

LOGLINEAR RASCH MODEL TESTS

HENDRIKUS KELDERMAN

TWENTE UNIVERSITY OF TECHNOLOGY
ENSCHEDE, THE NETHERLANDS

Existing statistical tests for the fit of the Rasch model have been criticized, because they are only sensitive to specific violations of its assumptions. Contingency table methods using loglinear models have been used to test various psychometric models. In this paper, the assumptions of the Rasch model are discussed and the Rasch model is reformulated as a quasi-independence model. The model is a quasi-loglinear model for the incomplete subgroup \times score \times item $1 \times$ item $2 \times \dots \times$ item k contingency table. Using ordinary contingency table methods the Rasch model can be tested generally or against less restrictive quasi-loglinear models to investigate specific violations of its assumptions.

Key words: latent-trait theory, Rasch model, quasi-loglinear model, quasi independence, incomplete contingency table, chi-square tests.

Introduction

Over the past decade, the Rasch (1960) model has become increasingly popular in constructing and scoring psychological tests (Fischer, 1974, 1978; Hambleton, Swaminatan, Cook, Eignor & Gifford, 1978; Mellenbergh, 1972; Renz & Bashaw, 1977). Its statistical properties have been studied extensively (Rasch, 1960, 1966a,b; Andersen, 1971, 1972, 1973a; Fischer, 1974, 1981), and efficient algorithms to estimate its parameters are developed (Gustafsson, 1970; Fischer & Scheiblechner, 1970; Wright & Mead, 1977). A number of statistical tests for the fit of a set of data to the Rasch model have been described (e.g., Andersen, 1973b; Fischer & Scheiblechner, 1970; Gustafsson, 1980; van den Wollenberg, 1979, 1982; Wright & Panchapakesan, 1969). Unfortunately, none of these tests is completely satisfactory, since they are only sensitive to specific violations of the assumptions of the Rasch model.

Gustafsson (1980) has recommended that the fit of a set of data to the Rasch model be investigated with respect to its different assumptions and emphasizes the need for a more exact definition of the assumption of unidimensionality. Moreover, Lumsden (1978) has pointed out that deviations from unidimensionality have been seriously neglected by both test constructors and test theorists.

In this paper the assumptions of the Rasch model are discussed. A general definition of unidimensionality is proposed that yields the assumption of local statistical independence as a special case.

Loglinear models (Andersen, 1980a; Bishop, Fienberg, & Holland, 1975; Fienberg, 1980; Goodman, 1978; Haberman, 1978, 1979) have been used for the estimation and testing of various psychometric models. Loglinear models, or their multiplicative equivalents, have been applied to Guttman's (1950) perfect scale model (Clogg & Sawyer, 1981; Davison, 1980; Dayton & Macready, 1980; Goodman, 1959, 1975). They have also been applied to Coombs' (1964) unfolding model (Davison, 1979) and to the model of item

The author thanks Gideon J. Mellenbergh, Pieter Vijn, Wim J. van der Linden, Ivo W. Molenaar, and Siebe J. Oosterloo for comments and Carolien Schamhardt and Anita Burchartz for typing the manuscript. Requests for reprints should be sent to Hendrikus Kelderman, Toegepaste Onderwijskunde, Technische Hogeschool Twente, Postbus 217, 7500 AE Enschede, The Netherlands.

homogeneity (Lienert & Raatz, 1981). Furthermore, Mellenbergh and Vijn (1981) use a logit-linear model to estimate the parameters in the Rasch model. In this paper the Rasch model is formulated as a quasi-loglinear model which can be tested generally or against less restrictive quasi-loglinear models using contingency table methods.

The Assumptions of the Rasch Model

An individual's responses to k test items are denoted by X_j ($j = 1, \dots, k$), which can take values 0 (negative, wrong, disagree) or 1 (positive, right, agree). The dichotomous Rasch model assumes that, for a given individual, the probability of a response x_j ($x_j = 0, 1$) depends on one latent ability (attitude etc.) parameter α (Rasch, 1961):

$$P(X_j = x_j | \alpha) = \exp(x_j(\alpha - \delta_j)) / (1 + \exp(\alpha - \delta_j)), \quad (1)$$

where δ_j is the item parameter describing the difficulty of item j on the latent continuum. Equation 1 is Rasch's special logistic model. For $x_j = 1$ it describes the ICC of item j . It follows from Equation 1 that the ICC's of the items differ only in location.

The second assumption of the Rasch model is the assumption of unidimensionality. Although it is implicit in model (1), it will be given a more precise meaning by expressing it as:

$$P(X_j = x_j | \alpha; y_1, \dots, y_q) = P(X_j = x_j | \alpha), \quad (2)$$

for all possible variables y_1, \dots, y_q which are not functions of x_j and α . Equation 2 represents Lord & Novick's (1968) interpretation of unidimensionality:

An individual's performance depends on a single underlying trait if, given his value on that trait, nothing further can be learned from him that can contribute to the explanation of his performance. The proposition is that the latent trait is the only important factor and, once a person's value on the trait is determined, the behaviour is random, in the sense of statistical independence (p. 538).

Assumption (2) follows from the fact that (1) must hold on the individual level, i.e.

$$P(X_j^{(v)} = x_j^{(v)} | \alpha^{(v)}) = P(X_j = x_j^{(v)} | \alpha), \quad \text{if } \alpha^{(v)} = \alpha$$

for all individuals v in the population of interest. But if (1) is invariant over individuals with the same α , (1) must be invariant over sets of individuals with the same y_1, \dots, y_q and α .

Special definitions of unidimensionality emerge if the y_1, \dots, y_q variables in Equation 2 are restricted to be specific types of variables. Definition (2) contains two well-known assumptions of the Rasch model as a special case: local (or conditional) independence and invariance of the ICC's for any subpopulation. The assumption of local independence (or measurement independence (Lord & Novick, 1968, p. 44)), can be obtained by setting all y variables in Equation 2 equal to the responses to the other items $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k$ (Lord & Novick, 1968, pp. 361, 538). It means that the observed dependence among the item responses is wholly explained by their dependence on the latent variable α . The second assumption, invariance of the ICC's over subpopulations (Lord & Novick, 1968, p. 359), is obtained by restricting the y variables in (2) to be characteristics of individuals such as sex and age.

In a discussion of the concept of local independence in the Rasch model, Goldstein (1980) correctly remarks that both special definitions of unidimensionality have not always been distinguished properly: invariance of the ICC's over subpopulations does not necessarily imply local independence and vice versa.

To obtain a coherent formulation of the Rasch model both basic assumptions, logistic ICC's (1) and unidimensionality (2), must be combined into one single equation. First, let $(y_1, \dots, y_q) = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k, z_1, \dots, z_w)$ in Equation 2, where z_1, \dots, z_w

describe individuals' characteristics. An alternative but equivalent formulation of equation (2) is then

$$P(x_1, \dots, x_k | \alpha; z_1, \dots, z_w) = P(x_1, \dots, x_k | \alpha) = \prod_{j=1}^k P(x_j | \alpha), \tag{3}$$

where $P(x_1, \dots, x_k | \alpha; z_1, \dots, z_w)$ is the joint distribution of item responses x_1, \dots, x_k for given values of α and z_1, \dots, z_w and the random variable notation is dropped. The equivalence of Equation 2 and 3 follows from elementary probability calculus (Mood, Graybill, & Boes, 1974, sec. 3.6). Inserting Equation 1 into 3, a unitary formulation of the Rasch model follows:

$$P(\mathbf{x} | \alpha; \mathbf{z}) = \prod_{j=1}^k \frac{\exp(x_j(\alpha - \delta_j))}{(1 + \exp(\alpha - \delta_j))} = \frac{\exp\left(\alpha t - \sum_{j=1}^k x_j \delta_j\right)}{\prod_{j=1}^k (1 + \exp(\alpha - \delta_j))} \tag{4}$$

where $\mathbf{z} = (z_1, \dots, z_w)$, $\mathbf{x} = (x_1, \dots, x_k)$ and $t = \sum_j x_j$ is the number of positive item responses (test score).

Equation 4 describes the structure of a 2^k contingency table of the x_1, \dots, x_k responses of an individual with ability α (cf. Goldstein, 1980). The hypothesis is that this structure can be explained by the marginal item response distributions, which depend only on the individual's ability α through the logistic function (1) and do not depend on the value of \mathbf{z} .

This formulation of the Rasch model has several advantages over the usual formulation (Equation 1). First, it contains the unidimensionality assumption in a more explicit manner. Second, it describes all variables relevant for testing the Rasch model. Third, the results of latent structure analysis (Andersen, 1980; Goodman, 1978; Lazarsfeld & Henry, 1968) can be applied to Formula 4, since it describes the probabilities of the manifest item response vector x_1, \dots, x_k in terms of the latent ability variable α . Furthermore, an important feature of model Formula 4 is that it is an exponential family distribution wherein the score t is a minimal sufficient statistic for the individual's ability α (Andersen, 1980a, p. 38). Consequently, the score t contains all information about ability α available in the data.

