# Modeling Individual Differences in Unfolding Preference Data: A Restricted Latent Class Approach

Ulf Böckenholt
University of Illinois

Ingo Böckenholt
University of Karlsruhe

A latent class scaling approach is presented for modeling paired comparison and "pick-any/$t$" data obtained in a preference study. Although the latent class part of the model identifies homogeneous subgroups that are characterized by their choice probabilities for a set of alternatives, the scaling part of the model describes the single-peakedness structure of the choice data. Procedures are suggested for examining the unfolding structure in an unrestricted latent class solution. Two applications are presented to illustrate the technique. In the first application, scaling solutions obtained from a latent class scaling model and a marginal maximum likelihood latent trait model are compared. *Index terms: latent class analysis, paired comparison data, pick any/t data, unfolding models.*

This research concerned two important data collection methods for measuring choice behavior. The first technique is a natural and economical one in which respondents are asked to select a preferred object from a set of objects. Coombs (1964) termed this procedure "pick-any" when the set of choice alternatives is unconstrained, and "pick-any/$t$" when the set of choice alternatives is fixed. In contrast to the first technique, the second technique requires people to compare pairs of objects and to select the preferred one. This is a more time-consuming method, but it is also more useful for obtaining reliable information about a person's preference order. For both data collection methods, a general approach is presented to account for individual differences and to scale the choice alternatives along a continuum. This approach is based on a reparameterization of a latent class model; thus, parameter estimation and model tests are straightforward. Moreover, simple tests can be performed to examine whether the choice data conform to an unfolding structure.

The restricted latent class analysis (LCA) used in this study complements a latent trait unfolding model recently introduced by Andrich (1988, 1989) for describing pick-any/$t$ and paired comparison data.. In order to represent individual preferences and the choice objects under study simultaneously, Andrich suggested the joint estimation of both person and object parameters. Such a procedure is complicated by the fact that, with increasing sample sizes, the number of person parameters to be estimated increases and standard limit theorems do not apply (Bock & Aitkin, 1981). A restricted latent class approach does not share these difficulties, but it also does not provide as much detailed information about individual differences. However, a restricted latent class approach may be the method of choice when a researcher is interested in detecting group structures consisting of respondents with similar choice patterns.

Unrestricted LCA is first reviewed for binary choice data. Next, a new approach is presented for modeling individual differences in choice data by combining the latent class (LC) model and several one-dimensional unfolding models for pick-any/$t$ and paired comparison data. After estimation and model tests are discussed, results are presented from two real datasets involving political and consumer judgments. The political judgment dataset was previously analyzed by Hoijtink (1990), who applied a marginal maximum likelihood latent trait approach. Finally, further extensions are discussed.

## Latent Class Analysis

The first systematic treatment of latent class models was given by Lazarsfeld and Henry (1968), and more recent developments are described by several texts and research monographs (Bartholomew, 1987; Formann, 1984; Goodman, 1978; Langeheine & Rost, 1988; McCutcheon, 1988; Rost, 1988b). Because these developments are well documented in these sources, only a summary of the fundamental assumptions of LCA (cf. Formann, 1985) is presented. Let $p_{m|\ell}$ be the probability that a person $s$ in class $\ell(\ell = 1, \ldots, q)$ selects object $m$ from a set of $t$ objects—that is,

$$P(R_{sm} = 1 \,|\, s \in K_\ell) = p_{m|\ell} \quad , \tag{1}$$

where $R_{sm}$ is a random 0,1 variable. The prior probability of a person belonging to latent class $\ell$ is denoted by $p_\ell$, and is usually interpreted as a class size parameter. Because the classes are defined to be exhaustive and disjunctive, the unconditional probability that person $s$ selects object $m$ is

$$P(R_{sm} = 1) = \sum_{\ell = 1}^{q} p_\ell p_{m|\ell} = p_m \quad , \text{ where } \sum_{\ell = 1}^{q} p_\ell = 1 \quad . \tag{2}$$

Person responses are also assumed to be stochastically independent within each latent class. Thus the probability of a person selecting pattern $R_s = \{R_{s1}, \ldots, R_{st}\}$ is

$$P(R_s \,|\, s \in K_\ell) = \prod_{m = 1}^{t} p_{m|\ell}^{R_{sm}} (1 - p_{m|\ell})^{1 - R_{sm}} \quad , \tag{3}$$

and the corresponding unconditional probability of observing a selection pattern $R_s$ is

$$P(R_s) = \sum_{\ell = 1}^{q} p_\ell P(R_s \,|\, s \in K_\ell) \quad . \tag{4}$$

Both sets of parameters, $p_\ell$ and $p_{m|\ell}$, are obtained by maximizing the kernel of the likelihood function

$$L = \prod_{j = 1}^{2^t} \left\{ \sum_{\ell = 1}^{q} p_\ell \left[ \prod_{m = 1}^{t} p_{m|\ell}^{R_{sm}} (1 - p_{m|\ell})^{1 - R_{sm}} \right] \right\}^{n_j} \quad , \tag{5}$$

where $n_j$ denotes the observed frequency for a response pattern $R_j$. The expectation maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977; Goodman, 1978, 1979; Takane & DeLeeuw, 1987) has proven useful for estimating latent class parameters.

