ADVANCES IN LONGITUDINAL AND MULTIVARIATE ANALYSIS IN THE BEHAVIORAL SCIENCES

edited by

JOHAN H.L. OUD

and

RIAN A.W. VAN BLOKLAND-VOGELESANG

its

Chapter 13

LOG-TRILINEAR MODELS FOR THREE-WAY CONTINGENCY TABLES

Ab Mooijaart
Leiden University

and

Peter G.M. van der Heijden
University of Utrecht

Abstract

In association models for two-way contingency tables, the categories of the variables can be quantified optimally by some well-defined criterion. An analogous quantification will be carried out in the case of three-way contingency tables. However, here some conceptual problems arise. These problems will be discussed in this contribution. Our general formulation of the model implies that other well-known models are special cases of our model. A real life data example will be discussed.

Index terms: log-trilinear models, three-way contingency tables, association models, low rank decomposition, PARAFAC.

1. Introduction

In the sixties and seventies a standard approach of log-linear analysis was developed. This led to widely available software for log-linear analysis, such as in SPSS-X and BMDP, and a number of standard works, such as Bishop, Fienberg and Holland (1975), Goodman (1978), and Haberman (1978, 1979).

Starting from a two-way contingency table with rows indexed by i (i = 1, ..., I) and columns indexed by j (j = 1, ..., J), the expected frequencies m_{ij} for the saturated log-linear model can be written as

$$m_{ij} = \exp\,[u + u_{1(i)} + u_{2(j)} + u_{12(ij)}] \ ,$$

where parameters add up to zero over each index. Since then, a new breakthrough that stimulated much research was a particular way to constrain the interaction parameters in a two-way contingency table. Goodman (1979) and Andersen (1980) proposed a constrained model for the interaction parameters, namely

$$m_{ij} = \exp\left[u + u_{1(i)} + u_{2(j)} + \lambda a_i b_j\right] ,$$

which was coined by Goodman (1979) the RC-model. Later in Goodman (1985, 1986) it was coined the RC-association model. The parameters a_i and b_j can be interpreted as scores for rows i and columns j, normalized by $\Sigma_i a_i = \Sigma_j b_j = 0$ and $\Sigma_i a_i^2 = \Sigma_j b_j^2 = 1$, and λ can be interpreted as a parameter for the strength of the association as described by a_i and b_j. The bilinear term $\lambda a_i b_j$ results in a model that is not log-linear but log-bilinear. Different motivations exist for this model. Goodman (1979) himself describes this model emphasizing an interpretation in terms of odds ratios. Later it becomes clear that if the two-way contingency table is a discretization of a bivariate normal distribution, then a_i and b_j are the quantifications for the categories indexed by i and j, and λ approximates the product moment correlation coefficient between the two variables (Goodman, 1981). A third interpretation is that the bilinear term $\lambda a_i b_j$ gives a rank 1 approximation of the matrix of association parameters $u_{12(ij)}$. All three interpretations indicate that these models might work out well in the situation that the categories of the row and the categories of the column variable are ordered, where the order is not necessarily known in advance.

In later papers (Goodman, 1985, 1986) this model was extended to the RC(P)-association model

$$m_{ij} = \exp\left[u + u_{1(i)} + u_{2(j)} + \sum_{p=1}^{P} \lambda_p a_{ip} b_{jp}\right] ,$$

where the matrix of association parameters $u_{12(ij)}$ is now described by a rank P matrix.

The ideas behind these models for the two-way association can be generalized to three-way contingency tables. Different kinds of generalizations for two-way association models to three-way association models were given by Clogg (1982a, 1982b), Agresti and Kezouh (1983), Goodman (1979, 1986), Gilula and Haberman (1988), Choulakian (1988), Becker and Clogg (1989), Becker (1989), and Mooijaart (1991). These papers deal with ordered categories, or explanatory and response variables, or a combination of log-bilinear models. Furthermore, in all these papers the maximum likelihood method is used for estimating the parameters. In D'Ambra and Kiers (1990) a least squares method is used for estimating the parameters in log-trilinear models.

The main issue in this paper is the decomposition of the three-way contingency table by a log-trilinear model where all variables are response variables. Some of the models discussed here are also discussed by Goodman (1986) and Choulakian (1988). Our general formulation of the model implies that other models are special cases of our model.