A Loglinear Rasch Model

Model Formula 4 defines a latent structure model. A basic concept in latent structure analysis is the so-called accounting equation. An accounting equation describes the probability $P(\mathbf{x})$ of an observed response \mathbf{x} in terms of conditional probabilities $P(\mathbf{x} | \alpha)$ and the distribution function $F(\alpha)$ of the latent variable (Lazarsfeld & Henry, 1968):

$$P(\mathbf{x}) = \int_{-\infty}^{+\infty} P(\mathbf{x} | \alpha) dF(\alpha). \tag{5}$$

To apply Equation 5 to the Rasch model (4), we must bring in the \mathbf{z} variables. Assume that z_1, \dots, z_w are categorical or categorized variables and, for simplicity, denote each distinct value of the joint variable \mathbf{z} by a single index i ($i = 1, \dots, m$), which will be referred to as the i th "subgroup". Equation 5 is now applied to each subgroup:

$$P_i(\mathbf{x}) = \int_{-\infty}^{+\infty} P_i(\mathbf{x} | \alpha) dF_i(\alpha), \tag{6}$$

where $P_i(\mathbf{x}) = P(\mathbf{x} | \mathbf{z})$ is the conditional distribution of response \mathbf{x} given \mathbf{z} (or i), $P_i(\mathbf{x} |$

$\alpha) = P(\mathbf{x} | \alpha; \mathbf{z})$ is the conditional distribution of \mathbf{x} given α and \mathbf{z} (or i) and $F_i(\alpha) = F(\alpha | \mathbf{z})$ is the conditional distribution function of α given \mathbf{z} (or i).

Inserting the Rasch model (4) into (6) yields:

$$\begin{aligned} P_i(\mathbf{x}) &= \int_{-\infty}^{+\infty} \frac{\exp\left(\alpha t - \sum_{j=1}^k x_j \delta_j\right)}{\prod_{j=1}^k (1 + \exp(\alpha - \delta_j))} dF_i(\alpha) \\ &= \exp\left(-\sum_{j=1}^k x_j \delta_j\right) \int_{-\infty}^{+\infty} \frac{\exp(\alpha t)}{\prod_{j=1}^k (1 + \exp(\alpha - \delta_j))} dF_i(\alpha). \end{aligned} \quad (7)$$

From (4) it also follows that

$$\begin{aligned} P_i(t | \alpha) &= \sum_{x_1} \cdots \sum_{x_k} P_i(x_1, \dots, x_k | \alpha); \quad x_1 + \cdots + x_k = t \\ &= \gamma_t(\delta_1, \dots, \delta_k) \frac{\exp(\alpha t)}{\prod_{j=1}^k (1 + \exp(\alpha - \delta_j))}, \end{aligned}$$

where

$$\gamma_t(\delta_1, \dots, \delta_k) = \sum_{x_1} \cdots \sum_{x_k} \prod_{j=1}^k \exp(-x_j \delta_j); \quad x_1 + \cdots + x_k = t$$

are the well-known, elementary symmetric functions and $P_i(t | \alpha)$ is the conditional probability of score t given ability α and subgroup i . The marginal score distribution $P_i(t)$ in each subgroup i is then:

$$P_i(t) = \gamma_t(\delta_1, \dots, \delta_k) \int_{-\infty}^{+\infty} \frac{\exp(\alpha t)}{\prod_{j=1}^k (1 + \exp(\alpha - \delta_j))} dF_i(\alpha) \quad (8)$$

Using (8), (7) can be written as

$$P_i(\mathbf{x}) = \frac{\exp\left(-\sum_{j=1}^k x_j \delta_j\right)}{\gamma_t(\delta_1, \dots, \delta_k)} P_i(t) \quad (9)$$

so that

$$P_i(\mathbf{x} | t) = P_i(\mathbf{x})/P_i(t) = \frac{\exp\left(-\sum_{j=1}^k x_j \delta_j\right)}{\gamma_t(\delta_1, \dots, \delta_k)} \quad (10)$$

Model (10) is the conditional Rasch (1960) model whereas in model (9) there is no conditioning on the score t . Both model (9) and model (10) can be written as a loglinear model. Model (9) corresponds to

$$\ln m_{itx_1 \dots x_k}^* = \sigma_{it}^* - \sum_{j=1}^k x_j \delta_j, \quad (11)$$

where

$$m_{itx_1 \dots x_k}^* = N_i P_i(\mathbf{x}) \quad \text{and} \quad \sigma_{it}^* = \ln(N_i P_i(t)/\gamma_t(\delta_1, \dots, \delta_k)), \quad (12)$$

with N_i the number of subjects in subgroup i and where m_{itx_1, \dots, x_k}^* is the expected number of responses x_1, \dots, x_k with score t in group i under model (9). Model (10) corresponds to

$$\ln m_{itx_1 \dots x_k} = \sigma_{it} - \sum_{j=1}^k x_j \delta_j \tag{13}$$

where

$$m_{itx_1 \dots x_k} = N_{it} P_i(\mathbf{x} | t) \quad \text{and} \quad \sigma_{it} = \ln (N_{it}/\gamma_i(\delta_1, \dots, \delta_k)),$$

with N_{it} the number of subjects in subgroup i having score t and where $m_{itx_1 \dots x_k}$ is the expected number of responses x_1, \dots, x_k with score t from group i .

In the conditional Rasch model (13) the observed subgroup \times score marginals N_{it} and the corresponding model parameters σ_{it} are considered as fixed by the sampling design. Consequently, there are no restrictions on σ_{it} . In the Rasch model (11) only the subgroup marginals N_i are considered as fixed and the parameter σ_{it}^* must satisfy equation (12). σ_{it}^* depends on the distribution $P_i(t)$ of the score in each of the subgroups. Because in (8) the distribution $P_i(t)$ depends on an underlying latent trait distribution, it can be shown that $P_i(t)$, and therefore σ_{it}^* , must satisfy certain constraints. Cressie and Holland (1983) studied these constraints. Using their results it can be shown (see Appendix I) that the σ^* -parameters must be consistent with the inequalities.

$$\det. \left(\left\| \exp \{ \sigma_{i(r+s)}^* \} \right\|_{r, s=0}^{q_1} \right) \geq 0$$

$$\det. \left(\left\| \exp \{ \sigma_{i(r+s+1)}^* \} \right\|_{r, s=0}^{q_2} \right) \geq 0$$

$i = 1, \dots, m$, where

$$q_1 = \begin{cases} k/2 & \text{if } k \text{ is even,} \\ (k-1)/2 & \text{if } k \text{ is odd,} \end{cases} \quad q_2 = \begin{cases} (k-2)/2 & \text{if } k \text{ is even,} \\ (k-1)/2 & \text{if } k \text{ is odd,} \end{cases}$$

and $\| \cdot \|_{r, s=0}^q$ denotes a matrix with row index r and column index s both running from zero to q . Throughout this paper we either assume that these constraints hold or we work with the conditional model (13).

In model Formula 13 (and 11) there is an obvious overparametrisation; adding a constant c to each item parameter δ_j and adding $c \cdot t$ to each subgroup \times score parameter σ_{it} , does not change the model. This indeterminacy can be removed by setting one item parameter equal to zero.

Model Formula 13 is a quasi-loglinear model for the incomplete subgroup \times score \times item $1 \times \dots \times$ item k contingency table with expected counts $m_{itx_1 \dots x_k}$ if $t = \sum_{j=1}^k x_j$ and structural (or a priori) zero cells otherwise (Bishop et al., 1975, sec. 5.4; Haberman, 1979, sec. 7.3). Table 1 shows the a priori pattern of expected counts and zero cells for $k = 3$ items in subgroup i .

Quasi-loglinear models describe quasi-independence structures in incomplete contingency tables such as Table 1. The quasi-independence concept was first introduced by Goodman (1968). The concept is also fundamental to other scaling models (Davison, 1979, 1980; Goodman, 1975). It means that there are no interactions between certain variables beyond those already imposed by the a priori incompleteness structure of the table, i.e. the pattern of structurally zero and structurally nonzero cells.

The quasi-loglinear Rasch model (13) then states that the item responses are quasi-independent of the subgroup, that the item responses are quasi-independent of the score, and that all item responses are quasi-independent of each other. The a priori incomplete-

TABLE 1

Expected Counts and Structural Zero's in Subgroup

$\underline{i} \times \text{Score} \times \text{Item 1} \times \text{Item 2} \times \text{Item 3 Table.}$

Item Response			Score \underline{t}			
			0	1	2	3
\underline{x}_1	\underline{x}_2	\underline{x}_3				
0	0	0	\underline{m}_{i0000}	-	-	-
1	0	0	-	\underline{m}_{i1100}	-	-
0	1	0	-	\underline{m}_{i1010}	-	-
0	0	1	-	\underline{m}_{i1001}	-	-
1	1	0	-	-	\underline{m}_{i2110}	-
1	0	1	-	-	\underline{m}_{i2101}	-
0	1	1	-	-	\underline{m}_{i2011}	-
1	1	1	-	-	-	\underline{m}_{i3111}

Note. Dashes denote structurally zero cells.

ness structure of the subgroup \times score \times item $1 \times \dots \times$ item k table is defined by the dependency $t = \sum_{j=1}^k x_j$ of the score variable on the item responses. Consequently, the quasi-loglinear Rasch model (13) states that the interactions of an item response with e.g. the subgroup, the score, or another item response are *only* explained by their contribution to the score.

