KM 42 (1993)
pg.128-142

Ordinal Latent Class Analysis for Single-peaked Items

Marcel Croon *

Abstract: In this paper we propose and develop a latent class model for the analysis of dichotomous items for which the functional relationship between the probability of a postive response and the subject's position on the latent continuum is not monotone, but single-peaked. Our model starts from the assumption that the set of latent classes is totally ordered, and imposes a partial 'umbrella' order relation on each item's response probabilities. By adapting an isotone regression procedure for the weighted least squares estimation of parameters subject to the inequality constraints implied by an 'umbrella' ordering, and by incorporating this procedure in an EM-algorithm, we succeed in estimating the item response probabilities under the constraint of single-peakedness. By way of illustration we discussion the application of our latent class model to two data sets. Finally, we relate the present model to some of our earlier work in the same field.

Key terms: Latent class analysis, ordered latent classes, nonmonotone items, single-peakedness, parallelogram analysis, 'umbrella' ordering.

*Send requests for reprints to: Marcel Croon, Department of Methodology, Faculty of Social Sciences, Tilburg University, P.O. Box 90153, 5000 LE Tilburg (The Netherlands)

1 Introduction

In his *Theory of Data* (1964), Coombs drew a fundamental distinction between data in which a subject's positive response to a particular item should be interpreted in terms of a dominance relation, and data in which such a response should be interpreted in terms of a proximity relation. For the former type of data, a positive response indicates that the subject has more of an underlying capacity or attitude than is strictly needed to respond positively to the item; for the latter kind of data, on the other hand, a positive response is an indication of the fact that the subject's position on the latent continuum is close to the position which corresponds with the item's content. The difference between these two types of data may be best illustrated by means of a one-dimensional geometric model in which both subjects and items are mapped onto points on a coordinate axis. Let β_a and δ_i represent respectively the subject and item coordinate. Then, for dominance data, we have that person a will respond positively to item i if and only if $\beta_a > \delta_i$. For proximity data, on the other hand, the subject's reaction is assumed to depend on the distance between the item's position and his own. In this case it is assumed that subject a will respond positively to item i if and only if $|\beta_a - \delta_i| < \tau$, for some critical threshold τ.

On the basis of these deterministic geometric models, different 'scaling' procedures have been developed for both types of data. These procedures should be thought of in the first place as methods for testing whether the basic model assumptions are appropriate for the data on hand. If that turns out to be the case, these scaling techniques allow the subjects and items to be scaled. For dominance items, Guttman (1944) developed scalogram analysis, which became a much used scaling procedure for dichotomous monotone items; Coombs' parallelogram analysis for the analysis of dichotomous proximity items, on the contrary, was used to a much lesser extent. However, it seems fair to say that the widespread use of both scaling procedures was thwarted by their intrinsic deterministic character. Both scalogram and parallelogram analysis often gave equivocal results when applied to data that contain some non-admissable response patterns. Scalogram analysis nor parallelogram analysis seemed to be able to accomodate in a satisfactory way to random deviations and fluctuations in the data. As a remedy for the apparent shortcomings of deterministic models, appropriate probabilistic scaling models were developed. These probabilistic models should provide more realistic representations of the response processes which determine a subject's reactions. Moreover, they should serve as a basis on which efficient estimation procedures for subject and item parameters can be developed and by means of which statistical hypotheses about these parameters can be tested.

The basic element in probabilistic item response models is the item trace line which describes how the probability of a positive response to a (dichotomous) items varies as a function of the subject's position on the latent continuum. In its most general formulation, we may represent the traceline of item i by $p_i(\beta)$ in which β is the subject's position on the latent continuum. Before specifying the functional form of $p_i(\beta)$ in more detail, some more qualitative observations may be made.

For dominance items it seems reasonable to suppose that $p_i(\beta)$ is a monotonically non-decreasing function of its argument:

$$\beta_1 \leq \beta_2 \Rightarrow p_i(\beta_1) \leq p_i(\beta_2) .$$

Items of this type are called *monotone* items: the higher a subject's position on the latent continuum, the higher his probability for a positive response. Monotone items are very ubiquitous in social and behavioral research. Almost every attainment test, in which the maximally attainable performance of the subjects is being assessed, consists of monotone items. Also many attitude questionnaires are composed of this type of items. A well known example of an item response model for dichotomous items is the Lord-Birnbaum (see Lord and Novick, 1968) model for which we have

$$p_i(\beta) = \frac{1}{1 + e^{\alpha_i(\beta - \delta_i)}}$$

in which the item discrimination parameter α_i and the item difficulty parameter δ_i are item specific parameters. Probably even better known is the model proposed by Rasch (1960), and which may be derived from the Lord-Birnbaum model by assuming that all items have equal discrimination parameters.

