KM 42 (1993)
pg.109-127

Analysis of 'pick any/n' data with no prior knowledge:
A Synthesis of Latent Class Analysis and Unfolding Models

Ulf Böckenholt

Abstract

A synthesis of unfolding models and latent class analysis is presented for modeling 'pick-any/n' data. The latent class part of the model identifies homogeneous subgroups that are characterized by their choice probabilities for a set of items, and the unfolding part accounts for the single-peakedness structure of the choice data. Two applications are presented to illustrate this approach, one of which involves the analysis of group differences.

Key words: Latent class analysis, pick any/n data, unfolding models.

The author is grateful to Herbert Hoijtink for helpful comments on an earlier draft of this paper.
Requests for reprints should be sent to Ulf Böckenholt, Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, Il 61820, U. S. A.

Introduction

Coombs' (1964) unfolding technique is a conceptually simple yet powerful approach for analyzing preference data. Coombs assumed that in a choice situation persons compare each alternative to their ideal or most preferred alternative. When asked to pick, for example, the most preferred out of several alternatives a person selects the alternative that is closest or least dissimilar to her ideal alternative. An important constraint of Coombs' unfolding approach is that, although persons may have different ideal alternatives, they agree on the similarity relationships among the choice alternatives. For example, if the choice alternatives share a common latent continuum, their positions along this continuum are perceived equally by all persons, however the positions of the individual ideal points may differ from person to person.

Numerous models for the analysis of choice data have been proposed that are based on Coombs' unfolding idea. However, a major problem is the treatment of individual differences when estimating the parameters of these unfolding models. In particular, joint estimation of the ideal points and the choice alternatives' positions by maximum likelihood methods is complicated by the fact that standard limit theorems do not apply because the number of ideal point parameters changes as a function of the sample size (Bock & Aitkin, 1981). By developing a synthesis of latent class and unfolding models, Böckenholt and Böckenholt (1990; 1991) provided a satisfactory solution to this problem. Their approach is based on the assumption that heterogeneity is limited and can be accounted for by assigning persons to different classes where each class is characterized by its ideal point and other class-specific parameters. This paper describes this approach for the analysis of 'pick any/n' data and presents results from the analysis of two data sets.

Unfolding Models for 'Pick Any/n' Data

Frequently, choice data are collected by asking respondents to select the preferred alternatives from a set of n alternatives. When Coombs (1964) introduced this 'pick any/n' procedure, he posited that a person selects those alternatives that are closest to the position of the most preferred alternative. More formally, let the positions of choice object i and of person's a ideal point be δ_i and β_a, respectively. A response of person a with respect to alternative i is denoted by the binary variable X_{ai}. Choice alternatives are chosen when their distance to the ideal point is smaller than some threshold τ_a specific to person a,

$$X_{ai} = 1 \text{ when } |\delta_i - \beta_a| \leq \tau_a$$

and are not chosen otherwise,

$$X_{ai} = 0 \text{ when } |\delta_i - \beta_a| > \tau_a.$$

Over the years this approach was refined in several ways. For example, the choice objects' representations were proposed to be multi-dimensional instead of unidimensional (Bennett & Hays, 1960), the positions of objects and/or ideal points were hypothesized to be random instead of fixed variables (Zinnes & Griggs, 1974). Similarly, distances between the ideal point and the alternatives' positions were proposed to be random (Ramsay, 1980), and the choice rule was modified to be probabilistic instead of deterministic (Schöneman & Wang, 1972). These and other extensions increased the applicability range of Coombs' original model and proved useful in a wide variety of studies (Bossuyt, 1990).

For the presentation of the synthesis of unfolding models and latent class analysis, we focus on two probabilistic unidimensional unfolding models that were recently proposed for binary 'pick any/n' data (Andrich, 1988; DeSarbo & Hoffman, 1986; Hoijtink, 1990). These models were selected for their simple structure and ease of implementation. However, because of their parsimonious form, these models may not always prove sufficient for the analysis of unfolding data. In this case, more complex probabilistic unfolding models may be combined with a latent-class approach (see Böckenholt & Böckenholt, 1991).

Both Andrich (1988), and DeSarbo and Hoffman (1986) proposed the unfolding threshold (UT) model for binary 'pick any/n' data. According to this model persons are characterized by their ideal point positions, β_a, and by a threshold parameter, τ_a, which describes the range of the continuum for a positive response. The probability of choosing an alternative is given by

$$\Pr(X_{ai}=1) = p_{ai} = \frac{1}{1 + \exp\{-\tau_a + (\delta_i - \beta_a)^2\}}. \tag{1}$$

The smaller the distance between the ideal point and the object position, the higher the probability that alternative i is chosen by person a. However, note that the probability of selecting an alternative when $\delta_i = \beta_a$ is only larger than .5 when $\tau_a > 0$. This dependency may introduce high correlations among the parameters of the UT model in empirical applications.