By Formula 3 through 13 it is proved that this quasi-independence hypothesis follows from the hypothesis that the data satisfy the assumptions of the Rasch model. This becomes more obvious if we realize that the score is a minimal sufficient statistic for the latent ability parameter α . Consequently, the score t contains all information about α available in a given set of data. Since the Rasch model states that the item response distribution varies only with the latent ability parameter α , the sufficiency property of the score t implies that for a given set of items the item response distribution varies only with the score t . The latter proposition, however, describes precisely what is meant by the quasi-independence hypothesis (Formula 13) for the incomplete subgroup \times score \times item $1 \times \dots \times$ item k contingency table. Consequently, this table contains all information that is available in a set of data to test the Rasch model. It will, therefore, be called: "The Rasch table".

If it is assumed that the individuals respond independently of one another, the quasi-loglinear Rasch model (11) describes a *parametric multinomial distribution* of the structurally nonzero observed counts $f_{itx_1 \dots x_k}$ in each subgroup i and the quasi-loglinear Rasch model (13) describes a *multinomial distribution* of the structurally nonzero observed counts in each subgroup i and scoregroup t (Andersen, 1980a; Bishop et al., 1975). Maximum likelihood estimates of the parameters in model (13) and likelihood ratio tests can be obtained by fitting the structurally nonzero expected cell counts $m_{itx_1 \dots x_k}$ to the observed cell counts $f_{itx_1 \dots x_k}$ (Bishop et al., 1975; Goodman, 1978; Haberman, 1978, 1979). It can be shown (see Appendix II) that the estimates of the item parameters in model (13) are identical to the well-known conditional maximum likelihood (CML) estimates (Andersen, 1973a). Algorithms to estimate the parameters in (quasi-) loglinear models have been described by Baker and Nelder (1978), Goodman and Fay (1974) and Haberman (1979).

In model (13), the effect of a negative item response is set to zero and the joint subgroup \times score effect is parametrized as a single parameter σ_{it} . Model (13) can be rewritten by using the ANOVA parametrisation or, equivalently, the “ u -terms” parametrisation often used in loglinear contingency table models (Bishop et al., 1975), that is,

$$\ln m_{itx_1 \dots x_k} = u + u_1(i) + u_2(t) + u_{12}(it) + u_3(x_1) + \dots + u_{k+2}(x_k) \tag{14}$$

with the ANOVA constraints

$$u_1(+) = u_2(+) = u_{12}(+t) = u_{12}(i+) = u_3(+) = \dots = u_{k+2}(+) = 0$$

where u is a constant term, $u_1(i)$ is the main effect on subgroup i , $u_2(t)$ is the main effect of score t , $u_{12}(it)$ is the effect of the combination of subgroup i and score t , and $u_3(x_1), \dots, u_{k+2}(x_k)$ are the effects of response x_j of item j ($j = 1, \dots, k$). A plus sign replacing an index means that the model parameters are summed over that index. To remove the indeterminacy between item parameters and score parameters one additional linear constraint must be imposed on the item parameters, e.g. $u_3(x_1) = 0$. A proof of the equivalence of both parametrisations (13) and (14) can be found in Bock (1975, pp. 50, 239).

The advantage of the u -term notation is that it allows us to specify alternative quasi-loglinear models for the Rasch table. The u -term notation can be used to specify alternative models with a broad range of complexity for any number of variables, from the fully specified or “saturated” model, to a more parsimonious or “restrictive” model. For example, a restrictive model may be specified by adding interaction terms $u_{13}(ix_1), \dots, u_{1(k+2)}(ix_k)$ to the Rasch model (14), to relax the quasi-independence hypothesis for subgroup \times item interactions.

By comparing the fit of the Rasch model with the fit of more complex models, by the methods described in what follows, various hypotheses on violations of the assumptions of the Rasch model can be tested.

Quasi-loglinear Models for the Rasch Table

Fienberg (1972; see also Bishop et al., 1975, sec. 5.4) presents a general theory for the analysis of incomplete contingency tables by quasi-loglinear models. In this section we apply Fienberg’s theory to the Rasch table.

The saturated or fully specified model for the Rasch table is:

$$\begin{aligned} \ln m_{itx_1 \dots x_k} = & u + u_1(i) + u_2(t) + u_3(x_1) + \dots + u_{(k+2)}(x_k) \\ & + u_{12}(it) + u_{13}(ix_1) + \dots + u_{(k+1)(k+2)}(x_{k-1}x_k) \\ & + u_{123}(itx_1) + \dots + u_{123 \dots (k+2)}(itx_1 \dots x_k) \end{aligned} \tag{15}$$

for $i = 1, \dots, m$; $x_1 = 0, 1; \dots; x_k = 0, 1$; $t = \sum_{j=1}^k x_j$. Note that the last implies that not all combinations of item and score indices can occur in the parameters of model (15). Model (15) has constraints:

$$\begin{aligned} u_1(+) &= u_2(+) = \dots = u_{(k+2)}(+) = u_{12}(+t) = u_{12}(i+) = u_{13}(+x_1) = u_{13}(i+) = \dots \\ &= u_{(k+1)(k+2)}(+x_k) = u_{(k+1)(k+2)}(x_{k-1}+) = \dots = u_{123}(+tx_1) = u_{123}(i+x_1) \\ &= u_{123}(it+) = \dots = u_{123 \dots (k+2)}(+tx_1 \dots x_k) = u_{123 \dots (k+2)}(i+x_1 \dots x_k) = \dots \\ &= u_{123 \dots (k+2)}(itx_1 \dots x_{k-1}+) = 0 \end{aligned} \quad (16)$$

The u -terms in model (15) describe main effects and interaction effects of subgroup i , score t and item responses x_1, \dots, x_k . The u -terms in expression (16) denote sums of parameters that occur in model (15), where a plus sign replacing an index indicates that the summation is over the replaced index. Fienberg (1972) gives a more precise but more elaborate expression for the constraints (16), using indicator variables to identify parameters that occur in model (15), but expression (16) and Fienberg's expression are equivalent. The constraints, however, are not sufficient to ensure that all parameters in model (15) are estimable. Like the Rasch model, additional constraints must be imposed to obtain a unique solution of the model parameters. This problem will be discussed later in this section.

Restrictive quasi-loglinear models for the Rasch table are defined by setting u -terms in (15) equal to zero. The only models considered here will be hierarchical, i.e. whenever a particular u -term is set to zero, all its higher order relatives must also be set to zero. For example, the loglinear Rasch model (14) is a restrictive model. It is obtained from (15) by setting all u -terms describing interactions with item responses equal to zero. As another example, consider the model

$$\begin{aligned} \ln m_{itx_1 \dots x_k} &= u + u_1(i) + u_2(t) + u_{12}(it) + u_3(x_1) + u_4(x_2) + \dots + u_{k+2}(x_k) + u_{23}(tx_1) \\ &\quad + u_{24}(tx_2) + \dots + u_{2(k+2)}(tx_k), \end{aligned} \quad (17)$$

where, as before, the model parameters are constrained to sum to zero over each index. Model (17) is obtained from model (15) by setting all u -terms describing interactions between item responses, interactions between item responses and the subgroup and interactions between two or more item responses and the score equal to zero. The Rasch model (14) is a special case of model (17); it can be obtained by setting the parameters $u_{23}(tx_1), \dots, u_{2(k+2)}(tx_k)$ in model (17) equal to zero.

If one model, say M , is a special case of another model, say M^* , model M can be tested against model M^* by the log-likelihood-ratio statistic:

$$G^2(m; m^*) = 2 \sum m^* \ln (m^*/m), \quad (18)$$

where m and m^* denote the expected counts of model M and M^* respectively. In Equation 18 indices are omitted for brevity, and the summation is over all structurally nonzero cells. Under the assumption of model M , $G^2(m; m^*)$ is asymptotically distributed as *chi square* with degrees of freedom equal to the difference between the numbers of estimable parameters of both models.

For example, the Rasch model (14) is a special case of model (17). Testing model (14) against model (17) is a test for the invariance of item difficulties over scoregroups. It can be shown (see Appendix III) that this test is equivalent to a well-known likelihood-ratio test by Andersen (1973b).

If M^* is the saturated model, the expected counts m^* are equal to the observed counts f , and G^2 becomes an overall goodness of fit statistic for model M :

$$G^2(m; f) = 2 \sum f \ln (f/m). \quad (19)$$

First, there are the usual indeterminacies in \mathbf{u} caused by the linear dependence of the categories of the observed variables, i.e.:

$$\begin{aligned} (S_0 + S_1 + S_2 + S_3) - G = 0, & \quad (R_0^{(1)} + R_1^{(1)}) - G = 0, \\ (R_0^{(2)} + R_1^{(2)}) - G = 0, & \quad \text{and} \quad (R_0^{(3)} + R_1^{(3)}) - G = 0, \end{aligned}$$

for the main effects, and

$$(S_0 R_{x'_1}^{(1)} + S_1 R_{x'_1}^{(1)} + S_2 R_{x'_1}^{(1)} + S_3 R_{x'_1}^{(1)}) - R_{x'_1}^{(1)} = 0, \quad x'_1 = 0, 1,$$

and

$$(S_{t'} R_0^{(1)} + S_{t'} R_1^{(1)}) - S_{t'} = 0, \quad t' = 0, \dots, 3,$$

for the interaction effects. As before, these indeterminacies are removed by constraining the parameters to sum to zero over each index, which leaves us with $1 + (4 - 1) + (2 - 1) + (2 - 1) + (2 - 1) + (4 - 1)(2 - 1) = 10$ linearly unrestricted parameters $u, u_1(t'), u_2(x'_1), u_3(x'_2), u_4(x'_3)$, and $u_{12}(t'x'_1)$ respectively.