For proximity data, on the other hand, the tracelines $p_i(\beta)$ cannot be monotonically increasing function of β. If, for this type of items, the probability of a positive response is a function of the distance between the subject and the item on the latent continuum, one may expect $p_i(\beta)$ to increase as a function of β up to some point on the latent continuum for which $p_i(\beta)$ reaches its maximum value, and to decrease as a function of β thereafter. In other words: for proximity items $p_i(\beta)$ should be a unimodal or single-peaked function of β. The more formal translation of this requirement is *quasi-concavity*. The traceline $p_i(\beta)$ satisfies quasi-concavity if:

$$\beta_1 \leq \beta_2 \leq \beta_3 \Rightarrow p_i(\beta_2) \geq \min\left[p_i(\beta_1), p_i(\beta_3)\right] .$$

Quite recently, Andrich (1988) and Hoijtink (1990) have developed parametric item response models for dichotomous non-monotone items in which the probability of a positive response is a function of the distance between the subject and item position. In the model proposed by Andrich, the item response function is given by

$$p_i(\beta) = \frac{1}{1 + e^{(\beta - \delta_i)^2}} .$$

In Hoijtink's model, the item response function is given by

$$p_i(\beta) = \frac{1}{1 + [(\beta - \delta_i)^2]^\gamma} .$$

In both models, δ_i is the item location parameter: it represents the value on the latent continuum for which the probability of a positive response reaches its maximum value. The item- independent power parameter γ in Hoijtink's model determines the steepness of the item response function. It is clear that both models lead to symmetric single-peaked (strongly quasi- concave) item response functions for the non-monotone items. However, both models have some clear drawbacks. Due to the fact that none of the two models belongs to the so- called family of exponential models, estimation of the unknown item and subject parameters proves to be very difficult from a numerical point of view. Andrich's procedure, which uses joint maximum likelihood estimation of subject and item procedure, may result in biased and even inconsistent estimates; Hoijtink's procedure, which uses marginal maximum likelihood estimation, has to cope with difficulties originating from the fact that the maximum likelihood function which is maximized has many discontinuity points. Moreover, in Andrich's model the maximal value of any response probability, which is attained for $\beta = \delta_i$, is one half; on which theoretical grounds this result might be defended is unclear.

An alternative approach to the development of statistically sound techniques for the analysis of non-monotone items was proposed by Formann (1988), who adapted the classical latent class model (Lazarsfeld and Henry, 1968) in a particular way. By assuming that he knows in advance the ordering of the items along the unidimensional continuum, and also knows the way in which the positive response regions of two or more adjacent items overlap on this continuum, Formann is able to postulate a specific number of latent classes, each of them corresponding to one 'perfect' response pattern. By formulating different assumptions about the way in which random deviations from these perfect patterns may occur, Formann succeeds in defining and implementing various latent class models whose main differences reside in the generality or specificity of 'error' or 'correct' response probabilities. More recently, Böckenholt and Böckenholt (1991) described another version of a latent class model for non-monotone models, in which the item response probabilities from different latent classes are assumed to satisfy a (possibly multi-dimensional) unfolding model which closely resembles Andrich's model.

In the present paper we will discuss another adaptation of the traditional latent class model for the analysis of non-monotone items. The approach proposed here is related to our earlier efforts to develop ordinal latent class models, *i.e.* latent class models in which the latent classes can be thought of as being ordered. For the development of ordinal latent class models for the analysis of monotone items, we refer to Croon (1990, 1991a).

2 The ordinal latent class model for quasi-concave items

2.1 The traditional latent class model: some notation and terminology

Assume n dichotomous items are given; an arbitrary item will be denoted by j. The total number of response patterns is then equal to 2^n. A particular response pattern will be denoted by ν, its observed frequency in a sample of N respondents by f_ν and its probability under some theoretical model by p_ν. Response patterns can also be represented by a binary indicator variable $x_{\nu j}$: $x_{\nu j} = 1$ if in response pattern ν the response to item j is positive; $x_{\nu j} = 0$ otherwise.