Hoijtink (1990) proposed a different unfolding model under the premise that a probabilistic unfolding model should have its deterministic version as a boundary case,

$$\Pr(X_{ai}=1) = p_{ai} = \frac{1}{1 + \{(\delta_i - \beta_a)^2\}^{\gamma}}, \tag{2}$$

where γ moderates the strength of the proximity relation on the choice probability whenever the distance between the ideal point and the object position differs from 0 or 1. This model, called here unfolding power (UP) model, makes the strong prediction that an alternative is chosen with certainty when its position coincides with the ideal point. In contrast, when an alternative's position coincides with the ideal point in the UT model, the probability of selecting the alternative depends also on the threshold parameter. As a result, the UP and the UT model may yield different unidimensional scales.

Latent Class Unfolding Models

It is well-known that the joint estimation of person (i.e., τ_a and β_a) and item parameters (i. e., δ_i) is problematic because with increasing sample size the number of person-specific parameters to be estimated also increases. As a result, standard limit theorems do not apply and parameter estimates are not consistent. One satisfactory solution to this problem is provided by the marginal maximum likelihood (MML) method, which in this context involves the estimation of the distribution of the ideal points (Hoijtink, 1990; Takane, 1983). However, the advantages of the MML method are gained at a considerable computational expense. A computationally more attractive approach follows from the assumption that heterogeneity is limited and can be described by grouping respondents with small intragroup and large intergroup differences. Because the grouping factor may not be observable or known a priori, Böckenholt and Böckenholt (1990; 1991) suggested a reparameterization of latent class analysis (Lazarsfeld & Henry, 1968) as a modeling framework for this problem. According to their approach, each group or latent class is represented by an ideal point and other class-specific parameters that describe individual differences in choice behavior. The unobserved classes are determined by invoking the principle of local independence which states that latent-class membership variables account completely for any relationships among the observed choices. In other words, choices among the alternatives are made independently of each other when class membership is known. Consequently, the probability of observing a choice pattern given that person a is a member of latent class t is

$$\Pr\{\mathbf{x}_a = (x_{a1}, \ldots, x_{ai}, \ldots, x_{an}) | \; a \epsilon \; t\} = \prod_{i=1}^{n} p_{i|t}^{\;x_{ai}} \; (1 - p_{i|t})^{1 - x_{ai}}.$$

Coombs (1964) also suggested latent-class analysis (LCA) for 'pick any/n' data. Distinguishing between techniques that do or do not require a priori knowledge about the ordering of choice alternatives, he argued that LCA is designed for the analysis of 'pick any/n' data with "no prior knowledge" (Coombs, 1964, p. 318). However, when applying LCA it is only assumed that every item has a certain probability of being selected and the underlying choice mechanism is left unspecified. As a result, LCA does not provide information about the underlying scale of the choice alternatives and the positions of the ideal points. However, this information can be extracted from the data by constraining the class-specific probabilities to conform to an unfolding structure. For example, we may constrain the class-specific probabilities to follow the UT model,

$$p_{i|t} = \Pr(X_{ai} = 1 | \; a \epsilon \; t) = \frac{1}{1 + \exp\{-\tau_t + (\delta_i - \beta_t)^2\}}. \tag{3}$$

Thus, every member of latent class t is characterized by an ideal point, β_t, and by a threshold parameter τ_t. Similarly, by combining the latent class model with the UP representation, we obtain

$$p_{i|t} = \Pr(X_{ai} = 1 \mid a \in t) = \frac{1}{1 + \{(\delta_i - \beta_t)^2\}^\gamma}. \tag{4}$$

The unconditional probability of observing the choice of an alternative is

$$p_i = \sum_{t=1}^{T} \pi_t \Pr(X_{ai} = 1 \mid a \in t),$$

where π_t denotes the probability of a person belonging to latent class t, and $\sum_{t=1}^{T} \pi_t = 1$. Similarly, the unconditional probability of observing a response pattern is

$$\Pr\{\mathbf{x}_a = (x_{a1}, \ldots, x_{ai}, \ldots, x_{an})\} = \sum_{t=1}^{T} \pi_t \prod_{i=1}^{n} p_{i|t}^{\,x_{ai}} \, (1 - p_{i|t})^{1 - x_{ai}}.$$

The combination of LCA with unfolding models has several advantages. First, for a given number of latent classes we can examine the unrestricted latent class-specific probabilities to determine whether they conform to an unfolding structure, or more specifically, to the unfolding structures predicted by the UP and the UT models. For example, according to both models choice probabilities are symmetric around the ideal point, and they satisfy strong stochastic transitivity as well as the condition of bilateral monotonicity (Bossuyt, 1990). If these characteristics are met by the class-specific probabilities, we can constrain them by Eqs. (3) or (4) to obtain a parsimonious description of the discrete choice data. Instead of estimating (n T) unconstrained choice probabilities, we only need to determine T ideal points, (n - 1) item parameters, and depending on the unfolding model, either T threshold parameters or one power parameter. Second, by jointly representing choice objects and ideal points on the class-level, we obtain an easily interpretable description of individual differences even when the number of respondents is large. Third, a decision regarding the number of latent classes which is usually unknown is not dependent on the specification of the unfolding model and may be based on the results of the unrestricted latent class analysis. Finally, estimation procedures for latent-class unfolding models are straightforward and easy to implement. This topic is discussed in more detail in the next section.