Second, some parameters are not estimable since they cannot occur in the model equations, that is, they correspond to structurally zero counts only. From model (22) it is seen that

$$S_0 R_1^{(1)} = S_3 R_0^{(1)} = 0,$$

since there is no zero or perfect score together with a positive or negative response respectively. The corresponding parameters $u_{12}(01)$ and $u_{12}(30)$ are not estimable, since there are no data to estimate them from. This indeterminacy can be removed by setting both parameters equal to zero which leaves $1 + (4 - 1) + (2 - 1) + (2 - 1) + (2 - 1) + (2 - 1)(2 - 1) = 8$ linearly unrestricted parameters.

Third, as in the Rasch model, there is an indeterminacy between the score and item parameters due to the linear dependence of the number right score on the item responses:

$$(0S_0 + 1S_1 + 2S_2 + 3S_3) - (R_1^{(1)} + R_1^{(2)} + R_1^{(3)}) = 0.$$

As before, this indeterminacy is removed by setting one item parameter, say $u_4(x_3)$, equal to zero. The total number of estimable parameters then becomes $1 + (4 - 1) + (2 - 1) + (2 - 1) + (2 - 1)(2 - 1) = 7$. Since there are $2^3 = 8$ structurally nonzero cells in this table, the number of degrees of freedom for the chi-square statistics Formula 19 and 20 is $8 - 7 = 1$. Obviously if the model becomes more complex it is better to determine the degrees of freedom by evaluating the rank of the design matrix numerically.

Finally, even if the model parameters are estimable by the structure of the model, the parameter estimates actually obtained from a particular set of data may not be unique. This phenomenon is well-known in the analysis of covariance structures (Kelderman, Mellenbergh, & Elshout, 1981; McDonald & Krane, 1979). Fischer (1981) has given necessary and sufficient conditions which the data must satisfy to obtain unique maximum-likelihood estimates of the parameters in the Rasch model. Fienberg (1972), (Bishop et al., 1975) has given necessary and sufficient conditions which the data must satisfy to obtain unique maximum-likelihood estimates of the expected cell counts in incomplete tables. Both Fischer and Fienberg employ graph-theoretic concepts (e.g. connectedness), whose treatment is beyond the scope of this paper.

A practical method is to determine the rank of the information matrix, which should be equal to the number of estimable parameters for a given set of data (cf. McHugh, 1956; Goodman, 1974). Baker and Nelder (1978, sec. 4.3) describes a weighted least-squares algorithm for the analysis of contingency tables, which estimates the parameters in a

sequential fashion. If a parameter is linearly dependent on the preceding parameters, or if there are no observations to estimate it from, the parameter is removed from the model, thus the information matrix is of full rank. This enables us to obtain the number of estimable parameters in the model. The reliability of this procedure, however, depends on the appropriateness of the numerical methods used to estimate the parameters from a particular set of data (McDonald & Krane, 1979). The rank of the information matrix, therefore, is not completely convincing evidence on the number of estimable parameters. In the following, we assume that there are no indeterminacies in the model parameters caused by the structure of the data.

Loglinear Rasch Model Tests

The set of u -terms that appear in the saturated model but not in the Rasch model, describes all possible deviations from the assumptions of the Rasch model. Comparing the Rasch model with the saturated model, or equivalently with the data itself, by the $G^2(m; f)$ or $Q(m)$ statistic therefore gives an overall test for the fit of a set of data to the Rasch model.

This test, however, is not very powerful against specific violations of the Rasch model. If one has some idea about the types of departures that are most likely to occur in a particular set of data, it is better to carry out a test that concentrates on these departures (Molenaar, 1983). If the hypothesized departures can be described by a (hierarchical) set of u -terms, they can be tested by the likelihood ratio statistic $G^2(m; m^*)$. The Rasch model M must then be tested against a less restrictive model M^* that contains both the hypothesized u -terms and the parameters of the Rasch model. If the value of $G^2(m; m^*)$ is large relative to its degrees of freedom the hypothesis that the hypothesized u -terms are zero must be rejected, indicating that the hypothesized departures of the Rasch model are present indeed. Table 3 (b through i), displays some of these hypotheses and the degrees of freedom of the corresponding $G^2(m; m^*)$ test.

The assumption on the special logistic form of the item characteristic curves (Formula 1) implies a single item difficulty parameter for each item. These parameters are independent of the individuals' scores on the latent trait. Therefore, if the item response parameters vary with the score, which is a sufficient statistic for the latent trait, the ICC assumption is violated. In test b, the Rasch model is compared with Model (17) containing parameters describing an interaction between item responses and the score, relaxing the quasi-independence assumption of item responses with the score. If test b is significant, these parameters are not zero, and the ICC assumption is clearly violated. It can be shown (see Appendix III) that test b is identical to the well-known conditional likelihood-ratio test by Andersen (1973b; 1980a, p. 253). In test i, the ICC hypothesis is tested with respect to only one item. Obviously, by specifying an appropriate comparison model, the ICC hypothesis can be tested with respect to any subset of items.

Tests c through h in Table 3 are each sensitive to a specific violation of the unidimensionality assumption (Formula 2). In tests c and h the hypothesis of unidimensionality is restricted to the invariance of the ICC's over subpopulations (Lord & Novick, 1968, p. 359). Goldstein (1980) has called this type of unidimensionality: "The unidimensionality of the between individuals space". If test c is significant, the items have different item difficulties in different subgroups, where the terms $u_{13}(ix_1), \dots, u_{1(k+2)}(ix_k)$ measure the departures in each subgroup i .

In test h this hypothesis is tested with respect to one item only. If test h is significant it means that the item measures a specific latent trait that is correlated with the subgroups. Test h can be used to detect biased items, i.e. items that are more difficult for some subgroups, e.g. minority groups (cf. Rudner, Getson, & Knight, 1980).

TABLE 3

Some Loglinear Rasch Model Tests.

Test	Null Hypothesis	Degrees of freedom
Comparison with less Restrictive Models		
a.	difference with saturated model	$\underline{m} \cdot 2^{\underline{k}} - (\underline{k} \cdot \underline{m} + \underline{k} + \underline{m} - 1)$
b.	$\underline{u}_{23}(\underline{t} \underline{x}_1) = \underline{u}_{24}(\underline{t} \underline{x}_2) = \dots = \underline{u}_{2(\underline{k}+2)}(\underline{t} \underline{x}_{\underline{k}}) = 0$	$(\underline{k} - 2)(\underline{k} - 1)$
c.	$\underline{u}_{13}(\underline{i} \underline{x}_1) = \underline{u}_{14}(\underline{i} \underline{x}_2) = \dots = \underline{u}_{1(\underline{k}+2)}(\underline{i} \underline{x}_{\underline{k}}) = 0$	$(\underline{m} - 1)(\underline{k} - 1)$
d.	$\underline{u}_{34}(\underline{x}_1 \underline{x}_2) = \underline{u}_{35}(\underline{x}_1 \underline{x}_3) = \dots = \underline{u}_{(\underline{k}+1)(\underline{k}+2)}(\underline{x}_{\underline{k}-1} \underline{x}_{\underline{k}}) =$ $\underline{u}_{345}(\underline{x}_1 \underline{x}_2 \underline{x}_3) = \dots = \underline{u}_{34 \dots \underline{k}+2}(\underline{x}_1 \underline{x}_2 \dots \underline{x}_{\underline{k}}) = 0$	$\frac{\underline{k}}{\underline{r}=2} \left(\frac{\underline{k}}{\underline{r}} - 1 \right)$
e.	$\underline{u}_{34}(\underline{x}_1 \underline{x}_2) = \underline{u}_{35}(\underline{x}_1 \underline{x}_3) = \dots = \underline{u}_{(\underline{k}+1)(\underline{k}+2)}(\underline{x}_{\underline{k}-1} \underline{x}_{\underline{k}}) = 0$	$\left(\frac{\underline{k}}{2} \right) - 1$
f.	$\underline{u}_{34}(\underline{x}_1 \underline{x}_2) = \underline{u}_{45}(\underline{x}_2 \underline{x}_3) = \underline{u}_{56}(\underline{x}_3 \underline{x}_4) = \dots =$ $\underline{u}_{(\underline{k}+1)(\underline{k}+2)}(\underline{x}_{\underline{k}-1} \underline{x}_{\underline{k}}) = 0$	$(\underline{k} - 1)$
g.	$\underline{u}_{(\underline{j}+2)(\underline{\ell}+2)}(\underline{x}_{\underline{j}} \underline{x}_{\underline{\ell}}) = 0$	1
h.	$\underline{u}_{1(\underline{j}+2)}(\underline{i} \underline{x}_{\underline{j}}) = 0$	$\underline{m} - 1$
i.	$\underline{u}_{2(\underline{j}+2)}(\underline{t} \underline{x}_{\underline{j}}) = 0$	$(\underline{k} - 2)$
Comparison with more Restrictive Models		
j.	$\underline{u}_{2(\underline{t})} = \underline{u}_{12}(\underline{i} \underline{t}) = 0$	$\underline{k} \cdot \underline{m} - 1$
k.	$\underline{u}_{12}(\underline{i} \underline{t}) = 0$	$\underline{k}(\underline{m} - 1)$
l.	$\underline{u}_{3(\underline{x}_1)} = \underline{u}_{4(\underline{x}_2)} = \dots = \underline{u}_{\underline{k}+2}(\underline{x}_{\underline{k}}) = 0$	$(\underline{k} - 1)$

Note. Degrees of freedom are the differences in rank of the design matrices of both models.

