(j-1)}\right) - \Phi\left(\mathbf{a}_i'\mathbf{u} + b_{i(j)}\right)\right\}^{x_{i(j)}}, \tag{13}$$

where $\mathbf{a}_i = \mathbf{c}_i/q_i$ (the same as in (11)) and $b_{i(j)} = -r_{i(j)}/q_i$. ($b_{i(0)} = \infty$ and $b_{i(n)} = -\infty$.) In the dichotomous case $b_{i(0)} = \infty$, $b_{i(1)} = b_i$, and $b_{i(2)} = -\infty$, so that (13) indeed reduces to (9) with (10).

The relationship between IRT and FA for ordered categorical data was first noted by de Leeuw (1983). The above proof generalizes his result to the multidimensional case.

As in the dichotomous case the logistic function is often used (e.g., Cox, 1966; Samejima, 1969; and Takane, 1983a, with some minor modifications) in place of the normal ogive model in (13). Again, the approximate relationship holds between FA and the logistic IRT model for ordered categorical data to the extent that the logistic distribution provides a good approximation to (13).

4. The Unordered Categorical Case

The case of unordered categorical data is slightly more complicated. First of all there is no factor analysis approach ever proposed for this case, although there have been a couple of significant proposals in the IRT approach (which will be discussed briefly toward the end of this section). Secondly, an element of $\tilde{\mathbf{y}}$ should be supplied for each nominal category of each item. That is, $\tilde{\mathbf{y}}' = (\tilde{\mathbf{y}}_1', \ldots, \tilde{\mathbf{y}}_n')$ where each $\tilde{\mathbf{y}}_i = (\tilde{\mathbf{y}}_{i(1)}, \ldots, \tilde{\mathbf{y}}_{i(n_i)})$ is an n_i-component vector. The factor analysis model is now written as

$$\tilde{\mathbf{y}} = \mathbf{m} + C\tilde{\mathbf{u}} + \tilde{\mathbf{e}}, \tag{14}$$

where $\mathbf{m}$, the mean vector, and C are partitioned in the same way as is $\tilde{\mathbf{y}}$. That is, $\mathbf{m}' = (\mathbf{m}_1', \ldots, \mathbf{m}_n')$ where $\mathbf{m}_i' = (m_{i(1)}, \ldots, m_{i(n_i)})$, and $C' = (C_1', \ldots, C_n')$ where $C_i' = (\mathbf{c}_{i(1)}, \ldots, \mathbf{c}_{i(n_i)})$. Without loss of generality we may assume $\mathbf{m}_i'\mathbf{1}_{n_i} = 0$ and $C_i'\mathbf{1}_{n_i} = \mathbf{0}$ for each i, where $\mathbf{1}_{n_i}$ is the n_i-component vector of ones. These restrictions remove indeterminacies of origin in $\mathbf{m}_i$ and C_i for each i. In the ordered categorical case there was only one $\tilde{y}_i$ for each item, and no comparison among $\tilde{y}_i$'s was involved. Consequently $\mathbf{m}$ could be set identically equal to a zero vector. The response pattern vector $\mathbf{x}$ has the identical form as that in the ordered categorical case.

Let $R_{i(j)}$ be the region such that $\tilde{y}_{i(j)} = \max(\tilde{y}_{i(1)}, \ldots, \tilde{y}_{i(n_i)})$ and $R_i = R_{i(j)}$ if $\tilde{x}_{i(j)} = 1$. Let R be the region defined by the direct product of R_i. Then (8) is still valid with

$$\int_R f(\mathbf{y}\,|\,\mathbf{u})\,d\mathbf{y} = \prod_i \int_{R_i} f_i(\mathbf{y}_i\,|\,\mathbf{u})\,d\mathbf{y}_i$$

$$= \prod_i \prod_j \left(\int_{R_{i(j)}} f_i(\mathbf{y}_i\,|\,\mathbf{u})\,d\mathbf{y}_i\right)^{x_{i(j)}} \tag{15}$$

Note that in this case $R_{i(j)}$ is not a parallelopiped, but a cone. However, it can be transformed into a rectangular region, using Lemma 2 in the appendix.

The same transformation is very effective in showing that when there are only two response categories, the unordered categorical case reduces to the dichotomous case discussed earlier. Let $\tilde{\mathbf{z}}' = (\tilde{z}_1, \ldots, \tilde{z}_n)$ where $\tilde{z}_i = \tilde{y}_{i(2)} - \tilde{y}_{i(1)}$. Then by Lemma 1 in the appendix,

$$\Pr(\tilde{\mathbf{x}} = \mathbf{x}) = \int_R h(\mathbf{y})\, d\mathbf{y}$$
$$= \int_{R^*} h^*(\mathbf{z})\, d\mathbf{z}, \tag{16}$$

where R^* is a multidimensional parallelopiped defined by the direct product of intervals, $R_i^*(i = 1, \ldots, n)$, where $R_i^* = (-\infty, 0)$ if $\tilde{x}_{i(1)} = 1 - \tilde{x}_i = 1$ and $R_i^* = (0, \infty)$ if $\tilde{x}_{i(2)} = \tilde{x}_i = 1$. Furthermore, in (15)