In association models for two-way contingency tables, the categories of the variables are quantified optimally, by some well defined criterion. An analogous quantification will be carried out in the case of three-way contingency tables. However, here some conceptual problems arise. For instance, decomposing the interaction between the first and the second variable by a log-bilinear model yields some quantification for the categories of the first variable. If the interaction between all three variables are decomposed by a log-trilinear model, some quantification for the categories of the first variable is computed too. The interesting problem then is: should the quantifications for the categories of the first variable be equal both for the decomposition of the two-factor interaction and for the three-factor interactions? We will show that a positive answer to this question will lead to a model that has a fruitful interpretation in the context of sets of two-way contingency tables.

2. Formulation of the model

Let n_{ijk} denote the observed frequency in the (i,j,k)-th cell of an $I \times J \times K$ contingency table, with $i = 1, \ldots, I$; $j = 1, \ldots, J$; $k = 1, \ldots, K$. The corresponding expected frequency will be denoted by m_{ijk}. The (one-dimensional) marginals of the first, second and third variables of the table with proportions will be denoted as p_{i++}, p_{+j+}, and p_{++k}, respectively. The model we assume for these expected frequencies is

$$m_{ij} = \exp\left[u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)} + u_{123(ijk)}\right] \ , \qquad (1)$$

where the u parameters are the main effects or interaction parameters as in log-linear analysis. Like in the well-known RC(P)-association model for two-dimensional tables of Goodman (1986), the two-factor interaction parameters will be decomposed. However, in three-dimensional tables there are three types of two-factor interactions. In the most general form we can write these interactions as

$$u_{12(ij)} = \sum_{p=1}^{P} \lambda_{12(p)} a_{ip} b_{jp} \ , \qquad (2a)$$

$$u_{13(ik)} = \sum_{q=1}^{Q} \lambda_{13(q)} c_{iq} d_{kq} \ , \qquad (2b)$$

$$u_{23(jk)} = \sum_{r=1}^{R} \lambda_{23(r)} e_{jr} f_{kr} \ . \qquad (2c)$$

In order to identify the model we impose the following centering restrictions:

$$\sum_{i=1}^{I} p_{i++} a_{ip} = \sum_{i=1}^{I} p_{i++} c_{iq} = \sum_{j=1}^{J} p_{+j+} b_{jp} = \sum_{j=1}^{J} p_{+j+} e_{jr} = \sum_{k=1}^{K} p_{++k} d_{kq} = \sum_{k=1}^{K} p_{++k} f_{kr} = 0 \ ,$$

and the following weighting restrictions:

$$\sum_{i=1}^{I} p_{i++} a_{ip} a_{ip'} = \sum_{j=1}^{J} p_{+j+} b_{jp} b_{jp'} = \delta^{pp'} \ ,$$

$$\sum_{i=1}^{I} p_{i++} c_{iq} c_{iq'} = \sum_{k=1}^{K} p_{++k} d_{kq} d_{kq'} = \delta^{qq'} \ ,$$

$$\sum_{j=1}^{J} p_{+j+} e_{jr} e_{jr'} = \sum_{k=1}^{K} p_{++k} f_{kr} f_{kr'} = \delta^{rr'} \ ,$$

where $\delta^{ss'} = 1$ if $s = s'$, and $\delta^{ss'} = 0$ else.

Two remarks have to be made. First, in (2a) – (2c) a reduced rank is specified for the two-factor interactions. However, these interactions can also be specified as completely free. For instance, if in (2a) P = min(I – 1, J – 1) then the I × J matrix with elements $u_{12(ij)}$ has maximum rank because of the centering restrictions and so the two-factor interactions $u_{12(ij)}$ are specified to be completely free.

Second, the parameters a_{ip}, b_{jp}, c_{iq}, d_{kq}, e_{jr} and f_{kr} can be conceived as category quantifications of the three variables. Notice that in this way the variable indexed by i is quantified in different ways: P times by the quantifications a_{ip} and Q times by the quantifications c_{iq}. A model which is more restricted than Equation 2 is the model in which the quantifications of a variable are equal in all two-factor interactions. I.e., in this more restricted model the maximum rank for each of these matrices is equal, i.e., P = Q = R, and $a_{ip} = c_{ip}$, $b_{jp} = e_{jp}$ and $d_{kp} = f_{kp}$. A proposal for such a model was given by Choulakian (1988). In this paper we will deal with the less restrictive model that the two-factor interactions are of reduced rank specified by Equations 2a, 2b and 2c. However, by imposing restrictions like P = Q = R, and $a_{ip} = c_{ip}$, $b_{jp} = e_{jp}$ and $d_{kp} = f_{kp}$ we can deal with more restricted models as well.