Parameter Estimation

To simplify the notation, subjects with identical choice patterns are grouped and f_v denotes the observed frequency for the v-th choice pattern $\mathbf{x}_v$. Under the assumption of random sampling of N persons, the likelihood function of the data may be written as

$$L = \prod_{v=1}^{2^n} \left\{ \sum_{t=1}^{T} \pi_t \prod_{i=1}^{n} p_{i|t}^{\,x_{vi}} \, (1 - p_{i|t})^{1 - x_{vi}} \right\}^{f_v}.$$

The expectation maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977; Goodman, 1979) has proven useful for estimating the parameters of latent class models. We present both the expectation (E) and the maximization (M) step and discuss the modifications necessary for estimating the parameters of the unfolding models.

E Step

Under the assumption of local independence, the joint probability of observing the choice pattern v in class t is computed as

$$p_{tv} = \pi_t \prod_{i=1}^{n} p_{i|t}^{x_{vi}} (1 - p_{i|t})^{1 - x_{vi}}$$

and

$$p_v = \sum_{t=1}^{T} p_{tv}.$$

Consequently, the posterior probability that a respondent (i.e., choice pattern) is a member of latent class t is $p_{t|v} = p_{tv}/p_v$. The expected number of persons selecting alternative i in latent class t is given by

$$E(f_{i|t}) = \sum_{v=1}^{2^n} f_v \, p_{t|v} \, x_{vi}.$$

Similarly, the expected number of persons who do not select alternative a in latent class t is given by

$$E(f_{\bar{i}|t}) = \sum_{v=1}^{2^n} f_v \, p_{t|v} (1 - x_{vi}).$$

Thus, the expected number of persons in class t is

$$E(f_t) = E(f_{i|t}) + E(f_{\bar{i}|t}) = n \, \pi_t.$$

M Step

In the M step the class size parameters and the parameters of the latent class models are determined by treating the expected values determined in the E-step as if they were observed and maximizing the kernel of likelihood function of the 'complete' data

$$L^* = \prod_{t=1}^{T} \pi_t^{E(f_t)} \prod_{i=1}^{n} p_{i|t}^{E(f_{i|t})} (1 - p_{i|t})^{E(f_{\bar{i}|t})},$$

where $p_{i|t}$ is either unconstrained or constrained by Eqs. (3) or (4). In the unconstrained latent class model, the class-specific probabilities are estimated by $\hat{p}_{i|t} = E(f_{i|t})/E(f_t)$, and the class size parameters, π_t, are estimated by $\hat{\pi}_t = E(f_t)/N$. These estimates are used to update the posterior probabilities and expected values determined in the E-step of the algorithm. The procedure iterates between the E-step and the M-step until convergence is obtained.

To estimate the parameters of the UT latent-class models an iterative estimation procedure must be used that maximizes the likelihood function of the complete data. Because the origin of the scale is arbitrary in the UT model, the positions of the items are estimated under the linear constraint that their sum is zero. This linear constraint is implemented by setting $\boldsymbol{\delta} = \mathbf{C}\,\boldsymbol{\delta}^*$, where $\boldsymbol{\delta}^*$ contains the reduced set of item parameters and $\mathbf{C}$ is a (n x s) matrix with element c_{iq}. For example, for four items, $\mathbf{C}$ can be specified as

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix},$$

and $\delta_4 = -(\delta_1{}^* + \delta_2{}^* + \delta_3{}^*)$. As a result, the following partial derivatives are derived for the UT model

$$\frac{\partial \ln \mathrm{L}^*}{\partial \tau_t} = \sum_{i=1}^{n} \mathrm{e}_{ti},$$

$$\frac{\partial \ln \mathrm{L}^*}{\partial \beta_t} = 2 \sum_{i=1}^{n} \mathrm{e}_{ti}\,(\beta_t - \delta_i),$$

$$\frac{\partial \ln \mathrm{L}^*}{\partial \delta_q{}^*} = -2 \sum_{t=1}^{T} \sum_{i=1}^{n} \mathrm{e}_{ti}\,(\beta_t - \delta_i)\,\mathrm{c}_{iq},$$

where $\mathrm{e}_{ti} = -\,\{E(\mathrm{f}_{\bar{i}|t})\,\mathrm{p}_{i|t} + E(\mathrm{f}_{i|t})\,(1 - \mathrm{p}_{i|t})\}$. In applications, the Davidson-Fletcher-Powell method (Luenberger, 1984) proved adequate for estimating the UT model parameters at each M-step. However, it is possible to accelerate the convergence of the estimation procedure by modifying the EM algorithm (Louis, 1982) or by combining it with a scoring algorithm. The estimation of the latent-class UP model follows a similar procedure but requires more care in the selection of initial starting values. We refer to Hoijtink (1990) for a detailed discussion of the relevant estimation issues for this model.