$$\int_{R_i} f_i(\mathbf{y}_i \mid \mathbf{u})\, d\mathbf{y}_i = \left(\int_{R_{i(2)}} f_i(\mathbf{y}_i \mid \mathbf{u})\, d\mathbf{y}_i\right)^{x_{i(2)}} \left(\int_{R_{i(1)}} f_i(\mathbf{y}_i \mid \mathbf{u})\, d\mathbf{y}_i\right)^{x_{i(1)}}$$
$$= \left(\int_0^{\infty} f_i(z_i \mid \mathbf{u})\, dz_i\right)^{x_i} \left(\int_{-\infty}^{0} f_i(z_i \mid \mathbf{u})\, dz_i\right)^{1 - x_i}, \tag{17}$$

where

$$\tilde{z}_i \mid \mathbf{u} \sim N((m_{i(2)} - m_{i(1)}) + (\mathbf{c}_{i(2)} - \mathbf{c}_{i(1)})'\mathbf{u},\ q_{i(1)}^2 + q_{i(2)}^2).$$

Thus, we obtain

$$\int_0^{\infty} h_i^*(z_i \mid \mathbf{u})\, dz_i = \Phi\left(\frac{\mathbf{c}_i'\mathbf{u} - r_i}{q_i}\right) \tag{18}$$

by setting $r_i = m_{i(1)} - m_{i(2)} = 2m_{i(1)}$, $\mathbf{c}_i = \mathbf{c}_{(2)} - \mathbf{c}_{i(1)} = -2\mathbf{c}_{i(1)}$ and $q_i^2 = q_{i(1)}^2 + q_{i(2)}^2$.

When the multivariate normal variates to be integrated over R_i are mutually independent with homogeneous variance, there is an excellent approximation method provided by the multivariate logistic function (Bock, 1975). (Note that $\tilde{y}_{i(j)} \mid \mathbf{u}, j = 1, \ldots, n_i$, are mutually independent, but their variances, $q_{i(j)}^2$, are generally not equal under the usual factor analysis assumptions. Thus, we need a more strict assumption of $q_{i(j)}^2 = q_i^2$ for all j for this approximation to be valid.) Namely,

$$p_{i(j)}(\mathbf{u}) = \Pr(\tilde{y}_{i(j)} = \max(\tilde{y}_{i(1)}, \ldots, \tilde{y}_{i(n_i)} \mid \mathbf{u}))$$
$$= \frac{\exp(\mathbf{a}_{i(j)}^{*\prime}\mathbf{u} + b_{i(j)}^*)}{\sum_k \exp(\mathbf{a}_{i(k)}^{*\prime}\mathbf{u} + b_{i(k)}^*)}, \tag{19}$$

where $\mathbf{a}_{i(j)}^*$ and $b_{i(j)}^*$ are approximately proportional to $\mathbf{c}_{i(j)}$ and $m_{i(j)}$, respectively. It is interesting to note that (19) provides a multidimensional generalization of Bock's (1972) unidimensional IRT model for unordered categorical data. It also generalizes Takane's (1983b) multivariate logistic unfolding model, which states

$$p_{i(j)}(\mathbf{u}) = \frac{\exp(-d_{i(j)}^2(\mathbf{u}))}{\sum_k \exp(-d_{i(k)}^2(\mathbf{u}))}, \tag{20}$$

where $d_{i(j)}^2 = (\mathbf{v}_{i(j)} - \mathbf{u})'(\mathbf{v}_{i(j)} - \mathbf{u})$ is the squared euclidean distance between $\mathbf{u}$ and the point representing category j of item i, whose coordinates are given by $\mathbf{v}_{i(j)}$. If we introduce a bias parameter, $w_{i(j)}$, for category j of item i, and replace $\exp(-d_{i(j)}^2(\mathbf{u}))$ by $w_{i(j)}$

$\exp(-d^2_{i(j)(u)})$, $j = 1, \ldots, n_i$, in (20), (20) will be identical to (19) with $\mathbf{a}_{i(j)} = 2\mathbf{v}_{i(j)}$ and $b_{i(j)} = \ln w_{i(j)} - \mathbf{v}'_{i(j)}\mathbf{v}_{i(j)}$. ($\mathbf{u}'\mathbf{u}$ cancels out in the numerator and the denominator.) Model (20) is considered as a combination of Coombs' (1964) unfolding model for preference data and Luce's (1959) choice model. Alternatively it can be viewed as a probabilistic generalization of dual scaling (Nishisato, 1980), homogeneity analysis (Gifi, 1981) or multiple correspondence analysis (Greenacre, 1984), which in turn is a special case of the unfolding model (Takane, 1980a; Heiser, 1981). It can be also regarded as a multiple-choice (as opposed to binary-choice) extension of Schönemann-Wang's (1972) individual difference preference model. Both Bock (1972) and Takane (1983b) proposed marginal maximum likelihood estimation methods for their models.

