Ab Mooijaart
Department of Psychology
University of Leiden
Hooigracht 15
2312 KM Leiden
The Netherlands

Title: Latent Class Analysis (LCA) with order restrictions on the latent parameters

1. Introduction

The Latent Structure Analysis (LSA) model is a general framework for models in which underlying variables are assumed. One very unrestrictive LSA model is the Latent Class Analysis (LCA) model. In this model the underlying (latent) variable(s) are categorical and no order restrictions (or even stronger restrictions) are imposed on the latent variable(s). In fact, the LCA model is the most basic LSA model. By specifying additional assumptions (e.g. order relations between the latent parameters, or by specifying a metric for the latent variable(s)) one can define other LSA models.

Some references regarding LSA models are: Lazarsfeld (1950), MCHugh (1956), Anderson (1959), Gibson (1959, 1960), Madansky (1960), McDonald (1967), Lazarsfeld & Henry (1968), Lord & Novick (with contributions by Birnbaum, 1968), Goodman (1974a, 1974b), Mooijaart (1978, 1980). Other references about LSA models in which more restrictive models are discussed are: Rasch (1960, 1966), Bock & Lieberman (1970), Andersen (1973, 1980), Christoffersson (1975, 1977), Muthén (1978).

Here we discuss LCA models and among them models in which we impose restrictions on the latent parameters. The following kinds of restrictions may be imposed.
a. Parameters may be fixed to a certain value, b. parameters may be equal to other parameters and c. order relations can be defined for the parameters. According my opinion LCA in which the user can specify order restrictions on the latent parameters is new in the literature. One nice consequence is that the problem of how to deal

with estimates of parameters which fall outside a permissible range (e.g. 0 - 1) can be solved. In many estimation procedures of LSA models this is a problem.
In this paper we start with formulating the LCA model. Then we discuss briefly how the latent parameters can be estimated and how to handle order restrictions on the latent parameters. As an illustration of the theory we discuss an example in which LCA with order restricitons was applied.

2. Notation

Scores on manifest (observed) variables are denoted by Latin letters and scores on latent (unobserved) variables by Greek letters. Vectors and matrices are written with a curl underneath to distinguish them from scalars. Frequently, so-called indicator vectors are used. An indicator vector represents a score on a categorical variable. The number of elements of an indicator vector is equal to the number categories of that variable. This vector consists of zero elements only, except for one element which is equal to 1. The position of this unit element corresponds to the category sample element (denoted by e) belongs.

Let $\underset{\sim}{x}_i$ be an indicator vector of order $(r_i \times 1)$. Then $E(\underset{\sim}{x}_i)$ is a vector of expectations which elements can be interpreted as probabilities. Element a of this vector will be denoted as $\pi_{i(a)}$. $E(\underset{\sim}{x}_i\underset{\sim}{x}_j')$ defines a two-dimensional cross-table of order $(r_i \times r_j)$. An element of this table is $\pi_{ij(ab)}$. In an analogous way more-dimensional tables can be defined.

In general, let V denote a set of variables and C an ordered set of specific categories of these variables. Then $E(x_{V(C)})$ will be denoted as $\pi_{V(C)}$. This defines the expectation of falling in the categories given in set C, corresponding to the variables given in set V. (For instance, in the two-dimensional example above: V = (i, j) and C = (a, b).)

When V consists of all available variables, all possible elements $\pi_{V(C)}$ (all elements of the highest dimensional table) may be collected in a vector, called $\underset{\sim}{\pi}$. However, in our estimation procedure we will mostly fit expectations of cross-products up to a certain order.