Assume furthermore that the data are analyzed by means of a latent class analysis with T latent classes. An arbitrary latent class will be denoted by t. The conditional probability that a subject from latent class t will respond positively to item j will be denoted by $p_{j|t}$. The proportion of subjects belonging to latent class t, its latent proportion, will be denoted by π_t. Now, assuming local independence within each latent class, the probability $p_{\nu|t}$ of observing response pattern ν in latent class t is given by

$$p_{\nu|t} = \prod_j p_{j|t}^{x_{\nu j}} (1 - p_{j|t})^{1-x_{\nu j}} .$$

We also have

$$p_\nu = \sum_t p_{\nu|t} \pi_t .$$

In the classical unconstrained latent class analysis for dichotomous items, the $n \times T$ item response probabilities and the T latent proportions are the unknown model parameters which have to be estimated on the basis of the data. In this respect, one usually resorts to a maximum likelihood estimation procedure which maximize the log likelihood function

$$\sum_\nu f_\nu \ln p_\nu$$

with respect to the unknown parameters. Most often, this estimation problems is solved by implementing an EM-algorithm (Dempster, Laird and Rubin, 1977). Each iteration of an EM- algorithm consists of two steps: an Expectation-step and a Maximization step:

- During the E-step the observed frequencies f_ν of the response patterns are redistributed over the T latent classes in the following way:

  $$e_{\nu t} = f_\nu \times p_{t|\nu} ,$$

  in which

  $$p_{t|\nu} = \frac{p_{\nu|t}\pi_t}{p_\nu}$$

  is the conditional probability that a respondent given his response pattern ν belongs to latent class t. This conditional probability is a function of the provisory estimates of the model parameters.

- During the M-step the maximum likelihood estimates of the model parameters are determined again on the basis of the 'complete' data, which consists of the estimated frequencies $e_{\nu t}$ with which each response patterm ν occurs in each latent class t. In the case of an unconstrained latent class analysis, this estimation procedure is very simple. Let a_{jt} be the frequency of a positive response to item j in latent class t and let e_{+t} be the frequency of latent class t. Note that the values of both quantities can be determined by counting in the appropriate way over the set of redistributed frequencies $e_{\nu t}$:

  $$a_{jt} = \sum_\nu x_{\nu j} e_{\nu t}$$

and

$$e_{+t} = \sum_{\nu} e_{\nu t} \ .$$

Then, the new estimates for the model parameters may be obtained by

$$p_{j|t} = \frac{a_{jt}}{e_{+t}}$$

and

$$\pi_t = \frac{e_{+t}}{N} \ .$$

Note also that it is during this M-step that constraints imposed on the model parameters have to be taken into account. For more detailed information on the implementation of the EM- algorithm in the context of latent class analysis, we refer to Hagenaars (1990) and Croon (1990, 1991a).

2.2 A latent class formulation for single-peaked items

In order to discuss our latent class formulation for single-peaked items we have to assume that the set of latent classes is totally ordered by means of a binary relation $\preceq$. Without loss of information we may assume that the classes are ordered in the following way:

$$1 \preceq 2 \preceq \cdots t \preceq t+1 \cdots \preceq T-1 \preceq T \ .$$

Latent class 1 is in some sense lower, or more to the left, than latent class 2, etc. In general, we say that latent class t is more to the left than latent class $t+1$. For single-peaked items j we now assume that there exists an item-specific latent class s_j such that the item response probabilities for item j satisfy the following system of inequalities:

$$p_{j|1} \leq p_{j|2} \leq \cdots \leq p_{j|s_{j-1}} \leq p_{j|s_j} \geq p_{j|s_{j+1}} \geq \cdots \geq p_{j|T-1} \geq p_{j|T} \ .$$

The response probabilities $p_{j|t}$ are increasing (or better, non-decreasing) with t up till latent class s_j; thereafter, they decrease (or, at least, do not increase) as the latent class number t runs from s_j to T. Note that we do not suppose that the parameter s_j is known in advance: the latent class s_j for wich the response probability for item j reaches its maximal value is itself an unknown parameter. Imposing this kind of 'umbrella' ordering on the response probabilities captures the essential feature of both Andrich's and Hoijtink's model: the single- peakedness of the item response functions.

In order to perform a latent class analysis under the restrictions of an 'umbrella' ordering we have to solve the following optimization problem: Maximize

$$\phi = \sum_{\nu} f_{\nu} \ln \, p_{\nu} \ ,$$

in which

$$p_\nu = \sum_t \pi_t \left(\prod_j p_{j|t}^{x_{\nu j}} (1 - p_{j|t})^{1-x_{\nu j}} \right) ,$$

with respect to the unknown parameters $p_{j|t}$ and π_t, and under the restrictions that for each item j there exists a latent class s_j such that

$$p_{j|1} \leq p_{j|2} \leq \cdots \leq p_{j|s_j-1} \leq p_{j|s_j} \geq p_{j|s_j+1} \geq \cdots \geq p_{|T-1} \geq p_{j|T} \; .$$

Of course, we also require $\sum_t \pi_t = 1$.