There are several ways for decomposing the three-factor interactions. We found inspiration for choosing specific decompositions in psychometrics, where several techniques exist for the decomposition of three-way matrices. For instance,

$$u_{123(ijk)} = \sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{u=1}^{U} \lambda_{stu} x_{is} y_{jt} z_{ku} \quad . \tag{3}$$

This decomposition is analogous to one of Tucker's decompositions for three-way data, see Tucker (1964, 1966) and Kroonenberg (1983). One problem with this decomposition is that identification of the parameters x, y, and z is not simple.

Therefore we choose a special case of this decomposition, which is equivalent to the PARAFAC decomposition, and which imposes that all λ's are zero except for s = t = u. Thus for the scores x_{is}, y_{jt} and z_{ku}, it holds S = T = U. See for a detailed discussion of the PARAFAC model Harshman and Lundy (1984a,1984b). So we can write

$$u_{123(ijk)} = \sum_{s=1}^{S} \lambda_s x_{is} y_{js} z_{ks} \quad . \tag{4}$$

We impose the following centering and weighting restrictions:

$$\sum_{i=1}^{I} p_{i++} x_{is} = \sum_{j=1}^{J} p_{+j+} y_{js} = \sum_{k=1}^{K} p_{++k} z_{ks} = 0 \quad ,$$

and

$$\sum_{i=1}^{I} p_{i++} x_{is}^2 = \sum_{j=1}^{J} p_{+j+} y_{js}^2 = \sum_{k=1}^{K} p_{++k} z_{ks}^2 = 1 \quad .$$

Notice that we only fix the scales for the parameters x_{is}, y_{js} and z_{ks}, and we do not restrict parameters of x_{is}, of y_{js} and or z_{ks} to be orthogonal for different s, like we did for the two-factor interactions. A problem of the PARAFAC model may be the existence of degenerate solutions (see Kruskal, Harshman, & Lundy, 1989). A typical degeneracy is that subsequent solutions for s = 1, 2, ... are very similar. To overcome such degeneracies one can assume orthogonality restrictions for one or two sets of the parameters x_{is}, y_{js} and or z_{ks}. We will not discuss these possibilities here.

3. Estimation of parameters

The parameters of the model will be estimated by the maximum likelihood method. Under Poisson sampling the kernel of the log likelihood function can be written as

$$\log L = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} n_{ijk}\log m_{ijk} - \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} m_{ijk} \ . \tag{5}$$

This likelihood function should be optimized with respect to the parameters. The two and three-factor interactions in the equation above are functions of the unknown parameters, see Equations 2 and 4. Estimation of these parameters will be done by some alternating estimation procedure.

Two remarks have to be made before, however. First, the centering restrictions may not be important in the algorithm. For instance, let the scores x_{is} be shifted with a constant. Then Equation 4 can be written as

$$u_{123(ijk)} = \sum_{s=1}^{S} \lambda_s (x_{is} + \delta_s) y_{js} z_{ks} - \sum_{s=1}^{S} \lambda_{123(s)} \delta_s y_{js} z_{ks} \ . \tag{6}$$

In the first term on the right hand side of Equation 6 the scores x_{is} are shifted with a constant, whereas the second term on the right hand side is in fact a two-factor interaction between the second and third variable. Depending on restrictions which hold for the two-factor interactions $u_{23(jk)}$ a simple correction may be applied without changing the expected frequencies m_{ijk} as given in Equation 1.

We can distinguish three situations.

A. The two-factor interaction parameters $u_{23(jk)}$ are completely free. In this case, the two-factor interaction $u_{23(jk)}$ has to be corrected with the second term of Equation 6.
B. The quantifications of variable 2 and 3 are equal in both the two- and three-factor interaction, i.e. S = R, and furthermore $e_{js} = y_{js}$, and $f_{ks} = z_{ks}$. The same correction as in A has to be applied.
C. In the general case where neither A nor B hold (for instance when some parameters are fixed to some constant) a shift of mean of the parameters x_{is} for each value of s can not be captured by a correction of lower factor interactions without changing the expected frequencies. In this case the algorithm becomes more complicated.

Of course, analogous results hold for a shift of the scores y_{js} and z_{ks} for each s. In addition, a shift of the scores of the two-factor interactions, for instance the scores a_{ip}, can also be applied. In this latter case a correction of some main effects parameters has to be applied.