It is well-known that the likelihood function of a latent class model may have multiple maxima (Haberman, 1974). Because the EM-algorithm does not ensure convergence to a global maximum different sets of starting values were used in the reported applications. These starting values were selected from an uniform (0, 1) distribution for the unconstrained latent class model. Starting values for the UP and UT models were obtained by preliminary least-squares analyses of the unconstrained class-specific probabilities. However, it should be emphasized that the use of different starting values does not eliminate the problem of local maxima.

Model tests

Provided standard regularity and identifiability conditions (Birch, 1964) are satisfied, the likelihood ratio (LR) statistic can be used for large sample tests of fit of any of the restricted or unrestricted latent class models,

$$G^2 = 2 \sum_{v=1}^{2^n} f_v \ln\left(f_v/\hat{f}_v \right),$$

where $\hat{f}_v = N \hat{p}_v$ denotes the expected frequency for the v-th choice pattern. Under the assumption of a multinomial distribution, there are $2^n - 1$ degrees of freedom. The unrestricted latent class model requires the estimation of $(T\ n)$ class-specific probabilities and $(T - 1)$ class size parameters. Constraints imposed by the unfolding models on the class-specific probabilities reduce the number of effective parameters. For example, the UT model requires the estimation of T ideal points, T class-threshold parameters, $(n - 1)$ item parameters, and $(T - 1)$ class size parameters. Thus, the degrees of freedom of this model are $(2^n - 3\ T - n + 1)$. Similarly, the degrees of freedom for the UP model are $(2^n - 2\ T - n)$.

For two nested hypotheses, the difference between the corresponding LR-statistics can be computed to assess the importance of the contribution to the LR-statistic by the additional constraints imposed by the stronger hypothesis. This difference, denoted by ΔG^2, is asymptotically distributed as a chi-squared statistic with degrees of freedom equal to the difference between the number of parameters in both models. This approach provides further guidance in selecting a parsimonious model. Unfortunately, no test is available for the comparison of latent class models that are not proper subsets of each other. For example, a LC model with $(t - 1)$ classes is not nested in a LC model with t classes. In this case, we choose the LC model that has an acceptable fit.

No completely satisfactory solutions are presently known when many of the expected frequencies are too small to justify the distributional assumptions of the LR-statistic. One strategy is to group response patterns until their expected frequencies exceed a specified minimum value and to compute the Pearson χ^2-statistic of the expected and observed frequencies for the K groupings

$$X^2 = \sum_{k=1}^{K} \frac{(f_k - \hat{f}_k)^2}{\hat{f}_k}.$$

This approach has two undesirable features. First, decisions regarding the grouping of the frequencies are to some extent arbitrary and may lead to different conclusions. Second, the test statistic of the grouped data does not follow a chi-square distribution under the null hypothesis. However, under certain regularity conditions a lower bound of the limiting distribution of X^2 is provided by the chi-square distribution with degrees of freedom computed as the difference between the by one reduced number of grouped choice patterns and the number of effective model parameters (Bishop, Fienberg, & Holland, 1975). As a result, asymptotic significance levels found from this reference distribution may be too small.

Applications

Analysis of nuclear energy data: In this section we analyze the binary responses ('I agree', 'I do not agree') of $N = 600$ persons to $n = 5$ items (in Table 1) selected from a questionnaire measuring general attitudes toward nuclear energy (Formann, 1988). The items seem to span a continuum which ranges from a more positive to a more negative attitude toward nuclear power.

Table 1

Items of Nuclear Energy Questionnaire

(1) In the near future alternative sources of energy will not be able to substitute nuclear energy.

(2) It is difficult to decide between the different types of power stations if one carefully considers all their pros and cons.

(3) Nuclear power stations should not be put into operations before the problem of radio-active waste has been solved.

(4) Nuclear power stations should not be put into operations before it is proven that the radiation caused by them is harmless.

(5) The foreign power stations now in operations should be closed down.