2.3 Some optimization problems and their solutions

The optimization problem described in the previous section can be solved by means of an EM algorithm in which, during the M-step of each iteration, the item response probabilities are estimatedunder the 'umbrella' ordering restriction. It is easy to see that the item response probabilities can be estimated separately for each item. In the M-step the provisory estimates $e_{\nu t}$ of the frequency with which response pattern ν occurs in latent class t are available. By counting in an appropriate way, we may determine the frequency a_{jt} of a positive response to item j in latent class t, and the frequency e_{+t} of respondents allocated to latent class t:

$$\begin{aligned} a_{jt} &= \sum_\nu x_{\nu j} e_{\nu t} \\ e_{+t} &= \sum_\nu e_{\nu t} \; . \end{aligned}$$

Using this notation, the log likelihood function which is being maximized in the M-step may be written as

$$\phi_M = \sum_j \sum_t [a_{jt} \ln p_{j|t} + (e_{+t} - a_{jt}) \ln (1 - p_{j|t})] + \sum_t e_{+t} \ln \pi_t$$

in which we impose, of course, that $\sum_t \pi_t = 1$. We immediately derive that the new estimates of the latent proportions are given by

$$\pi_t = \frac{e_{+t}}{N} \; .$$

Moreover, due to the structure of ϕ_M and the fact that the inequality constraints on the item parameters also pertain to separate items, the maximization of the first part of ϕ_M may proceed in an item-wise way. Hence, we are confronted with n successive but separate maximization problems of the following type: Maximize

$$\sum_t [f_t \ln p_t + (n_t - f_t) \ln (1 - p_t)]$$

under the constraint that there exists an integer $s : 1 \leq s \leq T$ for which

$$p_1 \leq p_2 \leq \cdots \leq p_{s-1} \leq p_s \geq p_{s+1} \geq \cdots \geq p_{T-1} \geq p_T \ .$$

Note that the latter system of inequalities defines a quasi-order on the set $\{p_1, \cdots, p_T\}$, since not all pairs of probabilities are comparable to each other. So, for example, although we impose $p_{s-1} \leq p_s$ and $p_{s+1} \leq p_s$, we do not compare directly p_{s-1} and p_{s+1}.

The latter maximization problem can be interpreted as the problem of estimating the success probabilities of T independent binomial distributions under the constraints of the 'umbrella' quasi-ordering. The latter problem may be solved by an appropriate adaptation of an isotonic regression procedure described by Geng and Shi (1990), based on previous theoretical work by Shi (1988).

The isotonic regression problem solved by the Geng and Shi procedure can be stated in the following way. Let $c_t : t = 1, \cdots, T$ be a given set of constants and let $w_t : t = 1, \cdots, T$ be a given set of positive weights. Then, determine the values of the unknown parameters $z_t : t = 1, \cdots, T$ which minimize the quadratic function

$$F = \sum_t w_t (z_t - c_t)^2$$

under the constraints of an 'umbrella' quasi-ordering on the z_t's: There should exist an integer s such that

$$z_1 \leq z_2 \leq \cdots \leq z_{s-1} \leq z_s \geq z_{s+1} \geq \cdots \geq z_{T-1} \geq z_T \ .$$

Their isotonic regression procedure for solving this minimization problem has the following structure:

1. For each integer value $u : 1 \leq u \leq T$, determine the isotonic regression estimates of the z_t's under the constraints

   $$z_1 \leq z_2 \leq \cdots \leq z_{u-1} \leq z_u \geq z_{u+1} \geq \cdots \geq z_{T-1} \geq z_T \ .$$

   To obtain these estimates, the following steps are needed:

   (a) Use the up-and-down block algorithm (see Barlow et al., 1972) on $\{c_1, \cdots, c_{u-1}\}$ with upward trend and on $\{c_{u+1}, \cdots, c_T\}$ with downward trend, and,

   (b) apply the maximum violator algorithm of Barlow et al. (1972) until the block with the largest weighted average contains the peak u.

   The preceding computations result in estimates of the parameters z_t under the assumption that the umbrella ordering attains his peak at the value u. The corresponding value of the quadratic loss function will be denoted by F_u.