These expectations (which we call lower order expectations or lower order cross-products) can also be collected in a vector. This vector will be denoted as $\underset{\sim}{\pi}^{o}$, where the superscript refers to the highest order of the fitted cross-products. It is easy to check that each element of $\underset{\sim}{\pi}^{o}$ can be derived from a summation of elements of $\underset{\sim}{\pi}$. For instance, an element $\pi_{S(A)}$ of $\underset{\sim}{\pi}^{o}$ can be derived form $\underset{\sim}{\pi}$ by $\pi_{S(A)} = \underset{\sim}{e}'_{S(A)}\underset{\sim}{\pi}$ where $\underset{\sim}{e}'_{S(A)}$ is a vector of zeroes and ones, defined in such a way that marginal elements of lower dimensional tables are given from elements of the highest dimensional table. Corresponding to each manifest variable there is a matrix $\underset{\sim}{\Lambda}_i$. The order of these matrices is defined by the number of categories of the variables and by a specific latent structure. Column t of this matrix is denoted by the vector $\underset{\sim}{\lambda}_{it}$ and an element of this matrix by $\lambda_{i(a)t}$ (this element corresponds to category a of variable i). The $\underset{\sim}{\Lambda}_i$ matrices are collected in a super-matrix

$\underset{\sim}{\Lambda}' = (\underset{\sim}{\Lambda}'_1 \vert \underset{\sim}{\Lambda}'_2 \vert \ldots)$.

3. Formulations of LSA

In this section we start with formulating the latent class class analysis model. As we shall show this is a very general (not very restrictive) model. From this model other models can be derived by adding restrictions on the latent parameters. Let e be a sample element, $\underset{\sim}{x}_i^e$ an indicator vector representing the answer of element e on variable i. Let $\underset{\sim}{\xi}$ be a latent indicator vector. An element of $\underset{\sim}{\xi}$ corresponds to a category of the latent variable. In fact we can define several latent variables, but for simplicity we shall use only one latent variable. The categories of this variable are the latent classes. Sample elements fall in one and only one latent class. Let T be a latent class. In the latent class model

$$(1) \quad \underset{\sim}{x}_i^e \vert e\varepsilon T = E(\underset{\sim}{x}_i \vert e\varepsilon T) + \underset{\sim}{\delta}_{it}^e .$$

$\underset{\sim}{x}_i^e \vert e\varepsilon T$ denotes that element e's score depends on the latent class it belongs to. $E(\underset{\sim}{x}_i \vert e\varepsilon T)$ is the expectation of $\underset{\sim}{x}_i$ given all elements falling in class T. In general, $E(\underset{\sim}{x}_i \vert e\varepsilon T)$ has some values between 0 and 1, whereas $\underset{\sim}{x}_i^e$ has elements 0 and 1, only. A consequence is that $\underset{\sim}{\delta}_{it}^{e}$ has values between 0 and 1. It is worth mentioning that $E(\underset{\sim}{x}_i|e\varepsilon T)$ is not a "true score" which might be observed. A corollary from (1) is

(2) $E(\underset{\sim}{\delta}_{it}|e\varepsilon T) = \underset{\sim}{0}.$

(This is easy to check by taking expectations over the elements of class T in the model equation.)

Let $\underset{\sim}{\lambda}_{it} = E(x_i|e\varepsilon T)$, so that

(3) $\underset{\sim}{x}_i^e|e\varepsilon T = \underset{\sim}{\lambda}_{it} + \underset{\sim}{\delta}_{it}^e$.

In matrix form the model equation can also be written as:

(4) $\underset{\sim}{x}_i^e|e\varepsilon T = \underset{\sim}{\Lambda}_i\underset{\sim}{\xi}_t + \underset{\sim}{\delta}_{it}^e,$

where $\underset{\sim}{\xi}_t$ is an indicator vector with element t equal to 1. Now we can write:

(5) $E(\underset{\sim}{x}_i|e\varepsilon T) = \underset{\sim}{\Lambda}_i\underset{\sim}{\xi}_t.$