Before we report the results obtained by constraining the latent class-specific probabilities to conform to an unfolding model, we examine the unconstrained latent class solutions. The G^2 statistics obtained for the one, two, three, and four class solutions are 249.5 (df = 26), 64.0 (df = 20), 29.6 (df = 14), and 8.4 (df = 8), respectively. Table 2 gives the corresponding class-specific 'agree'-probabilities and the class size parameter estimates for the two, three, and four class solutions. Overall, the pattern of the class-specific probabilities follows an unfolding or single peakedness pattern. Violations of this structure are predominantly a result of Item 2. However, the three and four class solutions are not well-determined. About 30% of the class-specific probabilities for both latent-class models are close or equal to their boundary values of 0 and 1. In addition, the smallest class of the three and four class solutions contains only about 31 and 13 respondents, respectively. Because these results indicate identifiability and stability problems, it seems preferable to focus on the solution of the two-class model. However, to better understand the reasons for the poor fit of this model, a detailed residual analysis was performed. This analysis showed that the two-class model does not account well for the association between items 3 and 4. Most likely, this association is a result of the items' similar phrasing. Significant fit improvements can thus be obtained by either adding an association parameter to the two-class model or by omitting one of the two items from the analysis. For example, when omitting item 4, the two-class model gives a satisfactory fit with $G^2 = 6.1$ (df = 6).

Table 2

Unconstrained LC Solutions with Two, Three, and Four Classes

		Class-Specific Probabilities				
Class	$\hat{\pi}$	Item 1	Item 2	Item 3	Item 4	Item 5
1	.54	.48	.58	.90	.70	.06
2	.46	.14	.34	.74	.95	1.00
1	.54	.48	.58	.90	.69	.08
2	.41	.15	.37	.83	1.00	1.00
3	.05	.00	.14	.00	.59	.98
1	.18	.54	.49	.78	.27	.00
2	.44	.46	.63	.97	.91	.27
3	.37	.08	.28	.72	1.00	.98
4	.02	.00	.15	.00	.00	1.00

As the unconstrained latent class model, the two-class UP and UT models do not yield a satisfactory fit to the data with $G^2 = 83.6$ (df = 22) and $G^2 = 102.7$ (df = 23), respectively. Consequently, they provide only a limited description of the information in the data. Table 3 displays the class-specific probabilities of the two-class UP and UT model and Figures 1 and 2 give a graphical representations of the parameter estimates. Each plot also contains the corresponding unconstrained class-specific probabilities (depicted by circles and squares for the first and second class, respectively). Although the respondents' class-memberships differ slightly for the unconstrained latent class, and the UP/UT models, it is clear that the UT model follows the unconstrained class-specific probabilities more closely than the UP model. As a result of the models' different parametrization the positions of the ideal points are different. However, the UP and the UT model agree with respect to the relative ordering of the items. Thus, the initial expectation that the items span a continuum seems confirmed. Favoring Item 3, the first latent class has a less extreme attitude toward nuclear power than the second latent class which agrees more strongly with Items 4 and 5.