The second remark we make is that the estimation of the λ parameters is not necessary in the first part of the algorithm, because the λ parameters are in fact just the result of some specific scaling of the scores. For instance, according to Equation 2a we can write $u_{12(ij)} = \Sigma_p \lambda_{12(p)} a_{ip} b_{jp} = \Sigma_p a_{ip}^* b_{jp}^*$ where $a_{ip}^* = a_{ip}/\lambda_{12(p)}^{1/2}$. So it is sufficient to estimate a_{ip}^* and b_{jp}^*, and afterwards decomposing $\Sigma_p a_{ip}^* b_{jp}^*$ to find a_{ip} and b_{jp} which meet the centering and weighting restrictions mentioned before. Such a decomposition of $\Sigma_p a_{ip}^* b_{jp}^*$ can be carried out by a generalized singular value decomposition (GSVD) (see Greenacre, 1984, Appendix).

The basic algorithm now runs as follows: Fix all the parameters for the scores, except one set of parameters, for example x_{i1}. Find a better estimate of the elements of these parameters, and repeat the whole procedure but now by improving the elements of some other set of parameters. Repeat the whole procedure till no change of the elements occurs.

As an example we show how to estimate the elements x_{is}. The method we use here is the unidimensional Newton-Raphson method, as was proposed by Goodman (1986). This method can be applied easily in case A and B mentioned above. In case C, or in cases where there are restrictions on the parameters, like for instance equality restrictions, a method discussed by Siciliano, Lauro and Mooijaart (1990) can be applied.

The first and the second derivatives of the log likelihood function (see Equation 5) with respect to the parameters x_{is}, are

$$l'(x_{is}) \equiv \frac{\partial \log L}{\partial x_{is}} = \sum_{j=1}^{J} \sum_{k=1}^{K} (n_{ijk} - m_{ijk}) y_{js} z_{ks} \ , \tag{7a}$$

$$l''(x_{is}) \equiv \frac{\partial^2 \log L}{(\partial x_{is})^2} = - \sum_{j=1}^{J} \sum_{k=1}^{K} m_{ijk} y_{js}^2 z_{ks}^2 \ . \tag{7b}$$

According to the unidimensional Newton-Raphson method a new update of the x scores can be obtained by

$$x_{is}^+ = x_{is} - \frac{l'(x_{is})}{l''(x_{is})} \ , \tag{8}$$

where the super-script "+" denotes the updated x score. So the updated x score can be written as

$$x_{is}^+ = x_{is} + \frac{\sum_{j=1}^{J} \sum_{k=1}^{K} (n_{ijk} - m_{ijk}) y_{js} z_{ks}}{\sum_{j=1}^{J} \sum_{k=1}^{K} m_{ijk} y_{js}^2 z_{ks}^2} \ . \tag{9}$$

Analogous formulae can be derived for other parameters.

The estimation method above can easily be generalized to models in which quantifications of categories of a variable are equal in different interactions. For instance, if $a_{is} = c_{is} = x_{is}$, some terms have to be added to the formulae for the first and second derivatives as given in Equation 7a and 7b. (These additional terms are very easy to derive.) Then applying Equations 8 and 9 gives update formulae which are analogous to the formulae given above.

4. Testing the model

Models can be tested by the likelihood ratio test, i.e. after the algorithm has converged we compute

$$G^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} n_{ijk}\log m_{ijk} - \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} n_{ijk}\log n_{ijk} \ . \qquad (10)$$

The G^2 value is chi square distributed with a number of degrees of freedom depending on a specific model. In table 1 the number of parameters to be estimated will be given. From this number the degrees of freedom can be derived.

Table 1
Number of parameters to be estimated

parameter	number of parameters	
	unrestricted	reduced rank
u	1	1
$u_{1(i)}$	$I - 1$	$I - 1$
$u_{2(j)}$	$J - 1$	$J - 1$
$u_{3(k)}$	$K - 1$	$K - 1$
$u_{12(ij)}$	$(I - 1)(J - 1)$	$P(I + J - 2) - P^2$
$u_{13(ik)}$	$(I - 1)(K - 1)$	$Q(I + K - 2) - Q^2$
$u_{23(jk)}$	$(J - 1)(K - 1)$	$R(J + K - 2) - R^2$
$u_{123(ijk)}$	$(I - 1)(J - 1)(K - 1)$	$S(I + J + K - 5)$