## A Synthesis of Unfolding Models and LCA

### Pick-Any/$t$ Data

Approaches suggested for modeling pick-any/$t$ data may be roughly classified as either deterministic or stochastic, and as either parametric or nonparametric. An example of a nonparametric and deterministic approach is a procedure suggested by Cliff, Collins, Zatkin, Gallipeau, and McCormick (1988) for reordering the rows and columns of a rank-order matrix in order to obtain as close a

parallelogram structure as possible. Andrich (1988) presented a parametric and stochastic model that describes explicitly the relationship between the distance of an object to a person's ideal point and the probability of endorsing it. Both approaches appear to be most useful at different steps of a modeling process. For instance, qualitative characteristics of the unfolding structure may be examined first by a nonparametric procedure, and in a further step, a parametric model may be fitted to obtain more precise information from the data.

The restricted latent class framework facilitates this sequential procedure. In the first step, the *unrestricted* latent-class-specific probabilities are examined to determine whether they conform to an unfolding structure. If this is the case, a parametric model is fitted to describe the unfolding relationship in a parsimonious way.

This research focused on two different parametric versions of stochastic unfolding models. These models were selected for their simple structure and their ease of implementation. However, there is little reason to believe that both models describe exhaustively unfolding preference data. For example, both models are symmetric around the ideal point. This characteristic may not always be found in preference data; if that occurs, other functional relationships may be specified that can be implemented easily in a latent class framework.

The first unfolding model was discussed and applied by Andrich (1988), and DeSarbo and Hoffman (1986, 1987). The latter authors specified a multidimensional version of this model. The class-specific probabilities of the LCA framework are defined here as a function of the distance between the ideal point of a latent class $\ell$ and an object coordinate, $x_m$,

Model 1: $p_{m|\ell} = 1/\left\{1 + \exp\left[-\tau_\ell + (i_\ell - x_m)^2\right]\right\}$ , $\qquad\qquad$ (6)

where $\tau_\ell$ is an additive constant for latent class $\ell$. Note that $p_{m|\ell}$ approaches 1 only when $(i_\ell - x_m) = 0$ and $\tau_\ell \to \infty$. Another version of an unfolding model has been recently introduced by Hoijtink (1990). His model can be represented in a latent class formulation as

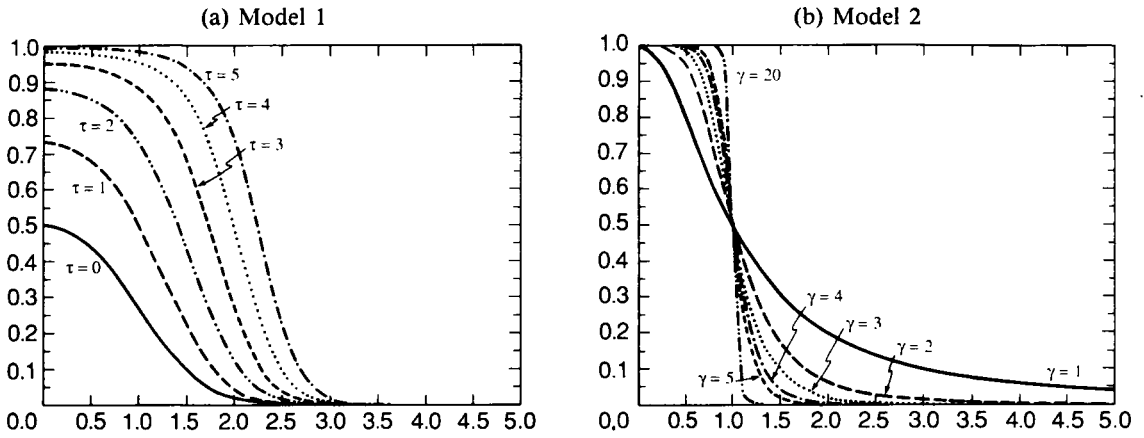Model 2: $p_{m|\ell} = 1/\left[1 + (i_\ell - x_m)^{2\gamma}\right]$ , $\qquad\qquad$ (7)

where $\gamma$ is positive. In contrast to the previous model, this model has class specific probabilities of 1 when $(i_\ell - x_m) = 0$. Thus, an object is always chosen when its position is equal to the ideal point.

Further differences between the models are illustrated in Figure 1, which depicts class-specific probabilities as a function of the distance between the ideal point and the coordinates of the objects. For Model 1 (Figure 1a), six curves are displayed that are the result of setting the additive constant $\tau = 0, 1, 2, 3, 4$, or 5. The curves are monotonically decreasing and do not intersect. Model 2, on the other hand (Figure 1b), is shown for $\gamma = 1, 2, 3, 4, 5$, or 20. For small values of $\gamma$ a slowly decreasing function is obtained, and for large $\gamma$ values the relationship approaches a step function. Not surprisingly, all curves intersect at $(i_\ell - x_m)^2 = 1$ and $p_{m|\ell} = .5$.