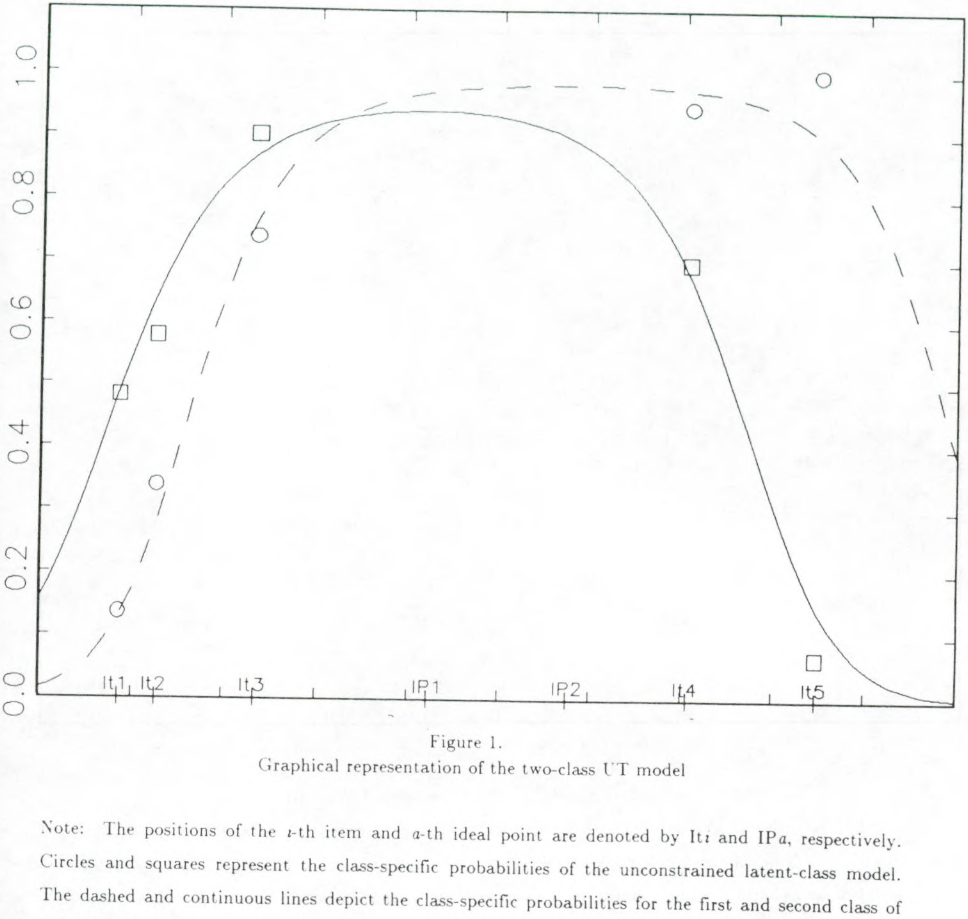

Figure 1.
Graphical representation of the two-class UT model

Note: The positions of the i-th item and a-th ideal point are denoted by Iti and IPa, respectively. Circles and squares represent the class-specific probabilities of the unconstrained latent-class model. The dashed and continuous lines depict the class-specific probabilities for the first and second class of the UT model.

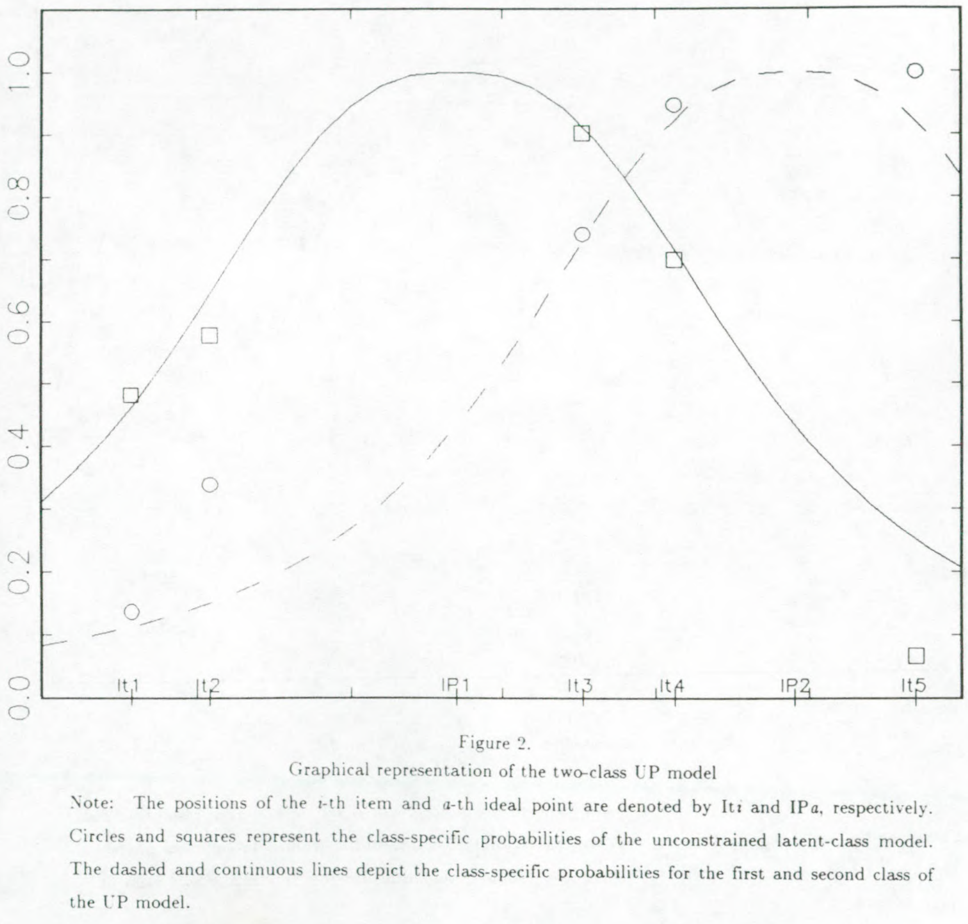

Figure 2.

Graphical representation of the two-class UP model

Note: The positions of the i-th item and a-th ideal point are denoted by Iti and IPa, respectively. Circles and squares represent the class-specific probabilities of the unconstrained latent-class model. The dashed and continuous lines depict the class-specific probabilities for the first and second class of the UP model.

Table 3

Parameter Estimates of Two Class UP and UT models

	Class	$\hat{\pi}$	Item 1	Item 2	Item 3	Item 4	Item 5
			Class-Specific Probabilities				
UP	1	.62	.46	.64	.91	.71	.25
	2	.38	.11	.15	.72	.92	.92
UT	1	.54	.49	.64	.87	.68	.14
	2	.46	.13	.28	.76	.98	.90

Attitude toward car-use and environment: The second data set contains binary responses ('I agree', 'I do not agree') to $n = 10$ items (in Table 4) that measure attitudes toward car-use and environment. The responses were collected from two independent samples of $N = 300$ respondents each before and after a pro-environment information campaign. The purpose of the following analyses is to examine whether the campaign changed the response behavior of the respondents.