5. Individual Differences Pair Comparison Models

Takane (1985) recently extended the "factorial" model (Takane, 1980*b*; Heiser & de Leeuw, 1981) and the wandering vector model (WVM; De Soete & Carroll, 1983) for pair comparison data to accommodate systematic individual differences in these models. He first introduced random vectors pertaining to the systematic individual differences (analogous to $\tilde{\mathbf{u}}$), and then marginalized them out to arrive at ACOVS (Jöreskog, 1970) formulations of these models. This clearly belongs to the FA approach. However, equivalent IRT formulations are also possible. Pair comparison data can be viewed as a special case of unordered categorical data with only two response categories, where the two categories are two stimuli to be compared. (Either stimulus A is chosen or B is chosen.) As such, they can be also considered as a special type of dichotomous data. (A stimulus is chosen or not chosen.) Peculiarity of the pair comparison data stems from the fact that the response categories (stimuli) are not nested within items (trials). The same stimuli repeatedly appear in different combinations.

Let $\tilde{\mathbf{x}}' = (\tilde{x}_{12}, \ldots, \tilde{x}_{(n-1)n})$ be a random vector of choice patterns, where

$$\tilde{x}_{ij} = \begin{cases} 1, & \text{if stimulus } i \text{ is chosen over stimulus } j, \\ 0, & \text{otherwise,} \end{cases}$$

for $i = 1, \ldots, n - 1$ and $j = i + 1, \ldots, n$. (n is the number of stimuli). For the "factorial" model of pair comparisons, let

$$\tilde{\mathbf{y}}^* = A\tilde{\mathbf{y}} + \tilde{\mathbf{e}}^* = A(\tilde{\mathbf{s}} + C\tilde{\mathbf{u}}) + \tilde{\mathbf{e}}^* \tag{21}$$

be the second order FA (or ACOVS) model, where $\tilde{\mathbf{s}} = \mathbf{m} + \tilde{\mathbf{e}} \sim N(\mathbf{m}, Q^2)$, A is the $n(n-1)/2$ by n design matrix for pair comparisons, and $\tilde{\mathbf{e}}^* \sim N(\mathbf{0}, K^2)$ is the error random vector for pair comparison trials with K^2 diagonal. Matrix A takes the difference between $\tilde{y}_i$ and $\tilde{y}_j$ for every distinct combination of stimuli i and j. $\tilde{\mathbf{e}}^*$ represents within-subject variation. The reason for this additional error term (over and above $\tilde{\mathbf{e}}$ which is part of $\tilde{\mathbf{s}}$) is that in the multiple-judgment pair comparison situation, a same stimulus is presented more than once to a same subject, and consequently a new error term is required that accounts for within-subject, across-trial variation. We then have

$$\tilde{\mathbf{y}}^* \sim N(A\mathbf{m}, A(CC' + Q^2)A' + K^2),$$

and

$$\tilde{\mathbf{y}}^* \mid \mathbf{w} \sim N(A(\mathbf{s} + C\mathbf{u}), K^2),$$

where $\mathbf{w}' = (\mathbf{s}', \mathbf{u}')$, for the "factorial" model.

In the WVM it is further assumed that $\mathbf{m} = C\mathbf{v}$, where $\mathbf{v}$ is the mean of the wandering vector, $\tilde{\mathbf{u}}^* = \tilde{\mathbf{u}} + \mathbf{v}$. We thus obtain the ACOVS formulation of the WVM as

$$\tilde{\mathbf{y}}^* = A(\tilde{\mathbf{e}} + C\tilde{\mathbf{u}}^*) + \tilde{\mathbf{e}}^*.$$

The marginal distribution of $\tilde{\mathbf{y}}^*$ is given by

$$\tilde{\mathbf{y}}^* \sim N(AC\mathbf{v}, A(CC' + Q^2)A' + K^2).$$

Similarly, the conditional distribution of $\tilde{\mathbf{y}}^*$ given $\mathbf{w}$, where $\mathbf{w}' = (\mathbf{e}', \mathbf{u}^{*\prime})$, is given by

$$\tilde{\mathbf{y}}^* \mid \mathbf{w} \sim N(A(\mathbf{e} + C\mathbf{u}^*), K^2).$$

Let R denote the multidimensional region defined by the direct product of unidimensional intervals, R_{ij}, which are either $(-\infty, 0)$ or $(0, \infty)$ depending on $\tilde{x}_{ij} = 0$ or 1, respectively. Then (8) is still valid with $\mathbf{y}$ and $\mathbf{u}$ in (8) replaced by $\mathbf{y}^*$ and $\mathbf{w}$, respectively, in both factorial and WVM. In the present case

$$\begin{aligned}\int_R f(\mathbf{y}^* \mid \mathbf{w})\, d\mathbf{y}^* &= \prod_{i,j} \int_{R_{ij}} f_{ij}(y_{ij}^* \mid \mathbf{w})\, dy_{ij}^* \\ &= \prod_{i,j} \left(\int_0^\infty f_{ij}(y_{ij}^* \mid \mathbf{w})\, dy_{ij}^* \right)^{x_{ij}} \left(1 - \int_0^\infty f_{ij}(y_{ij}^* \mid \mathbf{w})\, dy_{ij}^* \right)^{1-x_{ij}}. \end{aligned} \tag{23}$$