Both models have in common the fact that their class-specific probabilities are symmetric around the ideal point, they predict strong stochastic transitivity (SST), and they satisfy the condition of bilateral monotonicity. For example, consider the class-specific probabilities in Table 1 that are presented for different object coordinates and ideal points ordered by their size. The table depicts an unfolding structure with the main (block) diagonal of both matrices containing the largest values, and decreasing probabilities to both sides of the diagonal.

This particular structure exemplifies one necessary condition for fitting the two unfolding Models 1 and 2 in a latent-class framework. Thus, both unfolding models may be of use only if the *unrestricted* latent-class-specific probabilities can be reordered to conform to this pattern. Because the number of latent classes is frequently small, a reordering of the class-specific probabilities provides a straightforward examination of the appropriateness of the unfolding models. Moreover, by imposing

**Figure 1**
Representation of Models as a Function of the Distance to the Ideal Point

(a) Model 1

(b) Model 2

equality constraints on the class-specific probabilities in such a way that they conform to an unfolding structure, unfolding models can be estimated in the form of discontinuous functions. Applications of this approach can be found in Formann (1988).

## Paired Comparison Data

Only few approaches are available for modeling individual differences in paired comparison data. Takane (1987) developed a comprehensive methodology based on an analysis of covariance structures approach that facilitates a separation of the systematic individual differences components from random components. In addition, Takane's framework includes paired comparison models introduced by Bock and Jones (1968), and Bloxom (1972) as special cases. Restricted LCA is suggested as an alternative approach to account for the dependencies among the paired comparison judgments. This

**Table 1**
Class-Specific Probabilities Generated By Models 1 and 2
(Columns Represent Object Coordinates and
Rows Represent Different Ideal Point Positions)

| Model and Position | -3 | -2 | -1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|---|
| Model 1 ($\tau = 3$) | | | | | | | |
| -2.8 | .951 | .914 | .440 | .001 | .000 | .000 | .000 |
| -1.7 | .788 | .948 | .925 | .528 | .014 | .000 | .000 |
| -0.6 | .060 | .739 | .945 | .933 | .608 | .023 | .000 |
| 0.5 | .000 | .037 | .679 | .940 | .940 | .679 | .037 |
| 1.4 | .000 | .000 | .060 | .739 | .945 | .933 | .608 |
| 2.3 | .000 | .000 | .000 | .092 | .787 | .948 | .925 |
| Model 2 ($\gamma = 1$) | | | | | | | |
| -2.8 | .962 | .610 | .236 | .113 | .065 | .042 | .029 |
| -1.7 | .372 | .917 | .671 | .257 | .121 | .068 | .043 |
| -0.6 | .148 | .338 | .862 | .735 | .281 | .129 | .072 |
| 0.5 | .076 | .138 | .308 | .800 | .800 | .308 | .138 |
| 1.4 | .049 | .080 | .148 | .338 | .862 | .735 | .281 |
| 2.3 | .034 | .051 | .084 | .159 | .372 | .917 | .671 |

approach has also been applied by Formann (1989), who discussed an LCA version of the Bradley-Terry-Luce model (Luce, 1959).

In restricted LCA, $\binom{t}{2}$ paired comparisons are modeled. Because the number of possible response patterns increases rapidly with the number of objects, an LCA of paired comparison data is limited to a small number of objects. A complete paired comparison of $t$ objects yields $2^{\binom{t}{2}}$ possible response patterns. In contrast, for pick-any/$t$ data, $2^t$ response patterns may be observed. Consequently, incomplete paired comparison designs (Bock & Jones, 1968) may be of considerable usefulness in a restricted LCA. Andrich (1989) provides a detailed derivation of Model 1 for paired comparison data. Following his approach, Model 1 specifies the probability that object $m$ is preferred to object $h$:

$$\text{Model 3: } p_{mh|t} = 1/\{1 + \exp[x_m^2 - x_h^2 - 2i_t(x_m - x_h)]\}, \quad (x_m < x_h) \quad , \tag{8}$$

and the additive constant is eliminated. Similarly, for Model 2,

$$\text{Model 4: } p_{mh|t} = (i_t - x_h)^{2\gamma}/[(i_t - x_m)^{2\gamma} + (i_t - x_h)^{2\gamma}], \quad (x_m < x_h) \quad . \tag{9}$$

The main difference between both models is that the paired comparison Version 4 of Hoijtink's model makes the strong prediction that for an object value identical to the ideal point position, (i.e., $i_t = x_m$), object $m$ is always preferred to object $h$, regardless of the position of object $h$.

To illustrate some of the qualitative characteristics of Models 3 and 4, the class-specific probabilities are represented in the form of paired comparison matrices in Tables 2 and 3. These matrices were computed for the object values, $x = (-2, -1, 1, 2)$, and the ideal point positions, $i = (-1.7, -.6, .5, 1.4)$. Additionally, the rows and columns of the matrices were reordered so that the paired comparison probabilities increase from the left to the right columns and from the bottom to the top rows.

Greenberg (1965) provided a detailed discussion of this gradient pattern for unfolding paired comparison data. Models 3 and 4 may be of use only if the unrestricted latent-class probabilities can be reordered so that they conform to this pattern. This condition can be easily examined for a small number of objects and latent classes. For a large number of objects, a recursive dynamic programming strategy presented by Hubert and Golledge (1981) can be used to reorganize optimally the rows and columns of the paired comparison matrices for each latent class.