The cross-products of two variables conditional on latent class T is

(6) $\underset{\sim}{x}_i^e\,\underset{\sim}{x}_j^{e'}|e\varepsilon T = \underset{\sim}{\Lambda}_i\underset{\sim}{\xi}_t\underset{\sim}{\xi}_t'\underset{\sim}{\Lambda}_j' + \underset{\sim}{\Lambda}_i\underset{\sim}{\xi}_t\underset{\sim}{\delta}_{it}^{e'} + \underset{\sim}{\delta}_{it}^e\underset{\sim}{\xi}_t'\underset{\sim}{\Lambda}_j' + \underset{\sim}{\delta}_{it}^e\underset{\sim}{\delta}_{jt}^{e'}.$

The expectations of the cross-products is

(7) $E(\underset{\sim}{x}_i\underset{\sim}{x}_j'|e\varepsilon T) = \underset{\sim}{\Lambda}_i\underset{\sim}{\xi}_t'\underset{\sim}{\xi}_t\underset{\sim}{\Lambda}_j' + E(\underset{\sim}{\delta}_{it}\underset{\sim}{\delta}_{jt}'|e\varepsilon T).$

In derivation of (7) we used the property given in (2). We now assume that the δ variables (the residuals) are statistically independent, given class T. (This is analogous to the assumption of uncorrelated errors made in the factor analysis model for continuous variables.) This assumption implies that for $i \neq j$,

(8) $E(\underset{\sim}{\delta}_{it}\underset{\sim}{\delta}_{jt}'|e\varepsilon T) = E(\underset{\sim}{\delta}_{it}|e\varepsilon T)E(\underset{\sim}{\delta}_{jt}'|e\varepsilon T)$

$= \underset{\sim}{0}.$

This is in fact the so-called *local statistical independence* assumption, which is the basis assumption for LSA models.

It is worth mentioning that e.g. factor analysis for continuous variables is also a LSA model. In this case the factors are also defined by the local independence assumption, e.g. see the formulation of factor analysis by Jõreskog and Sõrbom (1979).

In fact factor analysis is an incomplete LSA model, because it uses the crossproducts of first and second order, only.

From (7) the expectation of the cross-products given class T is

$$(9)\quad E(\underset{\sim}{x}_i \underset{\sim}{x}_j' \mid e\varepsilon T) = \underset{\sim}{\Lambda}_i \underset{\sim}{\xi}_t \underset{\sim}{\xi}_t' \underset{\sim}{\Lambda}_j' .$$

In an analogous way, expectations of cross-products of more than two variables can be given. However, it is not convenient to formulate these expectations in terms of matrices. According to (5) and (9), and generalizing to higher order cross-products, we find

$$E(x_{i(a)} \mid e\varepsilon T) = \lambda_{i(a)t}$$

$$E(x_{i(a)}x_{j(b)} \mid e\varepsilon T) = \lambda_{i(a)t}\lambda_{j(b)t}$$

$$(10)\quad E(x_{i(a)}x_{j(b)}x_{k(c)} \mid e\varepsilon T) = \lambda_{i(a)t}\lambda_{j(b)t}\lambda_{k(c)t}$$

$$\vdots \qquad \vdots$$

So far we have formulated expectations of cross-products *conditional* on class T only. However, to specify the distribution of the manifest parameters we have to formulate the unconditional expectations of the cross-products. Let $\underset{\sim}{\omega} = E\underset{\sim}{\xi}$ be a vector of latent class sizes. From (10)

$$\pi_{i(a)} = Ex_{i(a)} = \sum_{t=1}^{s} \omega_t \lambda_{i(a)t}$$

$$\pi_{ij(ab)} = Ex_{i(a)}x_{j(b)} = \sum_{t=1}^{s} \omega_t \lambda_{i(a)t}\lambda_{j(b)t}$$

$$(11)\quad \pi_{ijk(abc)} = Ex_{i(a)}x_{j(b)}x_{k(c)} = \sum_{t=1}^{s} \omega_t \lambda_{i(a)t}\lambda_{j(b)t}\lambda_{k(c)t}$$

$$\vdots \qquad \vdots \qquad \vdots$$

Generally,

$$(12)\quad \pi_{V(C)} = \sum_{t=1}^{s} \omega_t \Pi \lambda_{v_i(c_i)t}$$

(11) and (12) define the so-called accounting equations. From these equations the latent parameters have to be estimated. We estimate the left hand side of (11) and (12) consistently by the proportions in the sample.