In test d through g the hypothesis of unidimensionality is restricted to the hypothesis of local independence (Lord & Novick, 1968, pp. 361, 538). Goldstein (1980) has called this: "The unidimensionality of the within individuals space". In test d the Rasch model is

compared with a model that contains interactions between item responses of all orders. If test d is significant one latent ability trait is not sufficient to explain the associations between the item responses. In that case, it is necessary to postulate additional latent traits to account for these item response interactions, and the unidimensionality assumption is violated. Note that these traits need not be correlated with the subgroups and may be entirely caused by the test items or conditions of test administration. For example, if, by some unlucky accident, an item is administered twice in the same test, we expect their association to be almost perfect, violating the local independence assumption. Since all examinees may recognize the identity of the items the "recognition trait" need not be associated with the subgroups. Therefore, unidimensionality of the between individuals latent space need not imply unidimensionality of the within individuals latent space (Goldstein, 1980; see also van den Wollenberg, 1979). More realistic examples of additional traits are fatigue or simply additional latent abilities that are needed to solve some items.

In tests e and g the Rasch model is compared with models describing more specific departures from the local independence assumption. In test e the violation is restricted to pairwise dependence between all item responses and in test g the violation is restricted to an interaction between one pair of items. Because of its greater statistical power, test e may be preferred to test d, especially if higher order interactions between item responses are unlikely. In test f the violation of local independence is restricted to pairwise dependence of consecutive item responses (cf. Kempf, 1974). Obviously, other specific violations of the local independence assumption may be tested by specifying an appropriate comparison model. Furthermore, several specific violations of the ICC assumption and the unidimensionality assumption may be combined into a single test.

If the Rasch model fits the data, we may go on further and compare it with more restrictive models, to test the hypothesis that one or more of its u -terms are equal to zero. Table 3 gives some of these tests (j, k and l).

In j and k the hypothesis tested concerns the distribution of the score. Therefore we may not use the conditional Rasch model (13), but we must use the Rasch model (11), where the score is considered as a random variable depending on a random latent ability variable α through Equation (8).

Test j is the most important one. In test j the Rasch model is compared with the model

$$\ln m_{itx_1 \dots x_k} = u + u_1(i) + u_2(x_1) + \dots + u_{k+2}(x_k) \quad (24)$$

which can be obtained from the Rasch model (14) by setting its score parameters $u_2(t)$ and $u_{1,2}(it)$ equal to zero. Model (24) is the complete independence model, where the item responses are independent of the subgroups and independent of each other. It represents the situation described by Wood (1978) where the item responses are generated by the tossing of k (biased) coins. Wood showed that the Rasch model fits simulated coin-tossing data very well. This is not very surprising since the complete independence model (24) is a special case of the Rasch model (14).

To test whether Rasch homogeneity can be distinguished from purely random responses we must, therefore, use test j. If test j fails to reach significance we must reject the hypothesis that the items measure a common latent trait. Consequently, it is not enough to test whether any of the assumptions of the Rasch model are violated, but it must also be tested against the null model (24).

In tests k and l, the Rasch model (14) is compared with more restrictive Rasch models. In test k the Rasch model is compared with a restricted Rasch model where the subgroup \times score interaction $u_{1,2}(it)$ is assumed to be zero.

TABLE 4

Observed and Expected Frequencies for the Lazarsfeld- Stouffer data.

Score	Item response				Observed frequency	Expected frequencies		
	X_1	X_2	X_3	X_4		Rasch	Equal item difficulty	complete independence
0	0	0	0	0	299	299.0(0.0)	299.0(0.0)	153.1(11.8)
1	1	0	0	0	199	199.6(0.0)	73.0(14.8)	228.8(-2.0)
1	0	1	0	0	52	43.8(1.2)	73.0(-2.5)	82.3(-3.3)
1	0	0	1	0	25	28.0(-0.6)	73.0(-5.7)	59.7(-4.5)
1	0	0	0	1	16	13.6(0.7)	73.0(-3.7)	32.0(-2.8)
2	1	1	0	0	96	97.1(-0.1)	39.2(9.1)	123.0(-2.4)
2	1	0	1	0	60	62.0(-0.3)	39.2(3.3)	89.2(-3.1)
2	0	1	1	0	45	45.9(-0.1)	39.2(0.9)	71.2(-3.1)
2	0	1	1	0	16	13.6(0.7)	39.2(-3.7)	32.1(-2.8)
2	0	1	0	1	8	10.1(-0.6)	39.2(-5.0)	25.6(-3.5)
2	0	0	1	1	10	6.4(1.4)	39.2(-4.7)	18.6(-2.0)
3	1	1	1	0	69	73.0(-0.5)	42.3(9.1)	47.1(3.0)
3	1	1	0	1	55	54.0(0.1)	42.3(2.0)	38.3(2.7)
3	1	0	1	1	42	34.5(1.3)	42.3(-0.0)	27.8(2.7)
3	0	1	1	1	3	7.6(-1.7)	42.3(-6.0)	10.0(-2.2)
4	1	1	1	1	75	75.0(0.0)	75.0(0.0)	14.9(15.6)

Note. Standardized residuals are between brackets.

If test k yields a significant outcome, it may be concluded that the score distributions are not the same in each subgroup. Using Equation (8), this implies that the distributions of the latent trait are not the same in each subgroup. Note that test k does not require

any assumptions on the distributional form of the latent trait. A more powerful test, assuming a normal distribution in each subgroup, is described by Andersen (1980b).

Finally in test l, the Rasch model is compared with a restricted Rasch model where item parameters $u_3(x_1), \dots, u_{(k+2)}(x_k)$ are assumed to be equal. Since there is one indeterminacy in the item parameters we may specify this model by setting all item parameters equal to zero. This restricted Rasch model is identical to Lienert and Raatz' (1981) model of item homogeneity. If test l is significant it is concluded that the item difficulties are not equal. In what follows some of the tests in Table 3 are applied to real data.

An Application

Table 4 shows the observed frequencies of response patterns by noncommissioned officers responding to four dichotomous items on attitudes towards the army (Lazersfeld, 1950; Stouffer, 1950).

Davison (1980) and Goodman (1975) reanalyzed this data using quasi-independence models to test Guttman's (1950) perfect scale model. Davison fitted a model describing two groups of subjects: one group whose responses are consistent with a perfect scale and one group responding randomly to the items. The model did not fit the data ($Q(m) = 26.09$, $df = 7$, $G^2(m; f)$ is not reported).

Goodman fitted a quasi-independence model (his H4) describing several groups of subjects responding on different Guttman scales and one subgroup responding randomly to the items. The model fits the data ($G^2(m; f) = 5.86$, $Q(m) = 5.47$, $df = 3$).

The Rasch model also fits the data very well ($G^2(m; f) = 10.93$, $Q(m) = 9.91$, $df = 8$). Moreover, it is more parsimonious and much easier to interpret. It may therefore be concluded that the item responses have a probabilistic rather than a deterministic relation with the latent trait.

Table 4 shows that the expected frequencies of the Rasch model are close to the observed frequencies. Table 5 gives the $G^2(m; m^*)$ values and degrees of freedom of some tests on specific violations of the Rasch model (test a through m). None of these tests show deviations from the Rasch model. Note that there are no data on subgroups, so that subgroup differences cannot be tested.

In tests n and o, the Rasch model is compared with more restrictive models. In test n the Rasch model is compared with the complete independence model. The corresponding $G^2(m; m^*)$ statistic is very large relative to its degrees of freedom, indicating a strong relation between the item response and the latent attitude trait, or equivalently, indicating steep ICC's. Moreover, Table 4 shows that the expected frequencies of the complete independence model differ substantially from the observed frequencies.

Finally the fit of the Rasch model is compared with the restricted Rasch model with equal item difficulties (test o). As can be seen from Table 5, the $G^2(m; m^*)$ value of this test is also very large relative to its degrees of freedom, indicating substantial differences in item difficulty. Moreover, Table 4 shows that the expected frequencies of this restricted Rasch model are very different from the observed frequencies. The estimates for the item parameters (and standard errors) in the Rasch model are 0.76(0.06), 0.98(0.06) and 1.13(0.06), for a negative response on item 2, 3 and 4 respectively, where the parameter of the first item is set to zero to fix the origin of the latent scale (the usual item difficulty parameters δ_j are twice $u_{(j+1)}(0)$). We may conclude that the item difficulty parameters are considerably different and can be measured with small standard errors. In sum, the Lazersfeld and Stouffer data fit a perfect Rasch scale rather than Guttman's perfect scale.

TABLE 5

Loglinear Rasch Model Tests for Lazarsfeld-Stouffer data.