In the factorial model, we have

$$\int_0^\infty f_{ij}(y_{ij}^* \mid \mathbf{w})\, dy_{ij}^* = \Phi\left(\frac{\mathbf{a}_{ij}'(\mathbf{s} + C\mathbf{u})}{k_{ij}} \right), \tag{24}$$

whereas in the WVM, we have

$$\int_0^\infty f_{ij}(y_{ij}^* \mid \mathbf{w})\, dy_{ij}^* = \Phi\left(\frac{\mathbf{a}_{ij}'(\mathbf{e} + C\mathbf{u}^*)}{k_{ij}} \right), \tag{25}$$

where $\mathbf{a}_{ij}'$ is the ij-th row vector of A and k_{ij} the ij-th diagonal element of K^2. Equation (23) along with (24) or (25) used in (8) provides the IRT formulation of the "factorial" model or the WVM.

Both the ACOVS and the IRT formulations of these two models can be easily generalized into ordered categorical ratings of pair comparisons (Sjoberg, 1967), although this case will not be discussed any further in this paper. De Soete, Carroll & DeSarbo (in press) recently proposed the wandering ideal point model based on Coombs' unfolding model. The model is conceptually similar to the WVM, and is applied to the same kind of pair comparison data. The ACOVS and the IRT formulations of this model is possible using squared euclidean distances (Takane, 1985). In fact they reduce to forms similar to those for the WVM, since the difference between two squared euclidean distances from a common ideal point reduces to a scalar product.

Some attempt has been made to incorporate systematic individual differences into Thurstone's pair comparison model (Bock & Jones, 1968, p. 143–161). This attempt belongs to the ACOVS approach. However, it is confined to the simplest possible covariance structure, namely equal variances and covariances. This corresponds with $K^2 = 0$, $CC' = d^2 r \mathbf{1}\ \mathbf{1}'$ and $Q^2 = d(1 - r)I$ in the "factorial" model, where d^2 and r are, respectively, the variance and correlation (assumed equal across stimuli).

6. Discussion

There are numerous instances of psychometric models involving subject parameters. This is because there are almost always some degree of individual differences in every psychological phenomenon, and so far as one uses repeated-measures designs, one cannot get away from the problem of dealing with the systematic individual differences. This means that it is almost impossible to develop realistic models without incorporating individual differences components into the models (Takane, 1985).

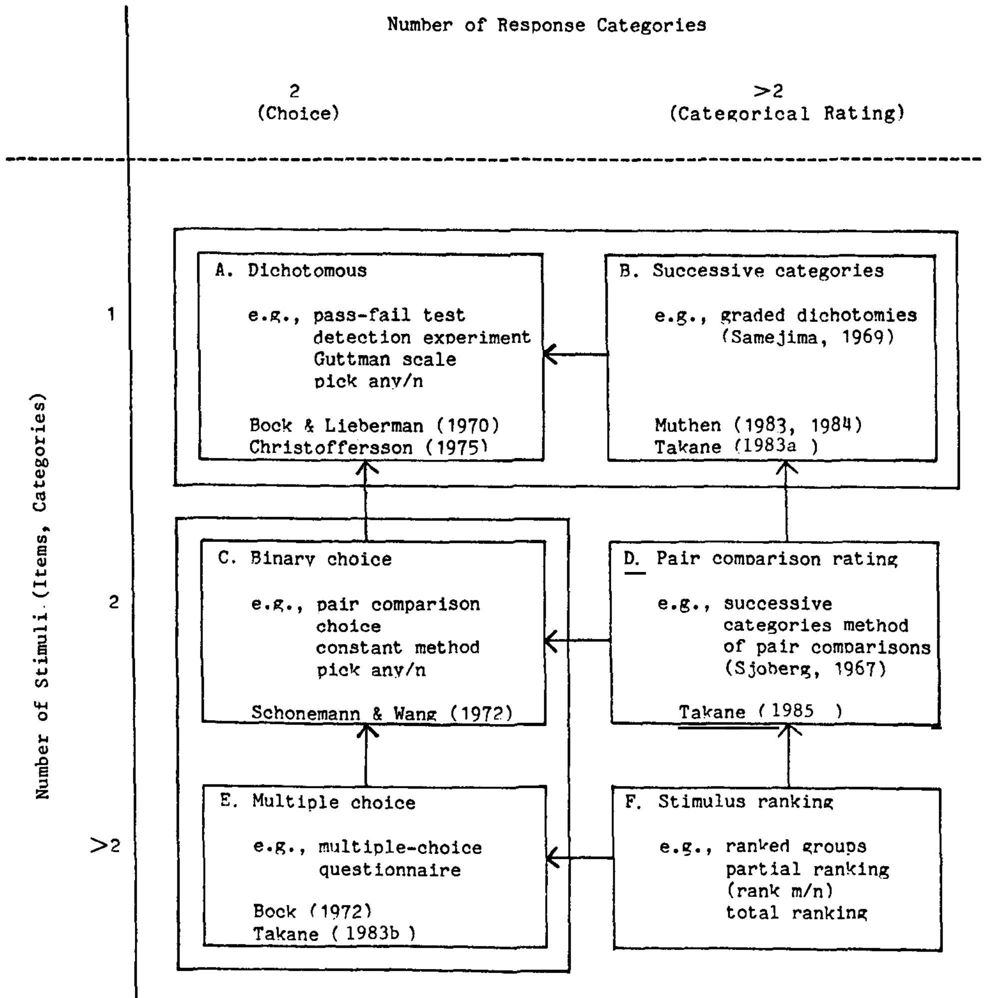
However, as argued in the introduction section, the subject parameters cause some difficulties in parameter estimation from a statistical viewpoint. Since they are incidental to the subjects, the number of parameters to be estimated generally increases with the number of observations, and consequently the asymptotic properties of ML or GLS estimators never hold. Marginalization of the subject parameters has been one of the major techniques to deal with the problem. (For other alternatives, see Basu, 1977.) This usually involves integrating out the subject parameters by assuming a population distribution on the subject parameters. Then the marginal probability of observed data is stated solely as a function of other (nonincidental) parameters related to, for example, stimuli, items, categories, and so forth. For a large sample the ML or GLS estimators (obtained through the marginal probability of the observed data) enjoy the usual asymptotic properties (BAN) of these estimators, provided that the model is correct, including the distributional assumption made on the subject parameters. Latent structure analysis (LSA; Lazarsfeld & Henry, 1968) and factor analysis are two classical examples of the marginalization model. The marginalization is indeed an intrinsic part of these models. Essentially the same approach has been proposed from a Bayesian perspective (Akaike, 1980) under the name "Bayesian modeling." Here the Bayesian predictive probability is maximized with respect to Bayesian hyperparameters (which correspond to our nonincidental, fixed-effect parameters.) This approach was proposed to deal with a large number of model parameters (which are not necessarily the subject parameters) and to incorporate certain desirable properties into the parameters.