4. Estimation of (latent) Parameters

So far we discussed the expectation of cross-products as functions of the latent parameters. These latent parameters can be collected in a vector $\underset{\sim}{\lambda}$.

We may specify $\underset{\sim}{\lambda}$ itself as a function of other parameters, say $\underset{\sim}{\theta}$. The unrestricted model specifies $\underset{\sim}{\pi} = \underset{\sim}{\pi}(\underset{\sim}{\lambda})$, whereas the restricted models are specified by $\underset{\sim}{\lambda} = \underset{\sim}{\lambda}(\underset{\sim}{\theta})$. A general method for estimating $\underset{\sim}{\theta}$ will be discussed here. In this method expectations of cross-products are estimated consistently by corresponding cross-products in the sample. For a discussion of the identification conditions of the parameters in the models we discuss in this paper, see Mooijaart (1980).
Estimation of parameters in models with order restrictions can be done by the method "Optimization by Manifold Suboptimization" (see Zangwill, 1969). However, we do not discuss this method here. (See Mooijaart, 1980, forthcoming.)

The general loss-function can be written as

$$(13)\quad S = (\underset{\sim}{p} - \underset{\sim}{\pi}(\underset{\sim}{\theta}))'W(\underset{\sim}{p} - \underset{\sim}{\pi}(\underset{\sim}{\theta})).$$

which defines the least squares loss function if $\underset{\sim}{W}$ is the identity matrix and if $\underset{\sim}{W}$ is specified as an inverse of the population covariance matrix of $\underset{\sim}{p}$ (say $\underset{\sim}{W} = \underset{\sim}{\Sigma}^{-}(\underset{\sim}{\theta})$) the generalized least squares loss function (see for more details and important theorems e.g. Rao (1965)). In practice it is hardly possible to use all information from the data, because the total number of cross-products (or the total number of response patterns) increases sharply with the number of variables. That is why we use limited information by fitting lower order cross-products only. In that case the loss-function is

$$(14)\quad S^{o} = (\underset{\sim}{p}^{o} - \underset{\sim}{\pi}^{o}(\underset{\sim}{\theta}))'\underset{\sim}{W}^{o}(\underset{\sim}{p}^{o} - \underset{\sim}{\pi}^{o}(\underset{\sim}{\theta})).$$

S^{o} is still asymptotically chi-square distributed if $\underset{\sim}{W}^{o}$ is an inverse of the covariance matrix of $\underset{\sim}{p}^{o}$, because $\underset{\sim}{p}^{o}$ and $\underset{\sim}{\pi}^{o}$ are linear functions of $\underset{\sim}{p}$ and $\underset{\sim}{\pi}$. By using limited information it is therefor still possible to test the model, but in this case with a fewer number of degrees of freedom. How to estimate elements of the matrix W can be found in Mooijaart (1980).

5. Example: Dutch people about abortion

In this example we discuss attitudes towards abortion. Data for a Dutch sample of people have been collected by Veenhoven and Hentenaar (1974). This was a project of the Dutch foundation "Stimezo" (Foundation of medical interruption of pregnancy).
This project was concerned with several issues like attitudes towards Capital Punishment, Abortion, Euthanasia, Sexual Liberty.
For our purpose we use a sample of 543 subjects for which there were no missing values on eight abortion items. For discussion of the results we have to translate these items. The translation of the questions resembles the wording of items of a similar American research project (see Davis et al., 1978) very much.
The wording of the abortion items is:

I think abortion should be possible for a pregnant woman if:

H: the woman's health is seriously endangered by the pregnancy.

O: the woman wants an abortion for some reason and if there are no medical objections.

D: there is a strong chance of serious defect in the baby.

W: the woman is not married and doesn't want to marry the man.

R: the woman became pregnant as a result of rape.

F: the woman already has a large family and cannot afford more children.