Table 2
Class-Specific Paired Comparison Probabilities Generated By
Model 3 ($x_1 = -2$, $x_2 = -1$, $x_3 = 1$, and $x_4 = 2$)

| $i_t = -1.7$ | | | | | $i_t = -.6$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Object | 1 | 2 | 3 | 4 | Object | 2 | 1 | 3 | 4 |
| 1 | .5 | .599 | .999 | 1.00 | 2 | .5 | .858 | .917 | .999 |
| 2 | .401 | .5 | .999 | 1.00 | 1 | .142 | .5 | .646 | .992 |
| 3 | .001 | .001 | .5 | .998 | 3 | .083 | .354 | .5 | .985 |
| 4 | .000 | .000 | .002 | .5 | 4 | .001 | .008 | .015 | .5 |
| $i_t = .5$ | | | | | $i_t = 1.4$ | | | | |
| Object | 3 | 2 | 4 | 1 | Object | 3 | 4 | 2 | 1 |
| 3 | .5 | .881 | .881 | .998 | 3 | .5 | .550 | .996 | 1.00 |
| 2 | .119 | .5 | .500 | .982 | 4 | .450 | .5 | .996 | 1.00 |
| 4 | .119 | .500 | .5 | .982 | 2 | .004 | .004 | .5 | .997 |
| 1 | .002 | .018 | .018 | .5 | 1 | .000 | .000 | .003 | .5 |

**Table 3**
Class-Specific Paired Comparison Probabilities Generated By
Model 4 ($x_1 = -2$, $x_2 = -1$, $x_3 = 1$, and $x_4 = 2$)

| $i_t = -1.7$ | | | | | $i_t = -.6$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Object | 1 | 2 | 3 | 4 | Object | 2 | 1 | 3 | 4 |
| 1 | .5 | .845 | .988 | .993 | 2 | .5 | .925 | .941 | .977 |
| 2 | .155 | .5 | .937 | .965 | 1 | .075 | .5 | .566 | .775 |
| 3 | .012 | .063 | .5 | .653 | 3 | .059 | .434 | .5 | .725 |
| 4 | .007 | .035 | .347 | .5 | 4 | .023 | .225 | .275 | .5 |
| $i_t = .5$ | | | | | $i_t = 1.4$ | | | | |
| Object | 3 | 2 | 4 | 1 | Object | 3 | 4 | 2 | 1 |
| 3 | .5 | .900 | .900 | .962 | 3 | .5 | .692 | .973 | .986 |
| 2 | .100 | .5 | .500 | .735 | 4 | .308 | .5 | .941 | .970 |
| 4 | .100 | .500 | .5 | .735 | 2 | .027 | .059 | .5 | .667 |
| 1 | .038 | .265 | .265 | .5 | 1 | .014 | .030 | .333 | .5 |

## Parameter Estimation: An EM Algorithm

The parameters of the restricted LC models can be estimated using a Newton-Raphson procedure by maximizing the likelihood function of the LC model under the constraints imposed by unfolding Models 1 through 4. However, it is computationally more convenient to modify the EM algorithm for estimating the unrestricted LC models. This approach has been adopted before by Rost (1985; 1988a), for example. The main steps necessary for implementing this approach are discussed below with the notation for pick-any/$t$ data. The corresponding changes in notation for paired comparison data are straightforward and are omitted.

### E Step

Let $n_j$ denote the observed frequency for response pattern $R_j$, and $p_j$ the probability of response pattern $j$. Let $n_{j,\ell}$ denote the frequency of the response pattern $j$ in latent class $\ell$, and let $p_{j,\ell}$ denote the joint probability of observing the response pattern $j$ in class $\ell$. Under the assumption of local independence, $p_{j,\ell}$ is computed

$$p_{j,\ell} = p_\ell \prod_{m=1}^{t} p_{m|\ell}^{R_{sm}}(1 - p_{m|\ell})^{1-R_{sm}}, \text{ and } p_j = \sum_{\ell=1}^{q} p_{j,\ell} \quad . \tag{10}$$

The expected number of respondents selecting object $m$ in latent class $\ell$ is given by

$$E(N_{m|\ell}) = \sum_{j,(R_m = 1)} n_j(p_{j,\ell}/p_\ell), \text{ for } m = 1, \ldots, t \quad , \tag{11}$$

and the expected number of respondents in class $\ell$ is given by

$$E(N_\ell) = \sum_{j=1}^{2^t} n_j(p_{j,\ell}/p_j) \quad . \tag{12}$$

### M Step

In the M step the class size parameters, $p_\ell$, are estimated by

$$\hat{p}_\ell = n^{-1}E(N_\ell) \quad , \tag{13}$$

where $n$ denotes the total number of respondents. Although the class-specific parameters $p_{m|\ell}$ for the unrestricted LC model are determined by

$$\hat{p}_{m|\ell} = E(N_{m|\ell})/E(N_\ell) \tag{14}$$

in the restricted LC model, they are constrained to be equal to one of the unfolding Models 1 through 4. Thus the class-specific probabilities constrained under Model 2, for example, are determined by maximizing the kernel of the likelihood function of the complete data,

$$L^* = \prod_{m=1}^{t} \prod_{\ell=1}^{q} \left[ 1 + (i_\ell - x_m)^{2\gamma} \right]^{-E(N_{m|\ell})} \left( 1 - \{1/[1 + (i_\ell - x_m)^{2\gamma}]\}^{-1} \right)^{E(\bar{N}_{m|\ell})} , \tag{15}$$

where $E(\bar{N}_{m|\ell})$ denotes the expected number of respondents who do not select object $m$ in latent class $\ell$. Values of $p_\ell$ and $p_{m|\ell}$ obtained in the M step are then used to determine the parameter estimates in the E step. This procedure continues until convergence is obtained.