2. As the optimal estimate of the peak s, take that value of u for which F_u is minimal:

   $$F_s = \min_{1 \leq u \leq T} F_u \ .$$

   The final estimates of the parameters z_t are the isotonic regression estimates which correspond with the optimal value s.

More technical details and a description of the algorithm can be found in the original paper by Geng and Shi (1990).

Our adaptation of the Geng and Shi isotonic regression procedure for solving the problem of estimating the success probabilities of independent binomial distributions under an 'umbrella' ordering is based on Theorem 1.5.2 of Robertson, Wright and Dykstra (1988). This theorem states that, under some regularity conditions, the canonical parameters of T independent univariate random variables belonging to an exponential family of distributions can be estimated under the restrictions of a quasi-order by means of an appropriately defined isotonic regression procedure. As a corollary to this theorem, the same authors point out that the maximum likelihood estimates of the success probabilities of T independent binomial distributions under the constraints of a quasi-ordering can be obtained by solving the corresponding isotonic regression problem in which constants and weights are (in our previously introduced notation) defined as

$$c_t = \frac{f_t}{n_t} ,$$

and

$$w_t = n_t .$$

Hence, it is clear that, with the preceding definition of constants and weights, Geng and Shi's isotonic regression procedure solves our 'umbrella' order constrained ML estimation problem.

2.4 Some Further Remarks

2.4.1 Determining the Optimal Number of Latent Classes

Up till now we have assumed that the correct number T of latent classes in the optimal solution was known in advance. In practice, this will never be the case. When analyzing real data, the value of T is best considered an additional unknown parameter which we have to estimate. In classical unconstrained latent class analysis the null hypothesis that a particular number of latent classes suffices to 'explain' the data can be tested against the alternative hypothesis that the distribution of the response merely follows a general unconstrained multinomial distribution by means of a log likelihood ratio test which is asymptotically chi-square distributed. Hence, the custom of performing successive latent class analyses with increasing numbers of classes until the point is reached for which the log likelihood ratio test fails to be significamt. As the final estimate of T is taken the smallest number of classes for which the test is not significant.

The latter procedure cannot be used for inequality constrained latent class analysis of any kind, since for this type of estimation problem the asymptotic distribution of the log likelihood ratio statistic is not yet known. A useful, but of course not waterproof, strategy we have often used to determine the appropriate number of latent classes in an inequality constrained latent class analysis is the following:

1. First, we perform a series of traditional unconstrained latent class analyses, and as a provisory estimate of T (say T_0) we take the smallest number of classes for which the log likelihood ratio statistic becomes non-significant.

2. We then perform the constrained analysis with T_0 classes, and check how many inequality constraints have to be made active, and how strongly this affects the value of the log likelihood.

   - If the difference between the log likelihood values for the constrained and unconstrained solutions is not too large compared to the number of free parameters lost, we accept the constrained solution with T_0 classes. We consider the constrained solution as acceptable if the ratio of the difference between the log likelihood values for both solutions and the number of free parameters lost is less than 2.5.

   - Otherwise, we repeat the constrained analyses with a systematically increasing number of classes. In some analyses increasing the number of latent classes leads to a considerable reduction of the log likelihood; however, we have also encountered situations in which the log likelihood value seemed to reach an unacceptably high asymptotic value when the number of classes increased. The latter situation probably occurs when the inequality constrained latent class model is inappropriate for the data under consideration.

2.4.2 Scaling the Respondents

In a traditional latent class analysis individual respondents are 'scaled' by assigning them to a particular latent class on the basis of their response pattern (Hagenaars, 1990). In this assignment procedure the posterior latent class probabilities

$$p_{t|\nu} = \frac{p_{\nu|t}\pi_t}{p_\nu}$$

play an essential role. The posterior probability distribution $\{p_{\nu|t} : t = 1, \cdots, T\}$ indicates how likely it is that a respondent with response vector ν belongs to each of the latent classes. By assigning a respondent to latent class t' for which the posterior probability $p_{t'|\nu}$ is maximal, one minimizes the overall misclassification error rate. Note that latent class t' is nothing but the mode of the posterior distribution. From a measurement point of view, assigning respondents to classes in this way represents a nominal measurement of the latent variable which is defined by the set of latent classes.