A: the woman is not married and is not able to marry the man.

L: it is likely that the child will be unhappy because the parents do not love him.

The underlining of some words above is done by me.
In the following the abortion items will be labeled by the letter given in front of them. These letters correspond with the underlined words.

We shall now discuss some unrestricted and restricted LCA models. In all estimation procedures we fit first, second and third order cross-products. In the tables below only estimators of the parameters of the yes-categories are given, the estimators of the parameters of the no-categories can simply be deduced from them.

In the table below we see the estimators of the unrestricted two class LCA model.

Table I
Unrestricted OLS estimates. Two classes.
Least squares value .235 . Mean deviation .020 .

Item	CLASS 1	CLASS 2
H	.93	1.00
O	.84	.10
D	.99	.66
W	.81	.00
R	1.00	.73
F	.89	.11
A	.87	.03
L	.84	.13
SIZE	.49	.51

Fitting first, second and third order proportions means here fitting 576 proportions. Because the least squares value depends on the number of fitted proportions a reasonable measure for the goodness of fit is S/NP in which S is the least squares value (see (14)) and NP the number of fitted proportions. This measure is called 'Mean deviation'.
In the table below estimators of the unrestricted three class LCA model is given.

Table II

Unrestricted OLS estimates. Three classes.
Least squares value .013 . Mean deviation .005 .

Item	CLASS 1	CLASS 2	CLASS 3
H	.96	.97	.74
O	.90	.20	.02
D	.97	.95	.17
W	.89	.06	.02
R	.99	.96	.35
F	.94	.24	.00
A	.95	.12	.00
L	.89	.25	.01
SIZE	.41	.40	.19

In table II we see that the least squares value (and the mean deviation) is much smaller than in table I. Up to now experience with these measures shows that in table II we have a reasonable good fit for the model. But more important is the following interpretation of this solution.

There are three latent classes of 41, 40 and 19 percent of the subjects. The first class is a class in which people most likely will say yes to all questions. Class two consists of people who are very likely to say yes to the items H-D-R. Class three consists of people who are very likely to

say no to all items, except for item H.
A labeling of the items H-D-R could be "medical" (see also Muthén), "hard" or "physical" and a labeling of the other items could be "social" or "soft" (of course other labels are possible).
Applying these labels the results can be interpreted as follows.
Class 1 consists of people who say yes to abortion for both physical and social reasons; class 2 consists of people who agree with abortion for physical reasons only and class 3 consists of people who do not agree with abortion for any reason (except for item H).
This interpretation suggests an ordering of the latent classes.
And indeed, the estimates show, almost consistently for all items, a specific ordering of the latent parameters over all classes. This could be interpreted as an underlying one-dimensional structure which can be labeled as a liberal versus non-liberal attitude towards abortion. Of course, this continuum is non-linearly related to the manifest variables.
Table III shows the estimates of the parameters of the model in which we imposed order restrictions for the latent parameters.

Table III

Restricted OLS estimates. Three classes.
For each item parameters are ordered from high to low over the classes. Least squares value .014 . Mean deviation .005 .

Item	CLASS 1	CLASS 2	CLASS 3
H	.96	.96	.75
O	.90	.19	.01
D	.97	.93	.12
W	.88	.05	.02
R	.99	.94	.32
F	.94	.23	.00
A	.94	.11	.00
L	.89	.24	.00
SIZE	.41	.41	.17

Table III shows similar results as in table II, so an underlying one-dimensional structure seems reasonable.

We have seen now an ordering of the latent classes. However, is there an ordering for the items too? Inspecting proportions of the items for the whole sample we find an ordering of items: H-R-D-F-L-O-A-W, in which item H has the highest proportion (.92) and item W the lowest proportion (.39). An interesting question is: does the given ordering of items hold not only for the whole sample of people but does it also hold within each latent class? It is possible to formulate this by a LCA model in which order restrictions hold for the parameters within each class. Estimates of the parameters are given in table IV.