Psychological mental testing situations are the ones in which individual differences are most pertinent. This in fact served as the basic motivation behind the IRT models developed for the mental testing situations. Curiously, however, it was not until 1970 (Bock & Lieberman, 1970) that the marginal maximum likelihood (MML) estimation was proposed for the IRT models. This is quite a contrast to LSA and FA, which included the marginalization as part of the models. This probably reflects a difference in the initial interest of these approaches. Whereas in IRT each subject's score was of primary interest, in LSA and FA how observations (response patterns) distributed in the population of subjects was the focus of interest. However, the MML estimation proposed to deal with inconsistent estimators in IRT has brought the two approaches together.

When the individual differences (in ability, attitude, preference, personality, etc.) are of interest, we may use estimates of structural (nonincidental) parameters obtained by MML to obtain EAP (expected *aposteriori* or the Bayes) or MAP (maximum *aposteriori* or the Bayes modal) estimators of the individual differences parameters (Bock & Aitkin, 1981). This corresponds with the estimation of factor scores in FA.

There are other models in psychometrics for which the marginalization may be useful. The unfolding model is designed to account for individual differences in preference. In this case coordinates of ideal points of subjects appear as incidental parameters. Perhaps Takane (1983a, 1983b) was the first to point out the necessity of treating the subject parameters as random effects, and to demonstrate the feasibility of MML in two specialized cases of the multidimensional unfolding model, drawing close relationships between his cases and the IRT test models. More recently Takane (1985) proposed the MML estimation for pair comparison models that take into account systematic individual differences. (See section 5.) A similar formulation is also possible for the wandering

Table 1. A taxonomy of data for the IRT and FA models



ideal point model recently proposed by De Soete, Carroll & DeSarbo (in press). Individual difference multidimensional scaling (Carroll & Chang, 1970) is another potential area to which the MML estimation might be effectively applied. As has been demonstrated by Weinberg, Carroll and Cohen (1984), the usual asymptotic results are too optimistic in this case. In a Bayesian framework Ramsay (1982) has proposed the marginalization of subject-specific data transformation parameters in his maximum likelihood multidimensional scaling.

The data and the models discussed in this paper are summarized in Table 1. The classification was made in terms of two criteria: (a) number of stimuli presented to the subject in each trial where the stimuli may be test items, categories (mostly unordered), et cetera, and (b) Number of response categories (usually ordered). There can be one, two, or more than two stimuli presented, which are either chosen (or not chosen), or rated (or ranked). Thus by combining the two criteria six data types emerge: (A) dichotomous; (B) successive categories; (C) binary choice; (D) pair comparison rating; (E) multiple choice; and (F) stimulus ranking. For each data type some representative cases (or situations in which the specific type of data are typically obtained) are given along with references to models designed for the particular type of data. Except for Schönemann-Wang's (1972) model, the models cited are restricted to those having marginalization elements in the models.

In the table an arrow indicates the data type at its tail is a generalization of the data type placed at the head of the arrow. For example, successive categories data (B) reduce to dichotomous data (A) when there are only two observation categories. Likewise multiple choice data (E) reduce to binary choice data (C) when there are only two alternatives to choose from. Models designed for more general types of data are usually applicable to more special types of data as well. When choice alternatives are nested within trials the binary choice data (C) are equivalent to the dichotomous data (A). This is because two unordered categories can be always arbitrarily ordered to obtain two ordered categories. (See also section 4.)