For instance, to determine the number of parameters for the interaction parameters $u_{12(ij)}$, we collect the parameters a_{ip} in a matrix **A** of order $I \times P$, and collect the parameters b_{jp} in a matrix **B** of order $J \times P$. Since $I > P$ and $J > P$, matrices **A** and **B** are of rank P. Taking into account the centering restrictions, the number of free elements in **A** and **B** are $P(I + J - 2)$. However, the product $\mathbf{AB}'$ is equal to $\mathbf{ASS}^{-1}\mathbf{B}'$, for any nonsingular matrix S of order $(P \times P)$. Therefore the total number of parameters to be estimated is $P(I + J - 2) - P^2$. Note that if $P = \min((I - 1), (J - 1))$ the number of parameters in the reduced case is equal to the number of parameters in the unrestricted case.

Analogously, the number of free elements in the matrices **X**, **Y**, and **Z** is equal to S(I + J + K − 3). Because the only non-uniqueness for these three matrices is the scaling of the columns, it holds that for each dimension in the three-factor interaction two elements are arbitrary, and may be fixed. Therefore the total number of elements to be estimated is here equal to S(I + J + K − 3) − 2S. The degrees of freedom can now be computed as IJK − the total number of estimated parameters.

5. An example

In the Netherlands children go at the age of 11-12 from primary school to secondary school. Distinct types of secondary education can be chosen, with two main types: vocational types of education and general types of education. A choice will depend on aspects such as, among others, capacities of children, interests, advice of the primary school teacher, advice of parents. In educational research much interest goes out to the way in which the social background of a child influences this choice.

In 1977 and 1981 data were collected from more than 37,000 children about their social background and aspects regarding their secondary education. Distinct variables were collected, see for a description CBS (1982) and Meester and De Leeuw (1983). We reanalyze part of the data that were published in Meester and De Leeuw (1983).

The variables we will use in our analysis are the scores on an intelligence test, sex and the level of education attained in 1981, i.e. after four years of secondary education. The intelligence test used was the (Dutch) Test for Intellectual Capacity (TIC), a figure exclusion test that consists of 33 items. The TIC scores were recoded by Meester and De Leeuw (1983) as 1 for 1 to 14 items correct, 2 for 15 to 17 correct, 3 for 18 to 20, 4 for 21 to 23, 5 for 24 to 26, 6 for 27 to 29, and 7 for 30 to 33 items correct. The response variable is the Level of education attained after 4 years, and these levels are 1. Dropped out (DO), 2. Junior vocational education (LBO), 3. General education, medium level (MAVO), 4. Senior vocational training (MBO), 5. General education, high level (HAVO) and 6. General education, preparing for university (VWO).

Meester and De Leeuw (1983) have eliminated all children having no TIC score (16,433 children). According to them, this elimination is not crucial because having no TIC score seemed to have been a random process. Further, children with a missing value on level of education attained (38) or on an education type called extraordinary lower education (646) were eliminated from the sample. Children having a father who is unemployed, or medically unfit for work were also eliminated (6,190). This is a more crucial elimination, and we should keep in mind that our analysis does not discuss children having these characteristics. After these selections there was a sample of 16,236 children.

Table 2 gives the three-dimensional frequency table, while Table 3 gives a summary of some models with their corresponding test statistics. In Table 3 "u" means that the corresponding interactions are unrestricted. The numbers 0, 1 and 2 mean that the corresponding interaction matrix has rank 0, 1 and 2, respectively. Note that the maximum rank of the interaction matrices (1,2) and (1,3) is 1. So for these interactions "unrestricted" is equal to rank 1. The sign "*" means that some special restrictions are imposed. We discuss these restrictions later on.

Table 2
Frequencies

Sex	boys							girls						
Intelligence	1	2	3	4	5	6	7	1	2	3	4	5	6	7
DO	75	77	105	125	89	38	17	51	60	115	123	78	56	9
LBO	216	305	495	522	389	168	34	144	223	382	370	290	107	26
MAVO	67	144	267	368	339	194	54	60	134	288	424	442	266	72
MBO	51	84	239	345	301	208	65	75	167	320	458	428	258	72
HAVO	26	65	200	332	383	258	98	23	68	211	373	450	402	169
VWO	12	27	104	216	325	321	178	5	9	77	183	307	326	209