## Model Tests

Large sample tests of fit of any of the restricted or unrestricted LC models or submodels are available using the likelihood ratio (LR) $\chi^2$ statistic:

$$G^2 = 2 \sum_{j=1}^{2^t} n_j \ln \left( \frac{n_j}{\hat{n}_j} \right) , \tag{16}$$

where $\hat{n}_j$ denotes the expected pattern frequency under the LC model. When one hypothesis is nested within another, differences between the likelihood ratio statistics can be computed to assess the importance of the contribution to the likelihood ratio statistic by the additional constraints imposed by the stronger hypothesis. This difference is asymptotically distributed as a $\chi^2$ statistic with degrees of freedom equal to the difference between the number of parameters in the unrestricted and restricted models.

This approach provides further guidance in selecting a parsimonious model. For the comparison of models that are not proper subsets of each other, the Aikake information criterion (AIC) statistic (Aikake, 1974) is applied. For example, the AIC statistic is useful for comparing the fit of unfolding Models 1 and 2, because these models are not nested. The AIC statistic is defined as

$$\text{AIC} = -2 \ln L_m + 2d , \tag{17}$$

where $L_m$ is the value of the model's maximum likelihood and $d$ corresponds to the number of effective model parameters.

Although $df = 2^t - 1$ under the saturated model, the unrestricted latent class model requires the estimation of $q \times t$ class-specific response probabilities and $q - 1$ class size parameters. The ideal point Models 1 through 4 are based on $t - 1$ scale values, $q$ ideal points, $q - 1$ class size parameters, and either one power parameter (Model 2, 4), or $q$ threshold parameters (Model 1). It is necessary to estimate only $t - 1$ object coordinates because the origin of the scale is arbitrary.

## Application I: Reanalysis of Political Candidate Data

The following example was selected to compare scaling solutions obtained from a marginal maximum likelihood (MML) latent trait model, and the restricted LC approach. Takane (1983a, 1983b) pointed to the usefulness of treating person parameters as random effects when modeling pick-any data. Further support for this approach was obtained by Hoijtink (1990), who developed a latent trait formulation of unfolding Model 2. In an application of his unfolding model, Hoijtink estimated the model parameters (without a priori assumption about the distribution of the person parameters) by MML methods.

Using the restricted latent-class approach, data were reanalyzed that were collected from 697 delegates who attended the Missouri party convention in March 1980; the delegates were asked to indicate their opinion, using a five-point favorableness scale, of nine political candidates (Abramowitz

& Stone, 1984). Six candidates were running for the presidency (Carter, Kennedy, Brown, Reagan, Bush, and Anderson), and three candidates were running for state offices. The names of these candidates are denoted by Governor, Senator 1, and Senator 2. The five-point rating scale with the labels "1 = very favorable" to "5 = very unfavorable" was recoded to a binary scale (1 = 1; 2, 3, 4, 5 = 0).

To fit his one-dimensional latent trait unfolding model, Hoijtink removed three politicians, Senator 1, Senator 2, and Brown, from the analysis. His approach was followed by analyzing the dataset using an unrestricted latent class approach. Both the LR test compared against the saturated model, and the AIC statistics in Table 4, suggest a three-class solution. The ordered latent-class-specific probabilities and the class size estimates are presented in Table 5. The first (24.0%) and second class (16.8%) favor the Democratic candidates, the third class (59.2%) favors the Republican candidates. It is interesting to note that the first two classes disagree considerably on their favorableness ratings for Carter and Kennedy, which could indicate a split in the Democratic party for the two candidates.

**Table 4**
Number of Latent Classes (LC), Results
of the Likelihood Ratio Test (LRT),
Degrees of Freedom (*df*), and Aikake's
Information Criterion (AIC) in
Application I for Unrestricted LCA