Table 4

Items of Information Campaign Questionnaire

(1) Car use cannot be abandoned. Some pressure on the environment has to be accepted.

(2) A cleaner environment demands for sacrifices like a decreasing car usage.

(3) The environmental problem justifies a tax burden on car driving so high that people quit using a car.

(4) Putting a somewhat higher tax burden on car driving is a step in the direction of a healthier environment.

(5) It is better to deal with other forms of an environmental pollution then car driving.

(6) Instead of environmental protection measures with respect to car use, the road system should be extended.

(7) Technically adapted cars do not constitute an environmental threat.

(8) Considering the environmental problems, everybody should decide for themselves how often to use the car.

(9) People who keep driving a car, are not concerned with the future of our environment.

(10) Car users should have to pay taxes per mile driven.

The sample size is small in relation to the number of items and only about 13% of the possible response patterns are observed. Because likelihood-based inferences for latent-class models may be problematic in such cases of sparse data tables, a subset of six items was chosen for further analysis. To ensure that changes regarding a pro-car and a pro-environment attitude can be assessed, three pro-environment items (2, 4, 10) and three pro-car items (5, 6, 8) were selected.

Table 5 contains goodness-of-fit results obtained from the analysis of the pre- and post-information campaign data by the unconstrained latent class model. G^2 is computed for the ungrouped data. Pearson's X^2 is obtained by grouping the data such that the minimum expected frequency exceeds one and the corresponding degrees of freedom are (K - 1 - number of estimated parameters). According to both statistics at least two latent classes are required for a satisfactory fit of the data. Overall, the fit of the latent class models seems better for the pre- than for the post-campaign data. The last two columns of Table 5 contain LR-statistics and their corresponding degrees of freedom obtained when testing the hypothesis that the latent-class parameters (class-specific probabilities and class sizes) are equal for both studies. For all latent-class solutions this hypothesis can be rejected which indicates that the information campaign may have had some effect on the respondents' attitudes toward car-use and environment.

Table 5

Goodness of Fit Statistics of Unconstrained Latent Class Models for

Pre- and Post Information Campaign Data

	Pre-Campaign				Post-Campaign				Equality Tests	
Class	G^2	df	X^2	df	G^2	df	X^2	df	ΔG^2	Δdf
1	295.5	55	565.0	53	271.8	55	440.9	47	21.5	6
2	56.2	48	31.4	31	82.2	48	46.8	28	35.0	13
3	35.7	41	10.9	22	61.5	41	28.8	20	44.3	20
4	28.0	34	10.2	17	44.2	34	17.4	11	57.8	27

To examine the effect of the information campaign in more detail, we need to investigate which subsets of the latent-class parameters differ significantly between the samples. For example, we may partition the LR-statistic ΔG^2 into two components by testing the equality of the class-specific probabilities with (6 T) degrees of freedom and by testing the equality of the class sizes with (T - 1) degrees of freedom. Although not presented in detail here, these partitions indicate that the main reason for the differences between the pre- and post-campaign data is related to changes in the class-size estimates. Table 6 contains the estimates of the three class solution with class-specific probabilities constrained to be equal for the pre- and post campaign data but different class sizes. With some minor exceptions the latent-class specific probabilities display a single-peakedness structure. As a result, a more parsimonious and informative description of the data may be obtained by an unfolding latent class model.

Table 6

Simultaneous Latent Class Analysis of Pre- and Post-Information Campaign Data

Class	It 6	It 5	It 8	It 2	It 10	It 4	$\hat{\pi}_1$	$\hat{\pi}_2$
1	.75	.98	.91	.33	.14	.16	.43	.25
2	.22	.74	.77	.84	.34	.32	.34	.51
3	.06	.40	.37	.98	.91	.85	.23	.24

For example, when applying the three-class UP model we obtain the parameter estimates given in Table 7. These results differ little from the unconstrained latent-class solution (in Table 6) and the Pearson X^2-statistic for the grouped data is 73.2 with 72 degrees of freedom. Although the overall fit of the latent-class UP model seems satisfactory, further residual analyses indicate that the UP model may be too restrictive for some items of the post-campaign data. In particular, the endorsement frequencies for Item 6 are not predicted well by the UP model. Overall, however, the three-class UP model gives an informative summary of the data and provides a simple explanation for the difference between the pre- and post-campaign data.