In the table (A) and (B) are the data types usually referred to as ordered categorical data, and (C) and (E) unordered (or nominal) categorical data. This distinction closely parallels the analogous distinction in models, similar to Bock's (1975) distinction of threshold and extremal concepts. In the former stimuli are supposedly compared with thresholds (or category boundaries), while in the latter "stimuli" are compared against other "stimuli." Takane (1983a) has shown that pick any/n data can be conceptualized in either way. That is, in the first approach it is assumed that preference of a stimulus is compared against a threshold, while in the latter relative strengths of two possible responses (pick or not pick) are compared against each other to determine if the stimulus is picked or not.

Relatively little attention has been paid to the two remaining data types (D and F). No marginalization models have yet been proposed for stimulus ranking. However, some plausible models may be developed for this case, using Takane & Carroll's (1981) directional ranking idea. This, however, is left to future investigations.

In this paper we have shown the equivalence of two marginalization techniques used to obtain marginal probabilities of observed categorical data in two related areas, item response test theory and factor analysis. Although useful exchanges of ideas and interplay between these two areas have already begun (e.g., Mislevy, in press), we hope this paper further facilitates this welcome trend.

Appendix

In this appendix we first give a couple of useful lemmas. We then show that two alternative formulations of the pair comparison model and the model of first choice by Thurstone (1927, 1945) closely parallel the two approaches (IRT and FA) we have discussed in this paper.

Lemma 1. Let $\tilde{u}$ and $\tilde{v}$ be two continuous random variables, each ranging from $-\infty$ to ∞, with their joint density function denoted by $f^*(u, v)$. Then

$$\Pr(\tilde{u} > \tilde{v}) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{u} f^*(u, v)\, dv \right) du = \int_{-\infty}^{\infty} g_U(u) \left(\int_{-\infty}^{u} f_{V|U}(v \mid u)\, dv \right) du$$

$$= \int_{-\infty}^{\infty} \left(\int_{v}^{\infty} f^*(u, v)\, du \right) dv = \int_{-\infty}^{\infty} g_V(v) \left(\int_{v}^{\infty} f_{U|V}(u \mid v)\, du \right) dv$$

$$= \int_0^\infty h(w)\,dw = \Pr(\tilde{u} - \tilde{v} > 0),$$

where $w = u - v$, and where g_U and g_V are marginal densities of $\tilde{u}$ and $\tilde{v}$, respectively, and $f_{U|V}$ and $f_{V|U}$ are conditional densities of $\tilde{u}$ and $\tilde{v}$, respectively, given v and u, respectively.

A proof of the above lemma is rather rudimentary, and will not be presented here. Note that Lemma 1 does not require either $\tilde{u}$ or $\tilde{v}$ to be normal. There are two important special cases to Lemma 1.

Corollary 1. Suppose $\tilde{u}$ and $\tilde{v}$ are independent. Then $f^*(u, v) = g_U(u) \times g_V(v)$. We then have

$$\Pr(\tilde{u} > \tilde{v}) = \int_{-\infty}^{\infty} g_U(u) G_V(u)\,du,$$

where

$$G_V(u) = \int_{-\infty}^{u} g_V(v)\,dv.$$

Corollary 2. Suppose $\tilde{u}$ and $\tilde{v}$ follow a bivariate normal distribution, namely,

$$\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} \sim N\left[\begin{pmatrix} m_u \\ m_v \end{pmatrix}\begin{pmatrix} s_u & s_{uv} \\ s_{uv} & s_v \end{pmatrix}\right],$$

then

$$\Pr(\tilde{u} > \tilde{v}) = \int_{-q}^{\infty} \Phi(z)\,dz,$$

where Φ is the density function of the standard normal distribution, and $q = (m_u - m_v)/(s_u + s_v - 2s_{uv})^{1/2}$.

Both Corollaries 1 and 2 were used by Thurstone. Corollary 1 was used for prediction of first choice (Thurstone, 1945), and Corollary 2 is the well known law of comparative judgment (Thurstone, 1927). Although Thurstone did not explicitly discuss the relationship of the two, Lemma 1 shows their equivalence when the conditions of both corollaries are simultaneously met.

A generalization of Lemma 1 to the multiple-choice situation is rather straightforward, which we state in Lemma 2.

Lemma 2. Let $\tilde{u}, \tilde{v}_1, \ldots, \tilde{v}_n$ be continuous random variables, each ranging from $-\infty$ to ∞, with their joint density function denoted by $f^*(u, v_1, \ldots, v_n)$. Then

$\Pr(\tilde{u} > \tilde{v}_1, \ldots, \tilde{u} > \tilde{v}_n)$

$$= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{u} \cdots \int_{-\infty}^{u} f^*(u, v_1, \ldots, v_n)\,dv_1 \cdots dv_n\right) du$$

$$= \int_{-\infty}^{\infty} g(u)\left(\int_{-\infty}^{u} \cdots \int_{-\infty}^{u} f(v_1, \ldots, v_n \mid u)\,dv_1 \cdots dv_n\right) du$$

$$= \int_0^\infty \cdots \int_0^\infty h(w_1, \ldots, w_n)\,dw_1 \cdots dw_n,$$

where $w_i = u - v_i (i = 1, \ldots, n)$, and where g and f are the marginal density of $\tilde{u}$ and the conditional density of $\tilde{v}_1, \ldots, \tilde{v}_n$ given u, respectively.