Table 3
Models and test statistics

	Interaction						
	(1,2)	(1,3)	(2,3)	(1,2,3)	G^2	d.f.	significance
M_1	u	u	u	0	61.62	30	s
M_2	u	u	u	1	19.91	20	ns
M_3	u	u	2	1	28.02	32	ns
M_4	u	u	1	1	45.71	40	ns
M_5	u	u	2*	2*	42.02	38	ns
M_6	u	u	1*	1*	73.46	48	s

Table 3 shows that M_1, in fact a standard log-linear model with no three-factor interactions, does not fit, whereas M_2, the model with unrestricted two-factor interactions and one component for the three-factor interactions, i.e. $S = 1$, does fit. Model M_2 can be restricted further by imposing restrictions on the interaction matrix of variable 2 and 3. The table shows that model M_4, the model in which the rank of the interaction matrix of variable 2 and 3 is just 1 instead of the maximum rank 5, does fit the data adequately.

An interesting question is the following: we find an interaction between Intelligence and Final Educational Level; however, are the patterns of this interaction equal for boys and girls, more specifically, does this pattern of interaction for boys and girls only differ in strength?

For answering this question, the model can be written as

$$m_{ijk} = \exp [u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + \sum_{r=1}^{R} \lambda_{23(r)} e_{jr} f_{kr} + \sum_{r=1}^{R} \lambda_{23(r)} x_{ir} e_{jr} f_{kr}] \ , \qquad (11a)$$

where it is specified that the quantifications of the categories of the second and third variable are equal in the two and three-factor interaction terms. The x parameters in the three-factor interaction term refer to the differences between the sexes. From model M_1 we know that these parameters can not be 0, because a model in which the three-factor interactions vanish does not fit.

This model can be written as

$$m_{ijk} = \exp [u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + \sum_{r=1}^{R} \lambda_{23(r)}(x_{ir} + 1)e_{jr}f_{kr}] , \tag{11b}$$

Note that multiplying $\lambda_{23(r)}$ with a constant and dividing $x_{ir} + 1$ by the same constant, results in the same expected frequency m_{ijk}. Therefore, the x scores of one category may be fixed to some constant values. In this example we fix the x scores of the boys equal to zero.

In Table 3 model M_5 and M_6 refer to the model above with $R = 2$ and $R = 1$, respectively. From the test statistics we see that model M_5 does fit the data. The solutions of the parameters e_{jr}, f_{kr} and x_{ir} are given in Table 4. The λ parameters are 0.403 and 0.028 for dimension 1 and 2, respectively.

Tabel 4
Quantifications of the categories according to model M_5

Intelligence			Final educational level			Sex		
dimension	1	2		1	2		1	2
1	−2.231	0.455	DO	−1.502	1.322	Boys	0.000	0.000
2	−1.591	1.663	LBO	−1.180	−0.242	Girls	0.239	−1.444
3	−0.682	−1.422	MAVO	−0.313	0.646			
4	−0.121	−0.398	MBO	0.301	−1.880			
5	0.398	0.924	HAVO	0.613	0.702			
6	1.139	−0.712	VWO	1.757	0.502			
7	1.851	1.197						

Figure 1 and 2 give the plots of the solutions for the category quantifications of the two variables. Our interpretation of the results is as follows:

1. The first dimension gives the ordering of the categories of the variable Intelligence, from low intelligence to high intelligence. This ordering is just in line with the ordering of the Educational levels of the different school types.
2. The second dimension can be best interpreted from the variable Final Educational Level. It shows a difference between the Vocational Categories (in particular MBO) and 'drop out', the general educations lying in between. However, for the variable Intelligence this second dimension is hard to interpret. It seems that this second dimension captures some

peculiarities in the data. Although the test statistic shows that this second dimension is needed for an acceptable fit of the model, it is hard to interpret.

3. In Table 5 we give the three-factor interaction scores of the variables for the first dimension. Because the interaction parameters for the boys are set equal to zero, we give the interaction scores for the girls only. The scores in this table should be interpreted as follows: a positive score means that relatively more girls than boys fall in the specific