Tests	Null Hypothesis	$G^2(m;m^*)$	Degrees of freedom
Comparison with less Restricted Models			
a.	$\underline{u}_{12}(t x_1) = \dots = \underline{u}_{15}(t x_4) = 0$	8.98	6
b.	$\underline{u}_{12}(t x_1) = 0$	4.28	2
c.	$\underline{u}_{13}(t x_2) = 0$	4.27	2
d.	$\underline{u}_{14}(t x_3) = 0$.71	2
e.	$\underline{u}_{15}(t x_4) = 0$	1.73	2
f.	$\underline{u}_{23}(x_1 x_2) = \underline{u}_{24}(x_1 x_3) = \dots =$ $\underline{u}_{45}(x_3 x_4) = 0$	5.27	5
g.	$\underline{u}_{23}(x_1 x_2) = 0$.67	1
h.	$\underline{u}_{24}(x_1 x_3) = 0$.10	1
i.	$\underline{u}_{25}(x_1 x_4) = 0$	2.72	1
j.	$\underline{u}_{34}(x_2 x_3) = 0$	1.25	1
k.	$\underline{u}_{35}(x_2 x_4) = 0$	1.10	1
l.	$\underline{u}_{45}(x_3 x_4) = 0$	1.64	1
m.	$\underline{u}_{23}(x_1 x_2) = \underline{u}_{34}(x_2 x_3) =$ $\underline{u}_{45}(x_3 x_4) = 0$	2.7	3
Comparison with More Restricted Models			
n.	$\underline{u}_1(t) = 0$	377.97	3
o.	$\underline{u}_2(x_1) = \underline{u}_3(x_2) = \underline{u}_4(x_3) =$ $\underline{u}_5(x_4) = 0$	458.57	3

Note. The computations are made by the GLIM programme.

Discussion

This paper shows that the Rasch model can be formulated as a quasi-loglinear-incomplete contingency table model. Quasi-independence models have already been applied to other scaling models by Davison (1979, 1980) and Goodman (1975). Furthermore, quasi-loglinear models have been applied to the Bradley-Terry model for paired comparisons (Fienberg, 1980).

Some of the results of this paper, i.e. Equation 7, are similar to results derived independently by Tjur (1982). Following a different line of reasoning, Tjur shows that the Rasch model can be formulated as a multiplicative Poisson model. The model can also be formulated as a loglinear model (Andersen, 1980a, p. 147). Tjur, however, does not use Fienberg's incomplete table methodology and Goodman's (1968) concept of quasi independence to describe the Rasch model.

Cressie and Holland (1983) do not formulate the Rasch model as a quasi-loglinear model but as an ordinary loglinear model for the complete item $1 \times \dots \times$ item k table. Their model includes interaction terms for every level of interaction which are subject to complicated constraints. As noted before, Cressie and Holland work with the "unconditional" model (11) rather than the conditional Rasch model (13). It can be shown that the conditional model is equivalent to their "generalized Rasch model".

The concept of quasi independence, accounting for the linear dependence of the item responses and the score, is fundamental in the application of loglinear models to the analysis of Rasch homogeneity. Mellenbergh and Vijn (1981) and Baker and Subkoviak (1981) use loglinear models with a score \times item $\# \times$ response table to analyze the fit of a set of data to the Rasch model. Their parameter estimates are very close to the CML estimates, but the table cannot be used for testing purposes, since the cell counts are statistically dependent by the linear dependence between the item responses and the score. Consequently, the cell counts do not follow a multinomial distribution so that the usual goodness-of-fit statistics may not be used with the score \times item $\# \times$ response table (Vijn & Mellenbergh, 1982).

In this paper we work with an incomplete table having 2^k structurally nonzero cells. If the number of items is large the table becomes very large and many cells will be empty. Most present day computer programs for the analysis of contingency tables by loglinear models (Baker & Nelder, 1978; Goodman & Fay, 1974) require the internal storage of the observed and expected table of counts, which is virtually impossible if the number of items is large. In principle, however, this is unnecessary. The table need only exist in theory. First, the data can be stored in an ordinary $N \times k$ data matrix, which avoids storage of structural and nonstructural zero's. Second, the parameter estimates may be calculated by solving the likelihood equations in terms of the minimal set of sufficient marginal tables (see e.g. Appendix II). Third, if the fit of two models is compared, the likelihood-ratio statistic may be calculated from the parameter estimates and their minimal sufficient statistics (see e.g. Appendix III).

Goodman (1964, 1968) describes an algorithm for the analysis of incomplete two way tables that calculates the parameter estimates from their sufficient statistics. The algorithm is easily generalized to analyze multi-way tables, but convergence is slow. For the case of the Rasch model, very efficient algorithms solving the likelihood equations (Appendix II, Equation II.4) in terms of minimal sufficient marginals are already developed by Fischer (1974, chapter 14) and Gustafsson (1977). These algorithms only compute the item parameters, but from Equation II.3 (Appendix II) the subgroup \times score parameters in the loglinear Rasch model can be calculated directly. The Fischer-Gustafsson algorithms can also be used with models (e.g. 17) that can be broken down into a set of

separate Rasch models e.g. models with different item parameters, within each subgroup or within each scoregroup (see also Andersen, 1980a, p. 251) or both. For other models, e.g. models with parameters describing interactions between item responses the algorithm is useless.

Another problem in the analysis of large Rasch tables concerns the approximation of the overall goodness-of-fit statistics $G^2(m; f)$ and $Q(m)$ to the chi-square distribution. If the expected cell counts become small, this approximation is known to be bad (Lancaster, 1961). A standard way to deal with this situation is to perform a grouping procedure (Andersen, 1980a, p. 94), that is, to add cells together to obtain a grouped table with higher expected counts. Although grouping procedures improve the agreement of the goodness-of-fit statistics with the chi-square distribution, this is achieved at the cost of a considerable loss of information about the structure of the data. Therefore, the statistics, being only sensitive to specific violations, may no longer be used as an overall test.

In conclusion, the loglinear model provides an overall test of the Rasch model if the number of items is not too large, and it provides a flexible parametric way of stating and testing several important departures from the Rasch model by comparing it with less restrictive quasi-loglinear models. Obviously, the class of quasi-loglinear models does not exhaust all possible alternative models against which one might want to test the Rasch model. For example, models with discrimination and guessing parameters cannot be formulated as quasi-loglinear models. It should, however, be noted that the quasi-loglinear model permits various generalisations of the Rasch model. For example, the Rasch table may have multiple subgroup variables, polychotomously scored item responses, there may be a linear structure on the item parameters (as in the linear logistic test model, see Fischer, 1983), or several Rasch scales may be analyzed simultaneously, e.g., to test whether the scales measure the same trait (Goldstein, 1980, sec. 4). Furthermore, Fienberg's incomplete table methodology seems preeminently suited to analyze Rasch tables with a priori missing entries, e.g. data where groups of individuals have made different but overlapping sets of items. Consequently, incomplete table methodology may be used with item sampling (Shoemaker, 1973), used for the vertical equating of Rasch scales (Lloyd & Hoover, 1980), and other practical applications of the Rasch model.

APPENDIX I: Constraints on the σ^* -parameters in Model 11

Inserting (8) in (12) and taking exponentials we have

$$\exp(\sigma_{it}^*) = n_i \int_{-\infty}^{+\infty} \frac{\exp(\alpha t)}{\prod_{j=1}^k (1 + \exp(\alpha - \delta_j))} dF_i(\alpha) \quad (I.1)$$

let

$$\phi = \exp(\alpha) \quad \text{and} \quad \varepsilon_j = \exp(-\delta_j), \quad j = 1, \dots, k$$

and let $H(\phi)$ be the distribution of ϕ on the interval $(0, \infty)$, then

$$\exp(\sigma_{it}^*) = n_i \int_0^{\infty} \frac{\phi^t}{\prod_{j=1}^k (1 + \phi \varepsilon_j)} dH_i(\phi) \quad (I.2)$$

Cressie and Holland (1981) show that the differential element

$$dG_i(\phi) = dH_i(\phi) / \left[P_i(\mathbf{0}) \prod_{j=1}^k (1 + \phi \varepsilon_j) \right] \quad (I.3)$$

defines a new distribution function $G_i(\phi)$. From (I.2) and (I.3) we have

$$\exp(\sigma_{it}^*) = n_i P_i(\mathbf{0}) \int_0^\infty \phi^t dG_i(\phi) = n_i P_i(\mathbf{0}) \mu_i(t) \tag{I.4}$$

where $\mu_i(t)$ is the t th moment of distribution $G_i(\phi)$ in subgroup i . Following Karlin and Studden ((1966) Theorem 10.1, chapter V), Cressie and Holland (1983) show that the moment $\mu_i(t)$ of an arbitrary positive random variable must satisfy the inequality constraints

$$\det. (\mathbf{A}_{i\ell}) \geq 0, \tag{I.5}$$

where

$$\mathbf{A}_{i\ell} = \begin{cases} \left\| \mu_i(r+s) \right\|_{r,s=0}^{\ell/2} & \text{if } \ell \text{ is even,} \\ \left\| \mu_i(r+s+1) \right\|_{r,s=0}^{(\ell-1)/2} & \text{if } \ell \text{ is odd,} \end{cases}$$