Table IV

Restricted OLS estimates.* Three classes.
Latent parameters are both ordered over classes and within each class. Least squares value .025 .
Mean deviation .007 .

Item	CLASS 1	CLASS 2	CLASS 3
H	.97	.97	.75
R	.97	.97	.35
D	.97	.96	.18
F	.94	.23	.02
L	.91	.23	.02
O	.91	.19	.02
A	.91	.14	.00
W	.88	.06	.00
SIZE	.42	.39	.20

* Note: the ordering of the items is not the same as in table I - III.

Substantially the interpretation of table IV is similar to that of table III, so the model seems good. In fact this model specifies double monotonous (holomorf) items. This is analogous to the definition of the Mokken-scale. The difference of our model and the Mokken-scale is that Mokken formulates quantitative underlying variables whereas LCA defines qualitative variables. So the LCA model with double monotonous items is weaker, but essentially both are the same.
In fact, Mokken's main concept is the concept of non-crossing trace-lines.
Of course, other orderings of items within the classes are possible. We chose for the ordering above because then the model specifies that the ordering of the items in the whole sample and in the latent classes are the same. The interpretation of the solution can now be stated as: an underlying continuum reflects the attitude towards abortion.
The extremes of this continuum can be labeled as a "liberal" and a "non-liberal" attitude.

Besides defining specific latent structures it is also meaningful to define so-called discrimination measures for each category. These measures denote how well a category gives information about the underlying structure. For instance, if the latent parameters of a category are equal for all latent classes, then that category does not give any information about the underlying structure. If, on the other hand, the parameters differ very much then the category tells something about the latent structure. The discrimination measure for category "a" of item "i" we use is:

$$D_{i(a)} = \sqrt{\sum_{t=1}^{s} \hat{\omega}_t (\hat{\lambda}_{i(a)t} - \bar{\hat{\lambda}}_{i(a)})^2}$$

in which $\bar{\hat{\lambda}}_{i(a)} = \sum_{t=1}^{s} \hat{\omega}_t \hat{\lambda}_{i(a)t} / s$. $\hat{\omega}_t$ and $\hat{\lambda}_{i(a)t}$ are estimates of latent parameters (see section 3).
In our example with an underlying latent scale these discrimination measures can be interpreted as measures of scalability of items.
In table V the items are ordered with respect to their scalability.
Also in this table results of three different analyses are given.
(For a more detailed discussion of these analyses see Albert Gifi, 1980.)

Table V

Ordering of items with respect to their scalability for 5 different analyses.

	HOMALS	Guttman	Mokken	LCA1	LCA2
most scalable	A	A	W	A	W
	W	W	R	W	A
	F	F	A	F	F
items	O	O	D	O	O
	L	L	F	L	L
	D	R	L	D	D
least scalable	R	D	O	R	R
	H	H	H	H	H

In table V discrimination measures of two LCA models are given: LCA1 is the unrestricted 3-class model and LCA2 is the double monotonous 3-class model. From the table we see that the ordering for the Mokken-scale is quite different from the other ones.
It is striking that LCA2 differs so much from Mokken's analysis because both models assume double monotonous items. (A more detailed comparisón of the two methods have to be carried out.) In all 5 analyses we see that item H is the least scalable item, so this item has almost nothing to do with the underlying structure.
Further in this table it is remarkable that Homals, Guttman's analysis, LCA1 and LCA2 do not differ very much with respect to the ordering of the scalability measures. So a not very risky conclusion may be that these analyses show about the same underlying structure.

Some final remarks

- In this example we have seen that the OLS procedure can be applied very useful for exploring the data.
- The proposed discrimination measure can be very helpful in interpreting a solution.

What we did not discuss in this paper

- How are background variablesrelated to an underlying structure.
- What are the GLS estimators and how to test a model.
- A more-dimensional underlying structure.

These last three points will be discussed in Mooijaart (1980, forthcoming).

References

Andersen, E.B. A goodness of fit test for the Rasch model. Psychometrika, 1973, 38, 123-140.