Figure 1
Intelligence

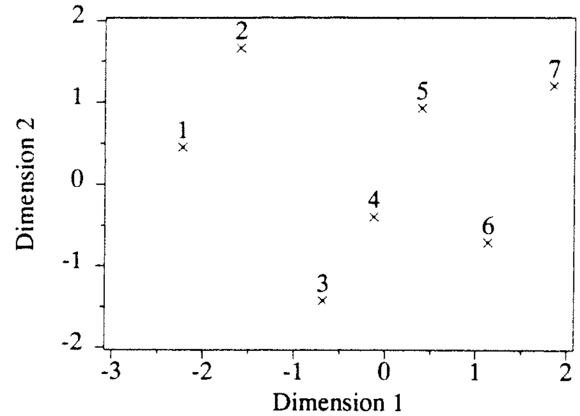
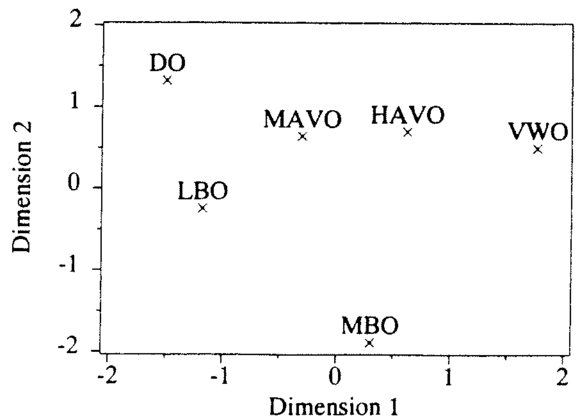


Figure 2
Final educational level



combination of row and column category. Remind the scores of boys are zero. On the other hand a negative score means just the opposite. The most interesting property of the table above is the nice ordering of the scores over the columns and the rows. This can be interpreted as

- In the lowest intelligence categories we see that relatively more girls than boys fall in the categories Drop Out, LBO and MAVO, and on the other hand we see, for this intelligence group, that relatively less girls than boys fall in the categories MBO, HAVO and VWO.

Table 5
Three-factor interactions for dimension 1 (girls only)

Intelligence Educ. Level	1	2	3	4	5	6	7
DO	0.32	0.23	0.10	0.02	−0.06	−0.17	−0.27
LBO	0.25	0.18	0.08	0.01	−0.05	−0.13	−0.21
MAVO	0.07	0.05	0.02	0.00	−0.01	−0.03	−0.06
MBO	−0.07	−0.05	−0.02	0.00	0.01	0.03	0.05
HAVO	−0.13	−0.09	−0.04	−0.01	0.02	0.07	0.11
VWO	−0.38	−0.27	−0.12	−0.02	0.07	0.19	0.31

- In the highest intelligence categories we see that relatively less girls than boys fall in the categories Drop Out, LBO and MAVO, and on the other hand we see, for this intelligence group, that relatively more girls than boys fall in the categories MBO, HAVO and VWO.

Table 6
Three-factor interactions for dimension 2 (girls only)

Intelligence Educ. Level	1	2	3	4	5	6	7
DO	−0.02	−0.09	0.08	0.02	−0.05	0.04	−0.06
LBO	0.00	0.02	−0.01	0.00	0.01	−0.01	0.01
MAVO	−0.01	−0.04	0.04	0.01	−0.02	0.02	−0.03
MBO	0.04	0.13	−0.11	−0.03	0.07	−0.05	0.09
HAVO	−0.01	−0.05	0.04	0.01	−0.03	0.02	−0.03
VWO	−0.01	−0.03	0.03	0.01	−0.02	0.01	−0.02

4. In Table 6 we give the three-factor interaction scores of the variables for the second dimension. Also here we give the interaction scores for the girls only.Obviously, because the second dimension is less important than the first dimension, the scores in table 6 are

closer to zero than the scores in table 5. From table 6 we see a less nice pattern of the scores over the columns and rows. The interpretation should be based here on specific combination of categories of the two variables Intelligence and Final Educational Level. For instance, in category 2 of Intelligence, more girls than boys fall in the category MBO, whereas in category 3 of Intelligence, more boys than girls fall in the category MBO.

6. Conclusion

In this contribution we have formulated a model for three-way contingency tables in which the two- and three-factor interactions are decomposed into a low rank matrix. This model is quite general. Sub-models can be specified by imposing restrictions on the parameters. It has been shown by a worked-out example that a simple log-trilinear model, in which the three-factor interactions differ from the two-factor interactions by just one parameter for each dimension, can result in a nice solution which fits the data well. Our conclusion is that association models for three-way contingency tables extend the two-way association models in a simple way and may be very suitable in analyzing real life data.

References

Agresti, A., & Kezouh, A. (1983). Association models for multidimensional cross-classifications of ordinal variables. *Communications in Statistics, Part A - Theory and Methods*, *12*, 1261-1276.