A similar assignment procedure for scaling individual respondents could be applied n the case of an ordinal latent class analysis on single-peaked items. However, assigning subjects to classes is the way described above does not explicitly take into account the fact that in an ordinal latent class analysis the set of latent classes is constrained to be ordered, implying that the measurement level of the latent variable defined by these classes is no longer merely nominal, but ordinal. It seems reasonable to require that a scaling procedure following any ordinal latent class analysis should result in a measurement of the latent variable on an ordinal level. One way to score subjects on an ordinal latent scale might consist in computing the median value of his posterior latent class distribution. This scaling procedure 'measures' the individual respondent on an appropriate level, given the data and the model used to analyze them. In this paper we will not illustrate this scaling procedure any further.

3 Some Applications

On the basis of the theoretical considerations as described in the previous sections the programm LCAPEAK was developed for the latent class analysis of single-peaked items. In this section we will discuss two applications of LCAPEAK to real data sets.

The first data set is taken from Formann (1988) and consists of the responses of 600 respondents to five dichotomous items that were used to measure the attitude towards nuclear energy. The five items cover a broad range of positive and negative attitudes towards nuclear energy. Item 1 is clearly the item whose content is most in favor of nuclear energy; item 5, on the other hand, is the item least in favor.

The second data set consists of the responses of 600 respondents to ten dichotomous items that were used to measure the attitude towards car use and its potential threats to the environment. Some of the items formulate an opinion strongly in favor of car use with little or no concern for its possibly damaging effects in the environment; others propose several measures to reduce car use and are, in this sense, more concerned with the possible environmental damages caused by unlimited car driving.

3.1 The Formann Data

We applied LCAPEAK to the Formann data with two, three and four latent classes. Table 1 contains the G^2- and χ^2-values that were observed.

TABLE 1
Global Fit Measures for Formann Data

t	G^2	χ^2
2	63.98	78.29
3	32.20	32.46
4	8.51	8.42

As we will note later, the LCAPEAK solutions were all identical to the corresponding unconstrained latent class solutions for which the log likelihood ratio test may be applied to determine the appropriate number of latent classes. The two- and three-class solutions proved to lead to significant G^2-values, but the probability level corresponding to the G^2-value of the four class solution was equal to .385, which is acceptable. Table 2 gives the estimates of the item response probabilities and the latent class proportions for this four class solution.

TABLE 2
Parameter Estimates Four Class Solution

	I	II	III	IV
1	.54	.46	.08	.00
2	.49	.62	.28	.15
3	.77	.97	.73	.00
4	.20	.91	1.00	.14
5	.00	.26	.98	1.00
π_t	.16	.45	.36	.03

From the information in this table we may conclude that the item response probabilities of all five items have a nice single- peaked relationship to latent class number. For item 1 the item response probabilities decrease as a function of latent class number; for item 5 they increase as a function of t. The three other items are single-peaked in the strict sense: items 2 and 3 attain their largest response probability in latent class 2, item 4 in latent class 3.

However, from the point of view of illustrating the possibilities of LCAPEAK, the present example is somewhat disappointing. As noted earlier, none of the inequalities corresponding to the 'umbrella' orderings had to made active in any of the solutions. This means that the unconditional latent class solution with four classes automatically satisfies the conditions imposed by the hypothesis of single-peakedness, which indicates that single-peakedness is strongly present in the Formann data.

The Formann data are nicely fitted by our ordinal latent class model for single-peaked items, with the attitudinal continuum running from 'favorable towards nuclear energy' at the left to 'opposed to nuclear energy' at the right. Moreover, if we look at where on the continuum each items reaches its maximal positive response probability, it seems safe to conclude that the set of five items covers this attitudinal range quite well. Finally, note that a large majority of the respondents (81 %) belongs to one of the two middle classes.

If we compare our own solution with the final solution preferred by Formann (1988) himself, we note that, whereas he needed nine classes (including one class that contains 42 % of the respondents which were deemed to be unscalable, *i.e.* did not belong to any of the pre- established latent classes which correspond to a 'perfect' response pattern) to describe the data by means of a relatively simple response model, we only need four classes. Our response model, which allows a different response parameter for each item in each latent class is indeed somewhat more complex than Formann's, but in our analysis all subjects are scalable.

3.2 The Car Use Data

In its application to the ten car use items, LCAPEAK was run with the number of classes running from two to seven. Table 3 contains the corresponding G^2- and χ^2-values.