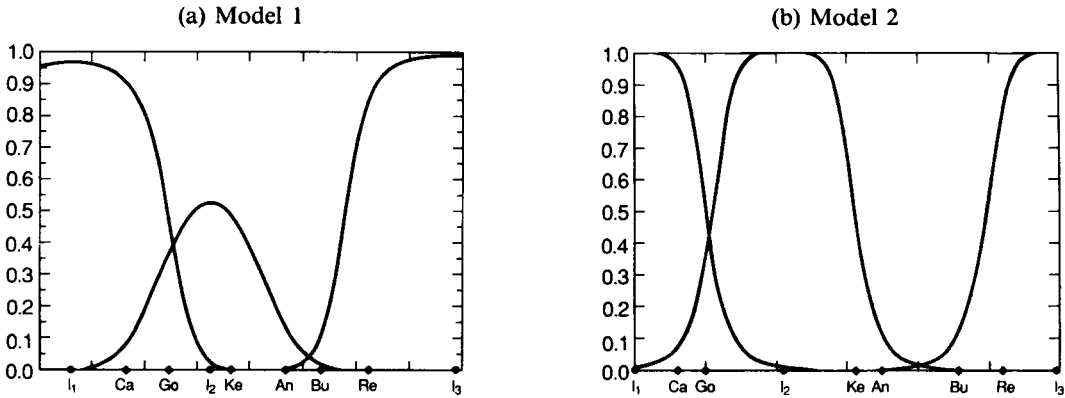| LC | LRT | *df* | AIC |
|----|------|----|----------|
| 1 | 8136.99 | 57 | 10756.33 |
| 2 | 116.44 | 50 | 2761.76 |
| 3 | 25.60 | 43 | 2672.93 |
| 4 | 13.14 | 36 | 2674.47 |
| 5 | 12.02 | 29 | 2687.35 |
| 6 | 9.41 | 22 | 2698.74 |

With one minor exception (Governor in the third class), the class-specific probabilities conform to an unfolding structure. Fitting Models 1 and 2 to the dataset resulted in likelihood ratio tests of $G^2 = 32.00$ ($df = 53$, AIC = 2658.32) for Model 1, and $G^2 = 32.10$ ($df = 55$, AIC = 2655.42) for Model 2. Both models provided a satisfactory fit to the data and were considerably more parsimonious than the unrestricted latent class solution. Model 2 seemed to be somewhat superior, because it is based on fewer parameters. The coordinates of the judged politicians are similar for both models. Figure 2 depicts graphical representations of the parameter estimates obtained from Models 1 and 2. Only one ideal point is enclosed by the object coordinates, which indicates that an unfolding representation is necessary only for the second class. Pearson correlations between the scale values obtained from Models 1 and 2, and Hoijtink's MML solution were .995 and .975, respectively, indicating a

**Table 5**
Class-Specific Selection Probabilities of Three Latent Class
Solutions (Ca = Carter, Ke = Kennedy, Re = Reagan,
Bu = Bush, An = Anderson, Go = Governor)

| Class | Class Size | Class-Specific Probabilities | | | | | |
|-------|-----------|------|-------|------|------|------|------|
| | | Ca | Go | Ke | An | Bu | Re |
| 1 | .240 | .992 | .468 | .022 | .012 | .002 | .002 |
| 2 | .168 | .007 | .394 | .454 | .109 | .002 | .002 |
| 3 | .592 | .002 | .009* | .002 | .005 | .116 | .831 |

*Violations of an unfolding structure.

**Figure 2**
Three-Class Solutions for Objects Ca = Carter, Ke = Kennedy, Re = Reagan, Bu = Bush,
An = Anderson, Go = Governor (Latent Class Positions are Denoted by $\ell_1$, $\ell_2$, and $\ell_3$)

(a) Model 1                                                    (b) Model 2



high agreement among the different scaling solutions.

In summary, the restricted latent class approach produced results similar to those obtained by Hoijtink's MML approach for this dataset. However, in addition to these similarities, the latent-class solution with three classes revealed a possible split in the Democratic party—a result not obtained by the MML latent trait procedure. Finally, problems with parameter estimation were not observed. The estimates of both unfolding models converged quickly to their ML solutions.

### Application II: A Consumer Study

In an unpublished paired comparison preference study (conducted by the first author), 50 people were asked to compare several brands of American cookies. Four cookie brands (Oreos, Chips Ahoy, Ding Dongs, and Twinkies) were selected from the dataset and submitted to an LCA. Table 6 contains the LR and AIC statistics of the one-class to four-class solutions. A two-class solution provided satisfactory fit, and the corresponding reordered class-specific probabilities are presented in the form of paired comparison matrices in Table 7. Although the gradient pattern was almost satisfied for the latent class probabilities of the first class, several violations occurred for the probabilities of the second latent class. Because these violations make the order of objects 2 and 4 ambiguous, two reordered paired comparison matrices are presented (A and B in Table 7) that seem equally acceptable.

**Table 6**
Number of Latent Classes (LC), Results
of the Likelihood Ratio Test (LRT),
Degrees of Freedom (*df*), and Aikake's
Information Criterion (AIC) in
Application II for Unrestricted LCA

| LC | LRT | *df* | AIC |
|----|-----|------|-----|
| 1 | 103.95 | 57 | 395.86 |
| 2 | 48.46 | 50 | 354.37 |
| 3 | 35.52 | 43 | 355.43 |
| 4 | 23.27 | 36 | 357.18 |

### Table 7
Reordered Class-Specific Paired Comparison Probabilities of Two Latent Class Solutions
(1 = Oreos, 2 = Chips Ahoy, 3 = Ding Dongs, and 4 = Twinkies)

Latent Class 1 With Class Size $p_1$ = 54

| Alternative | 4 | 3 | 1 | 2 |
|---|---|---|---|---|
| 4 | | .631 | .960 | .925* |
| 3 | .369 | | .919 | .923 |
| 1 | .040 | .081 | | .555 |
| 2 | .075* | .077 | .445 | |