Andersen, E.B. Comparing latent distributions. Psychometrika, 1980, 45, 121-134.

Anderson, T.W. Some scaling methods and estimation procedures in the latent class model. In U. Grenander (Ed.): Probability and statistics. New York: John Wiley & Sons, Inc., 1959, 9-38.

Bock, R.D. & Lieberman, M. Fitting a response model for n dichotomously scored items. Psychometrika, 1970, 35, 179-197.

Christoffersson, A. Factor analysis of dichotomized variables. Psychometrika, 1975, 40, 5-32.

Christoffersson, A. Two-step weighted least squares factor analysis of dichotomized variables. Psychometrika, 1977, 42, 433-438 (note).

Davis, J.A., Smith, T.W. & Stephenson, C.B. General social surveys, 1972 - 1978: Cumulative codebook. Chicago: National Opinion Research Center, 1978.

de Leeuw, J. Canonical Analysis of Categorical Data. Leiden: University of Leiden, Psychological Institute, 1973.

Gibson, W.A. Three multivariate models: factor analysis, latent structure analysis, and latent profile analysis. Psychometrika, 1959, 24, 229-252.

Gifi, Albert. Niet-lineaire multivariate analyse. Universiteit Leiden, afdeling Datatheorie, Leiden.

Goodman, L.A. The analysis of systems of qualitative variables when some of the variables are unobservable. Part I - A modified latent structure approach. American Journal of Sociology, 1974, 79, 1179-1259.

Goodman, L.A. Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika, 1974, 61, 215-231.

Jöreskog, K.G. A general method for estimating a linear structural equation system. In A.S. Goldberger & O.D. Duncan (Eds.): Structural equation models in the social sciences. New York: Seminar Press, 1973, 85-112.

Jöreskog, K.G. & Sörbom, D. Advances in factor analysis and structural equation models. Cambridge, Mass.: Abt Books, 1979.

Lazarsfeld, P.F. The logical and mathematical foundation of latent structure analysis. In S.A. Stouffer et al. (Eds.): Measurement and prediction. Princeton: Princeton University Press, 1950.

Lazarsfeld, P.F. & Henry, N.W. Latent structure analysis. Boston: Houghton Mifflin Company, 1968.

Lord, F. & Novick, H. (with contributions by A. Birnbaum) Statistical theories of mental test scores. Reading, Massachusetts: Addison-Wesley Publishing Company, 1968.

Madansky, A. Determinantal methods in latent class analysis. Psychometrika, 1960, 25, 183-198.

McDonald, R.P. Nonlinear factor analysis. Psychometric monograph, 1967, 15.

McHugh, R.B. Efficient estimation and local identification in latent class analysis. Psychometrika, 1956, 21, 331-347.

Mokken, R.J. A theory and procedure of scale analysis with applications in political research. Mouton, The Hague, 1970.

Mooijaart, A. Latent Structure Models (doctoral dissertation), Leiden: University of Leiden, Psychological Institute, 1978.

Mooijaart, A. Latent structure analysis for categorical variables. Department of Statistics, University of Uppsala, 1980, Research Report 80-4.

Mooijaart, A. Background and response variables in latent structure analysis. Department of Psychology, University of Leiden, 1980, Research Bulletin (in preparation).

Muthén, B. Contributions to factor analysis of dichotomous variables. Psychometrika, 1978, 43, 551-560.

Rao, C.R. Linear statistical inference and its application. New York: John Wiley & Sons, Inc., 1973.

Rasch, G. Probabilistic models for some intelligence and attainment test. Copenhagen: The Danish Institute for Educational Research.

Rasch, G. An item analysis which takes individual differences into account. British Journal of Mathematical and Statistical Psychology, 1966, 19, 49-57.

Veenhoven, R. and Hentenaar, F. Nederlanders over abortus. Stimezo-onderzoek, no 3, 1975.

Zangwill, W.I. Nonlinear Programming, a unified approach. Prentice Hall, Inc., London, 1969.