TABLE 3
Global Fit Measures Car Use Data

t	G^2	χ^2
2	581.57	2001.99
3	522.85	1682.37
4	485.31	1425.77
5	469.73	1325.91
6	455.32	1074.03
7	446.82	1082.82

The first thing that strikes the eye is the large discrepancy between the G^2- and χ^2-values for each number of classes. These large differences are almost certainly caused by the extremely large number of zero frequency response patterns in the data: not less than 79 % percent (810 out of 1024) of the response patterns were missing in this sample. This highly jeopardizes any appeal to asymptotic test results; this is the reason why we refrained from comparing the constrained and unconstrained G^2-values for this data set.

We decided to report on the solution with six classes. Increasing the number of classes from six to seven resulted in a relatively small reduction of G^2 with the value of χ^2 (which is not being minimized by our algorithm) even increasing. Table 4 contains the item response probabilities and the latent proportions for the six class solution.

TABLE 4
Parameter Estimates Five Class Solution

	I	*II*	*III*	*IV*	*V*	*VI*
1	.13	.84	.84	.85	.87	.97
2	1.00	.97	.91	.86	.52	.03
3	1.00	.34	.17	.17	.01	.00
4	.90	.83	.44	.44	.13	.13
5	.00	.51	.51	.92	.92	1.00
6	.04	.06	.06	.50	.50	1.00
7	.12	.27	.40	.77	.77	1.00
8	.07	.47	.62	.86	.86	.97
9	.63	.27	.27	.27	.03	.03
10	.90	.95	.37	.24	.24	.02
π_t	.04	.20	.17	.14	.38	.09

The six class solution, as shown in the preceding table, has a remarkable feature: For nine out of the ten items the item response probabilities were either non-decreasing (this was the case for items 1, 5, 6, 7, and 8), or non-increasing (for the items 2, 3, 4, and 9). Only for item 10 was a single- peaked relation in the proper sense found: Item 10 attains its maximal response probability in the second latent class, and its pattern of response probabilities is actually only marginally different from that of the other non-increasing items. Moreover, the set of non-decreasing items consists entirely of the items in facor of car use, while the set of non-increasing items, including item 10, express more critical attitudes towards unlimited car use. Result like these suggests that by simply reflecting for example the non-increasing items, *i.e.* converting an originally negative response to these items into a positive one, and *vice versa*, all items may ultimately be transformed to monotonous items and be analyzed conformingly.

As to the interpretation of the solution found, it is clear that the attitudinal continuum runs from 'in favor of restrictions on car use' at the left to 'in favor of free cae use' at the right. We would also like to point out the somewhat bimodal character of the latent distribution with its two modes being at the second and fifth latent class.

4 Some alternative developments

In this paper we have considered an adaptation of the latent class model for the analysis of single- peaked items. In our discussion we have implicitly assumed that the subjects' responses were observed by means of a single stimulus presentation paradigm: We assumed that the different items were presented separately to the subject who had to indicate whether he or she endorsed the item or not. For such data every conceivable binary response pattern may occur in the data.

However, single-peakedness of response tracelines may be relevant under many more different data collection procedures. More particularly, in investigating preferential choice data subjects are often confronted with entire sets of stimuli from which they have to select and/or rank a specified number of most preferred alternatives. As an example, we may take the following item from the recently held large scale cross-national survey 'European Values 90', organized by the 'European Value Systems Study Group' (see Halman, 1991). Subjects were asked to indicate with which of the following geographical groups they identified most strongly, and second most strongly:

1. Locality or town where you live;
2. Region of country where you live;
3. Your country as a whole;
4. Europe;
5. The world as a whole.

Although the subjects' responses to this item can be coded into a response pattern that pertains to five items, the model for analyzing these data should explicitly take into account that not all conceivable response pattern can occur. For example, if we neglect the relative ranking of the two selected items only response patterns with two 1's and three 0's are possible in the present example, and our basic latent class model should take this condition into account.

A first, perhaps obvious way to do just this is conditioning on the set of response patterns which satisfy the constraint. Let Ω be the set of response patterns which contain exactly two 1's and three 0's. Then, define

$$S_\Omega = \sum_{\nu \in \Omega} \prod_j p_{j|t}^{x_{\nu j}} (1 - p_{j|t})^{1 - x_{\nu j}}$$

so that we may define the properly adapted response probabilities in the following way:

$$\begin{aligned} p_{\nu|t} &= \frac{\prod_j p_{j|t}^{x_{\nu j}} (1 - p_{j|t})^{1 - x_{\nu j}}}{S_\Omega} \text{ for } \nu \in \Omega \\ p_{\nu|t} &= 0 \text{ for } \nu \notin \Omega \end{aligned}$$

Although this approach is conceptually attractive, it unfortunately leads to quite difficult numerical problems once we try to obtain the maximum likelihood estimates of the parameters. Because of the fact that we have conditioned on type of response patterns the log likelihood function cannot be maximized in an item-wise way any longer.