Lemma 2 is often used as two alternative formulations of the model of first choice. With the additional assumption of multivariate normality Lemma 2 is sometimes used as a reduction formula for multivariate normal integrals with certain patterned covariance matrices (Johnson & Kotz, 1974, pp. 43–53). Lemma 2 also indicates that there are two ways to evaluate a probability like $\Pr(\tilde{u} > \tilde{v}_1, \tilde{u} < \tilde{v}_2)$, namely,

$$\Pr(\tilde{u} > \tilde{v}_1, \tilde{u} < \tilde{v}_2) = \int_{-\infty}^{0} \int_{0}^{\infty} h(w_1, w_2)\, dw_1\, dw_2$$

$$= \int_{-\infty}^{\infty} g(u) \left(\int_{u}^{\infty} \int_{-\infty}^{u} f(v_1, v_2 \mid u)\, dv_1\, dv_2 \right) du.$$

The two ways of evaluating the probability of the above form correspond with the two ways of evaluating the marginal probability of a response pattern described in the main body of this paper. We will show this only for dichotomous data in the following. Generalizations of this argument to other situations are rather trivial.

Let $\tilde{v}_i$ be the random variable representing item difficulty of item i, where it is assumed that $\tilde{v}_i \sim N(r_i, q_i^2)$, $i = 1, \ldots, n$, independent of each other. Let $\tilde{u}_i = \mathbf{c}_i'\tilde{\mathbf{u}}$ be the random vector of subject ability relevant to item i. As before, we assume $\tilde{\mathbf{u}} \sim N(\mathbf{0}, I)$. We further assume $\tilde{u}_i$ and $\tilde{v}_i$ are independent. In accordance with the pair comparison model we may assume $\tilde{x}_i = 1$ when $\tilde{u}_i > \tilde{v}_i$, and $\tilde{x}_i = 0$ when $\tilde{u}_i < \tilde{v}_i$. Define $\tilde{w}_i = \tilde{y}_i - r_i$ where $\tilde{y}_i$ is the ith variable (pertaining to item i) in the factor analysis model (5). Then,

$$\Pr(\tilde{\mathbf{x}} = \mathbf{x}) = \int_W h^*(\mathbf{w})\, d\mathbf{w},$$

where $\mathbf{w}$ is the vector of w_i's and W is the multidimensional region defined by the direct product of W_i, which is obtained by downshifting R_i by r_i (i.e., $W_i = (-\infty, 0)$ if $\tilde{x}_i = 0$ and $W_i = (0, \infty)$ if $\tilde{x}_i = 1$). This is equivalent to (4). The difference formulation in Lemma 1 thus corresponds with the FA approach.

If, on the other hand, we use the conditional formulation in Lemma 1, we have

$$\Pr(\tilde{\mathbf{x}} = \mathbf{x}) = \int_{U^*} g^*(\mathbf{u}^*) \left(\int_V f(\mathbf{v})\, d\mathbf{v} \right) d\mathbf{u}^*$$

$$= \int_U g(\mathbf{u}) \left(\int_V f(\mathbf{v})\, d\mathbf{v} \right) d\mathbf{u}$$

where $\mathbf{u}^*$ and $\mathbf{v}$ are vectors of u_i and v_i, respectively, and where

$$\int_V f(\mathbf{v})\, d\mathbf{v} = \prod_i \int_{V_i} f_i(v_i)\, dv_i.$$

$$= \prod_i \left\{ \Phi\left(\frac{\mathbf{c}_i'\mathbf{u} - r_i}{q_i} \right) \right\}^{x_i} \left\{ 1 - \Phi\left(\frac{\mathbf{c}_i'\mathbf{u} - r_i}{q_i} \right) \right\}^{1 - x_i}$$

which is equivalent to (9) with (10). This corresponds with the IRT approach.

References

Akaike, H. (1980). Likelihood and the Bayes procedure. In J. M. Bernardo et al. (Eds.), *Bayesian statistics*. Valencia University Press.

Andersen, E. B. (1980). *Discrete statistical models with social science applications*. Amsterdam: North-Holland.

Basu, D. (1977). On elimination of nuisance parameters. *Journal of the American Statistical Association, 22,* 229–243.

Bartholomew, D. J. (1983). Latent variable models for ordered categorical data. *Journal of Econometrics. 22,* 229–243.

Bartholomew, D. J. (1985, July). A unified view of factor analysis, latent structure analysis and scaling. Invited paper given at the 4-th European Meeting of the Psychometric Society and Classification Societies, Cambridge.

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika, 37,* 29–51.

Bock, R. D. (1975). *Multivariate statistical methods in behavioral research.* New York: McGraw Hill.

Bock, R. D. (1984, June). *Full information item factor analysis.* Paper presented at the Joint Meeting of the Classification Society and the Psychometric Society, Sant Barbara.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika, 46,* 443–459.