Latent Class 2 With Class Size $p_2$ = .46

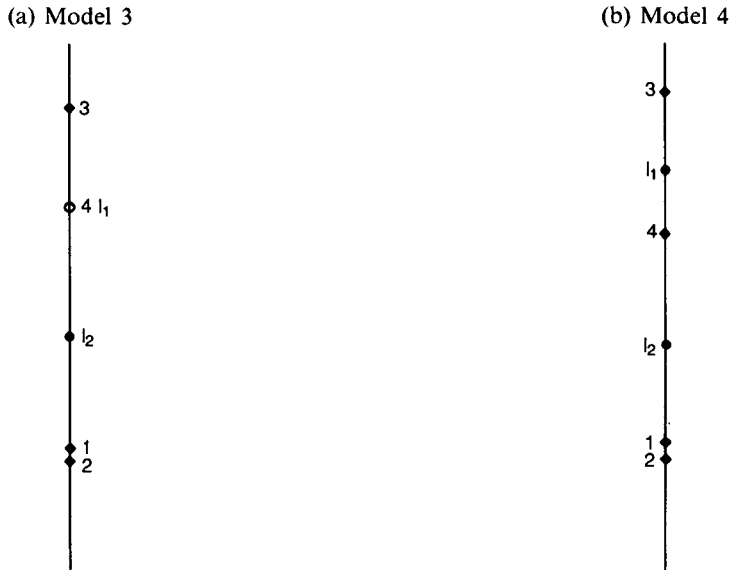| | Reordering A | | | | | Reordering B | | | |
|---|---|---|---|---|---|---|---|---|---|
| Alternative | 1 | 2 | 4 | 3 | Alternative | 1 | 4 | 2 | 3 |
| 1 | | .478* | .610 | .998 | 1 | | .610 | .478* | .998 |
| 2 | .522* | | .437* | .828 | 4 | .390 | | .563 | .782 |
| 4 | .390 | .563* | | .782 | 2 | .522* | .437 | | .828* |
| 3 | .002 | .172 | .218 | | 3 | .002 | .218* | .172 | |

*Violation of the gradient pattern.

By fitting Models 3 and 4 to the paired comparison data, a likelihood ratio test was obtained for Model 3 with $G^2$ = 54.95 ($df$ = 57, AIC = 346.86), and for Model 4 with $\gamma$ = 1, $G^2$ = 54.60 ($df$ = 57, AIC = 346.52). The scale values and ideal point positions obtained from Models 3 and 4 are displayed in Figure 3. Although the scale values obtained from both models are similar, the models differ in two respects. First, for Model 3 the scale value for Twinkies coincides with the ideal point position of latent class $\ell_1$. In contrast, for Model 4 the Twinkies scale value and the ideal point position for latent class $\ell_1$ are well separated, because a scale value and an ideal point coincide in this model only when an object is preferred all the time. Second, the preference rank order of latent class $\ell_2$ is 1, 2, 4, and 3 for Model 3, and 1, 4, 2, and 3 for Model 4. This result reflects the ambiguity in ordering the unrestricted latent-class-specific probabilities as discussed above. It is not clear which solution is preferable. More choice objects and larger sample sizes are needed for a more powerful comparison of the two models.

The two class size parameters were almost equal for both models, indicating that half of the people had a somewhat "sweeter tooth" and preferred Ding Dongs and Twinkies over Oreos and Chips Ahoy, while the remaining half showed (approximately) the reverse pattern. Moreover, the perceived difference between Chips Ahoy cookies and Oreos is considerably smaller than the one between Ding Dongs and Twinkies. Similar to the first application, problems with parameter estimation were not observed.

### Discussion

Although LCA has been used quite infrequently over the last 40 years, Goodman's (1979) formulation of the EM algorithm and the possibility of imposing constraints on the model parameters have led to a revival of this model. In particular, constrained latent class models have been applied successfully to various areas of psychology (Clogg & Goodman, 1984; Dayton & Macready, 1988; Formann, 1985, 1988, 1989; Rindskopf, 1983). Each of these applications demonstrates the usefulness and flexibility of restricted LCA in terms of implementation and parameter estimation. The present

**Figure 3**
Two-Class Solutions for 1 = Oreos, 2 = Chips Ahoy, 3 = Ding Dongs, and 4 = Twinkies
(Latent Class Positions are Denoted by $\ell_1$, and $\ell_2$)

(a) Model 3 (b) Model 4



research focused on the application of two one-dimensional unfolding models of restricted LCA for pick-any/$t$ and paired comparison data. A meaningful classification of the respondents was obtained, as well as a parsimonious representation of the unfolding structure in the data. These results indicate that, in particular for choice datasets with a large number of respondents, the synthesis of unfolding models and LCA may prove to be of considerable usefulness.

Although somewhat different in structure, the unfolding Models 1 and 2, as well as Models 3 and 4, yielded similar fit statistics for the binary choice data. This result may not be unexpected, yet it demonstrates that choice models cannot always be distinguished on the basis of fit statistics when the models satisfy certain qualitative characteristics of the data. Therefore, one important advantage of the proposed synthesis of LCA and unfolding models is that the unrestricted latent class probabilities for these qualitative characteristics (as demonstrated in the two applications) can be inspected before a particular model (or possibly a family of models) is specified and tested that may provide a more concise and parsimonious representation of the choice data.