$i = 1, \dots, m; \ell = 1, \dots, k$. That is all matrices $\mathbf{A}_{i\ell}$ must be positive semi-definite. The set of restrictions I.5, however, are redundant for $\ell < k - 1$ since $\det. (\mathbf{A}_{i(\ell-2)})$ is a principal minor of $\mathbf{A}_{i\ell}$ and if $\mathbf{A}_{i\ell}$ is positive semi-definite every principal minor of $\mathbf{A}_{i\ell}$ is non-negative (see Ayres, 1974, chapter 17, theorem XIII). Furthermore multiplying each element of $\mathbf{A}_{i\ell}$ by the positive numbers $n_i P_i(\mathbf{0})$ and using (I.4) gives the result, since in if Y is positive

$$\det. (Y \cdot \mathbf{A}_{n \times n}) = Y^n \det. (\mathbf{A}_{n \times n}) \geq 0 \text{ is equivalent to } \det. (\mathbf{A}_{n \times n}) \geq 0.$$

APPENDIX II: CML Estimates in the Loglinear Rasch Model

Using Haberman (1979) we can write the maximum likelihood equations of the loglinear Rasch model (13) as:

$$f_{it+\dots+} = m_{it+\dots+} \text{ and } f_{+\dots+x_j+\dots+} = m_{+\dots+x_j+\dots+}$$

for $i = 1, \dots, m; t = 0, \dots, k$ and $x_j = 0, 1; j = 1, \dots, k$, where a plus sign is defined as before. Expressing this in terms of the model parameters (13) yields:

$$\begin{aligned} f_{it+\dots+} &= \sum_{[x_1 \dots x_k]} \dots \sum \exp \left\{ \sigma_{it} - \sum_j x_j \delta_j \right\}; & \sum_j x_j &= t \\ &= \exp \{ \sigma_{it} \} \cdot \sum_{[x_1 \dots x_k]} \dots \sum \prod_j \exp \{ -x_j \delta_j \}; & \sum_j x_j &= t \\ &= \exp \{ \sigma_{it} \} \gamma_t(\delta_1, \dots, \delta_k) \end{aligned} \tag{II.1}$$

for $i = 1, \dots, m; t = 0, \dots, k$. And, for $x_j = 1$:

$$\begin{aligned} f_{+\dots+1_j+\dots+} &= \sum_{i \geq 1} \sum_t \sum_{[x_1 \dots x_k]} \dots \sum \exp \left\{ \sigma_{it} - \sum_l x_l \delta_l \right\}; & \sum_l x_l &= t; x_j = 1 \\ &= \sum_{i \geq 1} \sum_t \exp \{ \sigma_{it} \} \exp \{ -\delta_j \} \sum_{[x_1 \dots x_k]} \dots \sum \prod_l \exp \{ -x_l \delta_l \}; & \sum_l x_l &= t - 1; \ell \neq j \\ &= \sum_{i \geq 1} \sum_t \exp \{ \sigma_{it} \} \exp \{ -\delta_j \} \gamma_{t-1}(\delta_1, \dots, \delta_{j-1}, \delta_{j+1}, \dots, \delta_k) \end{aligned} \tag{II.2}$$

for $j = 1, \dots, k$, where $\gamma_t(\delta_1, \dots, \delta_k)$ and $\gamma_{t-1}(\delta_1, \dots, \delta_{j-1}, \delta_{j+1}, \dots, \delta_k)$ are elementary

symmetric functions. Solving σ_{it} from (II.1), i.e.,

$$\sigma_{it} = \ln f_{it+\dots+} - \ln \gamma_t(\delta_1, \dots, \delta_k) \tag{II.3}$$

and using it in (II.2) we get

$$f_{+\dots+1_j+\dots+} = \exp \left\{ -\delta_j \right\} \sum_i \sum_{t \geq 1} f_{it+\dots+} \frac{\gamma_{t-1}(\delta_1, \dots, \delta_{j-1}, \delta_{j+1}, \dots, \delta_k)}{\gamma_t(\delta_1, \dots, \delta_k)} \tag{II.4}$$

for $j = 1, \dots, k$; which is identical to the CML equations (Andersen, 1973a). The estimates of item parameters in the loglinear Rasch model are therefore identical to the CML estimates.

APPENDIX III: Andersen's Conditional Likelihood Ratio Test

For reasons of comparability we write the loglinear Rasch model and the comparison model (17) of test b, Table 3, as:

$$\ln \hat{m}_{itx_1 \dots x_k} = \hat{\sigma}_{it} - \sum_j \underline{x}_j \hat{\delta}_j \tag{III.1}$$

and

$$\ln \tilde{m}_{itx_1 \dots x_k} = \tilde{\sigma}_{it} - \sum_j x_j \tilde{\delta}_j^{(t)} \tag{III.2}$$

respectively, where (III.2) is an equivalent parametrisation of model (17) with $\tilde{\delta}_j^{(t)}$ the estimated difficulty of item j in scoregroup t . The likelihood equations of model (III.2) are (Haberman, 1979):

$$\tilde{m}_{it+} = f_{it+} \tag{III.3}$$

and

$$\tilde{m}_{+t+x_j+} = f_{+t+x_j+} \tag{III.4}$$

From (III.2) and (III.3), it follows (analogously to (II.1) and (II.3)) that

$$\tilde{\sigma}_{it} = \ln f_{it+} - \ln \gamma_t(\tilde{\delta}^{(t)}) \tag{III.5}$$

The likelihood-ratio statistic (19) can now be written as:

$$\begin{aligned} G^2(\hat{m}; \tilde{m}) &= 2 \sum_i \sum_t \sum_x \tilde{m}_{itx} \ln (\tilde{m}_{itx} / \hat{m}_{itx}) \\ &= 2 \sum_t \sum_i \sum_x \tilde{m}_{itx} [(\tilde{\sigma}_{it} - \sum_j x_j \tilde{\delta}_j^{(t)}) - (\hat{\sigma}_{it} - \sum_j x_j \hat{\delta}_j)] \\ &= 2 \sum_t \sum_i \sum_x \tilde{m}_{itx} [(-\ln \gamma_t(\tilde{\delta}^{(t)}) - \sum_j x_j \tilde{\delta}_j^{(t)}) - (-\ln \gamma_t(\hat{\delta}) - \sum_j x_j \hat{\delta}_j)] \\ &= 2 \sum_t [(-f_{+t+} \ln \gamma_t(\tilde{\delta}^{(t)}) - \sum_j f_{+t+1_j+} \tilde{\delta}_j^{(t)}) \\ &\quad - (-f_{+t+} \ln \gamma_t(\hat{\delta}) - \sum_j f_{+t+1_j+} \hat{\delta}_j)] \\ &= 2 \sum_t [(\ln L_c^{(t)}(\tilde{\delta}^{(t)}) - \ln L_c^{(t)}(\hat{\delta}))] = Z_c \end{aligned}$$

where the first equation follows by substitution of (III.1) and (III.2), the second equation follows by substitution of (III.5) the third follows by substitution of (III.3) and (III.4). The fourth equation proves the equality with Andersen's conditional likelihood ratio test,

where

$$L_c^{(t)}(\delta^{(t)}) = \frac{\exp\left(-\sum_j f_{+t+1j} + \delta_j^{(t)}\right)}{\gamma_t(\delta^{(t)})^{f_{+t+}}}$$

and

$$L_c^{(t)}(\delta) = \frac{\exp\left(-\sum_j f_{+t+1j} + \delta_j\right)}{\gamma_t(\delta)^{f_{+t+}}}$$

are Andersen's (1980, sec. 6.6) conditional likelihood functions and Z_c his conditional likelihood ratio statistics with $(k-2)(k-1)$ degrees of freedom.

REFERENCES

- Andersen, E. B. (1971). Asymptotic properties of conditional likelihood ratio tests. *Journal of the American Statistical Association*, 66, 630-633.
- Andersen, E. B. (1972). The numerical solution of a set of conditional estimation equations. *Journal of the Royal Statistical Society, Series B*, 34, 42-54.
- Andersen, E. B. (1973a). *Conditional inference models for measuring*. Copenhagen: Mentalhygiejnisk Forlag.
- Andersen, E. B. (1973b). A goodness of fit test for the Rasch model. *Psychometrika*, 38, 123-140.
- Andersen, E. B. (1980). *Latent structure Analysis*. Invited paper to the European Meeting of the Psychometric Society, Groningen.
- Andersen, E. B. (1980a). *Discrete statistical models with social science applications*. Amsterdam: North Holland.
- Andersen, E. B. (1980b). Comparing latent distributions. *Psychometrika*, 45, 121-134.
- Ayres, F. (1974). *Matrices*. New York: McGraw Hill.
- Baker, F. B. & Subkoviak, M. J. (1981). Analysis of test results via loglinear models. *Applied Psychological Measurement*, 5, 503-515.
- Baker, R. J. & Nelder, J. A. (1978). *The GLIM system: Generalized linear interactive modelling*. Oxford: The Numerical Algorithms Group.
- Bishop, Y. M. M., Fienberg, S. E. & Holland, P. W. (1975). *Discrete multivariate analysis*. Cambridge, Mass.: MIT Press.
- Bock, R. D. (1975). *Multivariate Statistical Methods in behavioral research*. New York: McGraw Hill.
- Clogg, C. C. & Sawyer, D. A. (1981). A comparison of alternative models for analyzing the scalability of response patterns. In Samuel Leinhardt (Ed.) *Sociological Methodology*. San Francisco: Jossey-Bass.
- Coombs, C. H. (1964). *A theory of data*. New York: Wiley.
- Cressie, N. & Holland, P. W. (1983). Characterizing the manifest probabilities of latent trait models. *Psychometrika*, 48, 129-142.
- Davison, M. L. (1979). Testing a unidimensional qualitative unfolding model for attitudinal or developmental data. *Psychometrika*, 44, 179-194.
- Davison, M. L. (1980). A psychological scaling model for testing order hypotheses. *British Journal of Mathematical and Statistical Psychology*, 33, 123-141.
- Dayton, C. M. & McReady, G. B. (1980). A scaling model with response errors and intrinsically unscalable respondents. *Psychometrika*, 45, 343-356.
- Evers, M. & Nambodiri, N. K. (1979). On the design matrix strategy in the analysis of categorical data. In Karl F. Schuessler (Ed.) *Sociological Methodology*. San Francisco: Jossey-Bass.
- Fienberg, S. E. (1972). The analysis of incomplete multi-way contingency tables. *Biometrics*, 28, 177-202. Corrig. 1973, 29, 829.
- Fienberg, S. E. (1980). *The analysis of cross-classified categorical data*. Cambridge Mass.: MIT Press.
- Fischer, G. H. (1974). *Einführung in die Theorie Psychologischer Tests (Introduction to the theory of psychological tests)*. Bern: Huber, (In German).
- Fischer, G. H. (1981). On the existence and uniqueness of maximum likelihood estimates in the Rasch model. *Psychometrika*, 46, 59-77.
- Fischer, G. H. (1978). Probabilistic test models and their applications. *The German Journal of Psychology*, 2, 298-319.
- Fischer, G. H. (1983). Logistic latent trait models with linear constraints. *Psychometrika*, 48, 3-26.