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

Becker, M.P. (1989). Models for the analysis of association in multivariate contingency tables. *Journal of the American Statistical Association*, *84*, 1014-1019.

Becker, M.P., & Clogg, C.C. (1989). Analysis of sets of two-way contingency tables using association models. *Journal of the American Statistical Association*, *84*, 142-151.

Bishop, Y.M.M., Fienberg, S.E., & Holland, P.W. (1975). *Discrete multivariate analysis.* Cambridge: MIT Press.

CBS, Netherlands Central Bureau of Statistics. (1982). *School career and origin of pupils in secondary education: Part 2. Cohort 1977: Choice of school type* (in Dutch). The Hague: Staatsuitgeverij.

Choulakian, V. (1988). Exploratory analysis of contingency tables by log-linear formulation and generalizations of correspondence analysis. *Psychometrika*, *53*, 235-250. (Errata on p. 593)

Clogg, C.C. (1982a). Some models for the analysis of association in multiway cross-classifications having ordered categories. *Journal of the American Statistical Association*, *77*, 803-815.

Clogg, C.C. (1982b). Using association models in sociological research: some examples. *American Journal of Sociology*, *88*, 114-134.

D'Ambra, L., & Kiers, H. (1990). Analysis of log-trilinear models for a three-way contingency table using PARAFAC/CANDECOMP. *Atti delle Giornate di Studio*, 101-114.

Gilula, Z., & Haberman, S.J. (1988). The analysis of multivariate contingency tables by restricted canonical and restricted association models. *Journal of the American Statistical Association*, *83*, 760-771.

Goodman, L.A. (1978). *Analyzing qualitative/categorical data: log-linear models and latent-structure analysis* (Reprints of several Goodman articles). Cambridge, MA: Abt.

Goodman, L.A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association*, *74*, 537-552.

Goodman, L.A. (1981). Association models and canonical correlation models in the analysis of cross-classifications having ordered categories. *Journal of the American Statistical Association*, *76*, 320-334.

Goodman, L.A. (1985). The analysis of cross-classified data having ordered and/or unordered categories: Association models, correlation models, and asymmetry models for contingency tables with or without missing entries. *Annals of Statistics*, *13*, 10-69.

Goodman, L.A. (1986). Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables. *International Statistical Review*, *54*, 243-309.

Greenacre, M.J. (1984). *Theory and applications of correspondence analysis*. New York: Academic Press.

Haberman, S.J. (1978). *Analysis of qualitative data: Vol. 1. Introductory topics*. New York: Academic Press.

Haberman, S.J. (1979). *Analysis of qualitative Data: Vol. 2. New Developments*. New York: Academic Press.

Harshman, R.A., & Lundy, M.E. (1984a). The PARAFAC model for three-way factor analysis and multidimensional scaling. In H.G. Law, C.W. Snyder, J.A. Hattie, & R.P. McDonald (Eds.), *Research methods for multimode data analysis* (pp. 122-215). New-York: Praeger.

Harshman, R.A., & Lundy, M.E. (1984b). Data preprocessing and the extended PARAFAC model. In H.G. Law, C.W. Snyder, J.A. Hattie, & R.P. McDonald (Eds.), *Research methods for multimode data analysis* (pp. 216-284). New-York: Praeger.

Kroonenberg, P.M. (1983). *Three-mode principal component analysis*. Leiden: DSWO Press.

Kruskal, J.B., Harshman, R.A., & Lundy, M.E. (1989). How 3-MFA data can cause degenerate PARAFAC solutions, among other relationships. In R. Coppi & S. Bolasco, *Multiway data analysis*.

Meester, A., & De Leeuw, J. (1983). *Intelligence, social economic status and the school career* (in Dutch). Leiden: Department of Data Theory.

Mooijaart, A. (1991). Three-mode association models. *Proceedings of the International Workshop on Multidimensional Data Analysis, Meeting of Dutch and Italian Schools, October 2-5, Anacapri, Italy.*

Siciliano, R., Lauro, N.C., & Mooijaart, A. (1990). Exploratory approach and maximum likelihood estimation of models for non-symmetrical analysis of two-way multiple contingency tables. *Compstat '90*, 157-162.

Tucker, L.R. (1964). The extension of factor analysis to three-dimensional matrices. In H. Gulliksen & N. Frederiksen (Eds.), *Contributions to mathematical psychology* (pp. 110-119). New York: Holt, Rinehart and Winston.

Tucker, L.R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika*, *31*, 279-311.