A second, probably more easily implemented approach to the development of latent class models for data from 'pick k out of n' or 'rank k out of n' experiments consists in formulating appropriate log linear models for the ranking or selection probabilities. For 'pick 2 out of n' data the basic building block of such a latent class model would be the 'quasi-independent' log-linear model for the probability p_{ij} that stimuli i and j are chosen from the set of alternatives:

$$\ln p_{ij} = m + a_i + a_j$$

in which the parameters like a_i represent the 'values' of the alternatives. If different latent classes or subpopulations of respondents exist with different stimulus scale values we would write

$$\ln p_{ij|t} = m_t + a_{it} + a_{jt}$$

Similar models have been worked out for the analysis of data that consist of partial preference rankings. For more information in this respect we refer to Croon (1989a, 1989b, 1991) and to Croon and Luijkx (1992).

REFERENCES

Andrich, D. (1988). The application of an unfolding model of the PIRT type to the measurement of attitude. *Applied Psychological Measurement*, **12**, 33-51.

Barlow, R.E., Bartholomew,D.J., Brenner, J.M., and Brunk, H.D. (1972). *Statistical*

Inference under Order Restrictions. New York: Wiley.
Böckenholt, U. and, Böckenholt, I. (1991). Constrained latent class analysis: Simultaneous classification and scaling of discrete choice data. *Psychometrika*, **56**, 699-716.
Coombs, C.H. (1964). *A Theory of Data.* New York: Wiley.
Croon, M.A. (1989a). The analysis of partial rankings by log-linear and latent class models. In: R. Coppi and S. Bolasco (eds): *Multiway Data Analysis.* Amsterdam: Elsevier, pp.497-506.
Croon, M.A. (1989b). Latent class models for the analysis of rankings. In: G. De Soete, H. Feger and K.C. Klauer (Eds): *New developments in psychological choice models.* Amsterdam: Elsevier, pp.99-121.
Croon, M.A. (1990). Latent class analysis with ordered classes. *British Journal of Mathematical and Statistical Psychology,***43**, 171-192.
Croon, M.A. (1991a). Investigating Mokken scalability of dichotomous items by means of ordinal latent class analysis. *British Journal of Mathematical and Statistical Psychology*, **44**, 315-331.
Croon, M.A. (1991b). Log-linear analysis of partial preference rankings. *Methodika*, **IV**, 63-86.
Croon, M.A. and, Luijkx, R. (1992, to appear). Latent structure models for ranking data. In: J. Verducci and M. Fligner (Eds): *Probability models and statistical analyses for rankings.* New York: Spinger Verlag (Lecture Notes in Statistics).
Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). Maximum likelihood estimation from incomplete data via the EM-algorithm. *Journal of the Royal Statistical Society, Series B*, **39**, 1-22.
Formann, A.K. (1988). Latent class models for nonmonotone dichotomous items. *Psychome- trika*, **53**, 45-62.
Geng, Z., and, Shi, N.Z. (1990). Isotonic regression for umbrella orderings. *Applied Statistics*, **39**, 397-424.
Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review*, **9**, 139-150.
Hagenaars, J.A. (1990). *Categorical Longitudinal Data. Log-Linear, Panel, Trend, and Cohort Analysis.* Newbury Park: Sage.
Halman, L. (1991). *Waarden in de Westerse Wereld.* [*Values in the Western World*] Dissertation. Tilburg: Tilburg University Press.
Hoijtink, H. (1990). A latent trait model for dichotomous choice data. *Psychometrika*, **55**, 641-656.
Lazarsfeld, P.F. and, Henry, N.W. (1968). *Latent Structure Analysis.* Boston: Houghton Mifflin.
Lord, F.M. and, Novick. M.R. (1968). *Statistical Theories of Mental Tests Scores.* Reading (Mass.): Addison-Wesley. Rasch, G. (1960). *Probabilistic Models for some Intelligence and Attainment Tests.* Copenhagen: Paedagogiske Institut.
Robertson, T., Wright, F.T., and Dykstra, R.L. (1988). *Order Restricted Statistical Inference.* New York: Wiley.
Shi, N.Z. (1988). A test of homogeneity for umbrella alternatives and tables of the level probabilities.*Communications in Statistics. Theory and Methods*, **17**, 657-650.