- Fischer, G. H. & Scheiblechner, H. (1970). Algorithmen und Programma für das probabilistische Testmodell von Rasch. *Psychologische Beiträge*, 12, 23–51 (in German).
- Goldstein, H. (1980). Dimensionality, bias, independence and measurement scale problems in latent trait test score models. *British Journal of Mathematical and Statistical Psychology*, 33, 234–246.
- Goodman, L. A. (1959). Simple statistical methods for scalogram analysis. *Psychometrika*, 24, 29–43.
- Goodman, L. A. (1964). A short computer program for the analysis of transaction flows. *Behavioral Science*, 9, 176–186.
- Goodman, L. A. (1968). The analysis of cross-classified data: Independence, quasi-independence, and interactions in contingency tables with or without missing entries. *Journal of the American Statistical Association*, 63, 1091–1131.
- Goodman, L. A. (1974). Exploratory latent structure analysis. *Biometrika*, 61, 215–231.
- Goodman, L. A. (1975). A new model for scaling response patterns: An application of the quasi-independence concept. *Journal of the American Statistical Association*, 70, 755–768.
- Goodman, L. A. (1978). *Analyzing qualitative/categorical data: Loglinear models and latent structure analysis*. London: Addison Wesley.
- Goodman, L. A. & Fay, R. (1974). *ECTA program, description for users*. Chicago: Department of Statistics University of Chicago.
- Gustafsson, J. E. (1977). *The Rasch model for dichotomous items: Theory, applications and a computer program*. Reports from the Institute of Education, University of Göteborg, no. 63, ED154018.
- Gustafsson, J. E. (1980). Testing and obtaining the fit of data to the Rasch model. *British Journal of Mathematical and Statistical Psychology*, 33, 205–233.
- Guttman, L. (1950). The basis for scalogram analysis. In S. A. Stouffer et al (Eds.), *Measurement and prediction: Studies in social psychology in World War II*. Vol. 4, Princeton: Princeton University Press.
- Haberman, S. J. (1978). *Analysis of qualitative data*, Vol. 1. New York: Academic Press.
- Haberman, S. J. (1979). *Analysis of qualitative data: New developments*, Vol. 2. New York: Academic Press.
- Hambleton, R. K. Swaminatan, H., Cook, L. L., Eignor, D. R. & Gifford, J. A. (1978). Developments in latent trait theory: Models technical uses and applications. *Review of Educational Research*, 48, 467–510.
- Karlin, S. & Studden, W. J. (1966). *Tchebycheff systems: with applications in analysis and statistics*. New York: Interscience Publishers.
- Kelderman, H., Mellenbergh, G. J. & Elshout, J. J. (1981). Guilford's facet theory of intelligence: An empirical comparison of models. *Multivariate Behavioral Research*, 16, 37–62.
- Kempf, W. (1974). Dynamische Modelle zur Messung sozialer Verhaltensdispositionen. (Dynamic models for attitude measurement.) In: W. Kempf (Ed.) *Probabilistische Modelle in der Sozialpsychologie (Probabilistic models in social psychology)*, Bern: Huber, 13–55 (in German).
- Lancaster, H. O. (1961). Significance tests in discrete distributions. *Journal of the American Statistical Association*, 56, 223–234.
- Lazersfeld, P. F. (1950). The interpretation and computation of some latent structures. In Samuel A. Stouffer et al. (Eds.), *Measurement and prediction in World War II*, Vol. 4. Princeton: Princeton University Press.
- Lazersfeld, P. F. & Henry, N. W. (1968). *Latent structure analysis*, Boston: Houghton-Mifflin.
- Lienert, G. A. & Raatz, U. (1981). Item homogeneity defined by multivariate axial symmetry. *Applied Psychological Measurement*, 5, 263–269.
- Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, Mass.: Addison-Wesley.
- Loyd, B. H. & Hoover, H. D. (1980). Vertical equating using the Rasch model. *Journal of Educational Measurement*, 17, 179–193.
- Lumsden, J. (1978). Tests are perfectly reliable. *British Journal of Mathematical and Statistical Psychology*, 31, 19–26.
- McDonald, R. P. & Krane, W. R. (1979). A Monte Carlo study of local identifiability and degrees of freedom in the asymptotic likelihood ratio test. *British Journal of Mathematical and Statistical Psychology*, 32, 121–132.
- McHugh, R. B. (1956). Efficient estimation and local identification in latent class analysis. *Psychometrika*, 21, 331–347.
- Mellenbergh, G. J. (1972). Applicability of the Rasch model in two cultures. In L. J. Cronbach & P. J. D. Drenth (Eds.), *Mental tests and cultural adaptation*. The Hague: Mouton.
- Mellenbergh, G. J. & Vijn, P. (1981). The Rasch model as a loglinear model. *Applied Psychological Measurement*, 5, 369–376.
- Molenaar, I. W. (1983). Some improved diagnostics for failure of the Rasch model. *Psychometrika*, 48, 49–73.
- Mood, A. M., Graybill, F. A. & Boes, D. C. (1974). *Introduction to the theory of statistics*. Tokyo: McGraw-Hill Kogakuska.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Paedagogiske Institut.
- Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In J. Neyman (Ed.), *Proceedings of the Fourth Symposium on Mathematical Statistics and Probability*, Vol. 4. Berkeley: University of California Press, 321–333.

- Rasch, G. (1966a). An item analysis that takes individual differences into account. *British Journal of Mathematical and Statistical Psychology*, 19, 49–57.
- Rasch, G. (1966b). An individualistic approach to item analysis. In P. F. Lazarsfeld & N. W. Henry (Eds.), *Readings in Mathematical Social Science*. Cambridge, Mass.: MIT Press, 89–107.
- Renz, R. R. & Bashaw, W. L. (1977). The National reference scale for reading: An application of the Rasch model. *Journal of Educational Measurement*, 14, 161–179.
- Rudner, L. M., Getson, P. R. & Knight, D. L. (1980). Biased item detection techniques. *Journal of Educational Statistics*, 5, 213–233.
- Shoemaker, D. M. (1973). *Principles and procedures of multiple matrix sampling*. Cambridge, Mass.: Ballinger.
- Stouffer, S. A. (1950). An overview of the contributions to scaling and scale theory. In S. A. Stouffer et al. (Eds.), *Measurement and prediction: Studies in Social Psychology in World War II*, (Vol. 4), Princeton: Princeton University Press.
- Tjur, T. (1982). *A connection between Rasch's item analysis model and a multiplicative Poisson model*. *Scandinavian Journal of Statistics*, 9, 23–30.
- Van den Wollenberg, A. L. (1979). *The Rasch model and time-limit tests*. Doctoral dissertation, Nijmegen.
- Van den Wollenberg, A. L. (1982). Two new test statistics for the Rasch model. *Psychometrika*, 47, 123–140.
- Vijn, P. & Mellenbergh, G. J. (1982). *The Rasch model versus the loglinear model*. Révész Berichten no. 42, Psychologisch Laboratorium, Universiteit van Amsterdam.
- Wood, R. (1978). Fitting the Rasch model—A heady tale. *British Journal of Mathematical and Statistical Psychology*, 31, 27–32.
- Wright, B. D. & Mead, R. J. (1977). *BICAL: Calibrating items and scales with the Rasch model*. Research Memorandum, 23, Statistical Laboratory, Department of Education, University of Chicago.
- Wright, B. D. & Panchapakesan, N. (1969). A procedure for sample-free item analysis. *Educational and Psychological Measurement*, 29, 23–48.

Manuscript received 3/24/82

First revision received 6/14/83

Final revision received 1/4/84