Finally, one-dimensional unfolding models were focused on in this research. Preference data may be frequently multidimensional, and this limits the range of applicability of these models. Böckenholt and Böckenholt (1990) applied constrained LCA to multidimensional preference data (that may be ordinal) to obtain a wider and more flexible set of models that can account for systematic individual differences.

## References

Abramowitz, A. I., & Stone, W. J. (1984). *Nomination politics: Party activists and presidential choice.* New York: Praeger.

Aikake, H. (1974). A new look at statistical model identification. *IEEE Transactions on Automatic Control,* AC-19, 716–723.

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

Andrich, D. (1989). A probabilistic IRT model for unfolding preference data. *Applied Psychological Measurement, 13,* 193–216.

Bartholomew, D. J. (1987). *Latent variable models and factor analysis.* New York: Oxford University Press.

Bloxom, B. (1972). The simplex in pair comparisons. *Psychometrika, 37,* 119–136.

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

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

Böckenholt, U., & Böckenholt, I. (1990). *Constrained latent class analysis: Simultaneous classification and scaling of discrete data.* Unpublished manuscript, University of Illinois, Champaign.

Cliff, N., Collins, L. M., Zatkin, J., Gallipeau, D., & McCormick, D. J. (1988). An ordinal scaling method for questionnaire and other ordinal data. *Applied Psychological Measurement, 12,* 83–97.

Clogg, C. C., & Goodman, L. A. (1984). Latent structure analysis of a set of multidimensional contingency tables. *Journal of the American Statistical Association, 77,* 803–815.

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

Dayton, C. M., & Macready, G. B. (1988). A latent class covariate model with applications to criterion-referenced testing. In R. Langeheine & J. Rost (Eds.), *Latent trait and latent class models* (pp. 129–146). New York: Plenum.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM-algorithm. *Journal of the Royal Statistical Society B, 39,* 1–22.

DeSarbo, W. S., & Hoffman, D. L. (1986). A new unfolding threshold model for the spatial representation of binary choice data. *Applied Psychological Measurement, 10,* 247–264.

DeSarbo, W. S., & Hoffman, D. L. (1987). Constructing MDS joint spaces from binary choice data: A multidimensional unfolding threshold model for marketing research. *Journal of Marketing Research, 24,* 40–54.

Formann, A. K. (1984). *Die Latent-Class-Analyse: Einführung in Theorie und Anwendung.* [*Latent class analysis: An introduction to theory and applications.*] Weinheim: Beltz Verlag.

Formann, A. K. (1985). Constrained latent class models: Theory and applications. *British Journal of Mathematical and Statistical Psychology, 38,* 87–111.

Formann, A. K. (1988). Latent class models for nonmonotone dichotomous items. *Psychometrika, 53,* 45–62.

Formann, A. K. (1989). Constrained latent class models: Some further applications. *British Journal of Mathematical and Statistical Psychology, 42,* 37–54.

Goodman, L. A. (1978). *Analyzing qualitative/categorical data: Log-linear models and latent-structure analysis.* Cambridge: Abt Books.

Goodman, L. A. (1979). On the estimation of parameters in latent structure analysis. *Psychometrika, 44,* 123–128.

Greenberg, M. G. (1965). A method of successive cumulations for scaling of pair-comparison judgments. *Psychometrika, 30,* 441–448.

Hoijtink, H. (in press). A latent trait model for dichotomous choice data. *Psychometrika.*

Hubert, L. J., & Golledge, R. G. (1981). Matrix reorganization and dynamic programming: Applications to paired comparisons and unidimensional seriation. *Psychometrika, 46,* 429–441.

Langeheine, R., & Rost, J. (1988). *Latent trait and latent class models.* New York: Plenum.

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

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

McCutcheon, A. L. (1988). *Latent class analysis.* Newsbury Park: Sage.

Rindskopf, D. (1983). A general framework for using latent class analysis to test hierarchical and non-hierarchical learning models. *Psychometrika, 48,* 85–97.

Rost, J. (1985). A latent class model for rating data. *Psychometrika, 50,* 37–49.

Rost, J. (1988a). Rating scale analysis with latent class models. *Psychometrika, 53,* 327–348.

Rost, J. (1988b). *Quantitative and qualitative probabilistische Testtheorie.* [*Quantitative and qualitative probabilistic test theory.*] Bern: Huber.

Takane, Y. (1983a). *Choice model analysis of the "pick-any/n" type of binary data.* Paper presented at the European Psychometric and Classification Meetings, Jouy-en-Josas, France.

Takane, Y. (1983b). *An item response model for multidimensional analysis of multiple-choice data.* Paper presented at the Psychometric Society Meeting, Los Angeles CA, U.S.A.

Takane, Y. (1987). Analysis of covariance structures and probabilistic binary choice data. *Cognition and Communication, 20,* 45–62.

Takane, Y., & DeLeeuw, J. (1987). On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika, 52,* 393–408.

## Acknowledgements

## Author's Address

Send requests for reprints or further information to Ulf Böckenholt, Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign Il 61820, U.S.A.