Bock, R. D., & Jones, L. V. (1968). *The measurement and prediction of judgment and choice.* San Francisco: Holden Day.

Bock, R. D., & Lieberman, M. (1970). Fitting a response model for *n* dichotomously scored items. *Psychometrika, 35,* 283–319.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an *N*-way generalization of "Eckart-Young" decomposition. *Psychometrika, 35,* 283–319.

Christoffersson, A. (1975). Factor analysis of dichotomized variables. *Psychometrika, 40,* 5–32.

Coombs, C. H. (1964). *A theory of data.* New York: Wiley.

Cox, D. R. (1966). Some procedure connected with the logistic qualitative response curve. In F. N. David (Ed.), *Research papers in statistics: Festschrift for J. Neyman* (pp. 55–71). New York: Wiley.

de Leeuw, J. (1983). Models and methods for the analysis of correlation coefficients. *Journal of Econometrics, 22,* 113–137.

De Soete, G., & Carroll, J. D. (1983). A maximum likelihood method for fitting the wandering vector model. *Psychometrika, 48,* 553–566.

De Soete, G., Carroll, J. D., & DeSarbo, W. S. (in press). The wandering ideal point model: A probabilistic multidimensional unfolding model for paired comparisons data. *Journal of Mathematical Psychology.*

Gifi, A. (1981). *Non-linear multivariate analysis.* Unpublished manuscript, University of Leiden, Department of Data Theory.

Greenacre, M. (1984). *Theory and application of correspondence analysis.* London: Academic Press.

Heiser, W. J. (1981). *Unfolding analysis of proximity data.* Unpublished doctoral Dissertation, University of Leiden.

Heiser, W., & de Leeuw, J. (1981). Multidimensional mapping of preference data. *Mathematique et Sciences Humaines, 19,* 39–96.

Johnson, N. L., & Kotz, S. (1974). *Distribution in Statistics: Multivariate distributions.* New York: Wiley.

Jöreskog, K. G. (1970). A general method for analysis of covariance structures. *Biometrika, 57,* 239–251.

Lazarsfeld, P. F., & Henry, N. (1968). *Latent structure analysis.* New York: Houghton Mifflin.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison Wesley.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis.* New York: Wiley.

Mislevy, R. J. (in press). Recent developments in factor analysis of categorical variables. *Journal of Educational Statistics.*

Muthén, B. (1978). Contributions to factor analysis of dichotomous variables. *Psychometrika, 43,* 551–560.

Muthén, B. (1979). A structural probit model with latent variables. *Journal of the American Statistical Association, 24,* 807–811.

Muthén, B. (1983). Latent variable structural equation modeling with categorical data. *Journal of Econometrics, 22,* 43–65.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika, 49,* 115–132.

Muthén, B., & Christoffersson, A. (1981). Simultaneous factor analysis of dichotomous variables in several groups. *Psychometrika, 46,* 407–419.

Ramsay, J. O. (1982). Some statistical approaches to multidimensional scaling. *Journal of the Royal Statistical Society, Series A, 145,* 285–312.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph No. 17, 34*(4, Pt. 2).

Schönemann, P. H., & Wang, M. M. (1972). An individual difference model for the multidimensional analysis of preference data. *Psychometrika, 37,* 275–309.

Sjoberg, L. (1967). Successive categories scaling of paired comparisons. *Psychometrika, 32*, 297–308.

Takane, Y. (1980a). Analysis of categorizing behavior by a quantification method. *Behaviormetrika, 8*, 75–86.

Takane, Y. (1980b). Maximum likelihood estimation in the generalized case of Thurstone's model of comparative judgment. *Japanese Psychological Rsearch, 22*, 188–196.

Takane, Y. (1983a, July). Choice model analysis of the "pick any/*n*" type of binary data. Handout for the talk given at the European Psychometric and Classification Meeting, Jouy-en-Josas.

Takane, Y. (1983b, June). An item response model for multidimensional analysis of multiple-choice data. Handbook for the talk given at the Annual Meeting of the Psychometric Society, Los Angeles.

Takane, Y. (1985, June). Probabilistic multidimensional pair comparison models that take into account systematic individual differences. Transcript of the talk given at the 50-th Anniversary Meeting of the Psychometric Society, Nashville, TN.

Takane, Y., & Carroll, J. D. (1981). Nonmetric maximum likelihood multidimensional scaling from directional rankings of similarities. *Psychometrika, 46*, 389–405.

Thissen, D. (1982). Marginal maximum likelihood estimation for the one-parameter logistic model. *Psychometrika, 47*, 175–186.

Thurstone, L. L. (1927). A law of comparative judgments. *Psychological Review, 34*, 273–286.

Thurstone, L. L. (1945). The prediction of choice. *Psychometrika, 10*, 237–253.

Weinberg, S. L. Carroll, J. D., & Cohen, H. S. (1984). Confidence regions for INDSCAL using the jackknife and bootstrap techniques. *Psychometrika, 49*, 475–491.

Manuscript received 3/18/86
Final version received 9/16/86