Table 7

Latent Class-Specific Probabilities and Parameter Estimates of Simultaneous Pre- and Post-Campaign Analysis by Unfolding Power Model

Class	It 6	It 5	It 8	It 2	It 10	It 4	$\hat{\beta}$	$\hat{\pi}_1$	$\hat{\pi}_2$
1	.72	.95	.92	.29	.13	.12	-.956	.44	.26
2	.22	.75	.80	.88	.40	.37	.004	.34	.52
3	.10	.31	.34	.97	.88	.85	.723	.22	.22
$\hat{\delta}$:	-1.65	-.64	-.58	.46	1.18	1.23		$\hat{\gamma} = 1.26$	

According to the UP model, the items can be ordered along a continuum ranging from the attitude that "car-use does not pose an environmental problem" to the position that "car-use damages the environment and should be restricted by some governmental interventions". In both the pre- and post-campaign data, about 22% of the respondents favor a strong pro-environment position. Before the campaign a substantial number of the respondents (44%) did not consider car-use to be an environmental problem. One effect of the information campaign was to reduce the size of this group and to increase the number of respondents who acknowledge the negative influence of car-usage on the environment. Perhaps not surprisingly, however, the campaign did not increase the number of respondents favoring a tax increase as a means to reduce car usage.

One obvious disadvantage in examining only six instead of the original ten items is a loss in power for detecting effects of the information campaign. Adopting a more exploratory approach we therefore analyzed the complete questionnaire by the unconstrained and the unfolding latent class models. Although these analyses revealed three distinct classes similar to the ones found when using six items, no satisfactory unidimensional representation of the items was obtained by the latent-class UP or UT models. Instead a two-dimensional UT model (Böckenholt & Böckenholt, 1991) provided a significantly better fit than its one-dimensional counterpart. One major result obtained by applying this model to the pre- and post-campaign data is that there is not only a change in the class sizes but also that the ideal-points shift from more extreme positions (either pro-environment or pro-car) to more moderate ones. Thus, another effect of the information campaign (not captured by the analysis of the six items) seems to be that respondents of the post-campaign survey were more homogeneous in their attitudes toward car-use and environment.

Conclusion

One of the most important problems in choice modeling is to account for the effects of preference heterogeneity in a parsimonious and versatile way. A satisfactory solution to this problem is provided by the synthesis of unfolding models and latent class analysis. This synthesis yields a general approach for modeling individual differences in choice data and is not restricted to the analysis of 'pick any/n' data or to a particular choice model. For example, the approach can be extended to 'pick k/n' data. In this case only certain choice patterns can occur and a constrained quasi-latent class approach may be utilized to describe the unfolding structure in the data. Restricted LCA can also be used to account for individual differences in other kinds of choice data such as paired comparisons and (partial) rankings (Böckenholt, 1992; Croon, 1989; Formann, 1992). Similarly, other choice models, such as the vector model (Carroll, 1980) or other types of constraints, such as inequality constraints for testing ordinal restrictions on the class-specific probabilities (Bossuyt, 1990; Böckenholt, 1990; Croon, 1991), may prove more appropriate in some applications and can easily be integrated into a latent class framework. Thus, restricted LCA provides a general and flexible framework well-suited for the joint modeling of the respondents' preferences and the choice alternatives' similarity structure.

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.

Bennett, J. F., & Hays, W. L. (1960). Multidimensional unfolding: Determining the dimensionality of ranked preference data. Psychometrika, 25, 27-43.

Birch, M. W. (1964). A new proof of the Fisher-Pearson theorem. Annals of Mathematical Statistics, 35, 817-824.

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

Böckenholt, U. (1990). Latent class analysis under order constraints. Paper presented at the Annual Meeting of the Classification Society, Utah State University, Logan.

Böckenholt, U. (1992). Applications of Thurstonian models to ranking data. In: M. Fliegner & J. Verducci (Eds.), Probability Models and Statistical Analyses for Ranking Data (in press). New York: Springer.

Böckenholt, U, &. Böckenholt, I. (1990). Modeling individual differences in unfolding preference data: A restricted latent class approach. Applied Psychological Measurement, 14, 257-269.

Böckenholt, U, &. Böckenholt, I. (1991). Constrained latent class analysis: Simultaneous classification and scaling of discrete choice data. Psychometrika, 56, 699-716.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. Psychometrika, 46, 443-459.

Bossuyt, P. (1990). A comparison of probabilistic unfolding theories for paired comparisons data. New York: Springer.

Carroll, J. D. (1980). Models and methods for multidimensional analysis of preferential choice data. In: E. D. Lantermann & H. Feger (Eds.), Similarity and choice (pp. 234-289). Bern: Huber

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

Croon, M. A. (1989). Latent class models for the analysis of rankings. In: H. Feger, K. Klauer, G. de Soete (Eds.), New developments in psychological choice modeling (pp. 99-121). Amsterdam: North-Holland.

Croon, M. (1991). Latent class models with ordered latent classes. Paper presented at the 7th European Meeting of the Psychometric Society. Trier, Germany.

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.

Formann, A. K. (1988). Latent class models for non-monotone dichotomous items. Psychometrika, 53, 45-62.

Formann, A.K. (1992) Linear logistic latent class analysis for polytomous data. Journal of the American Statistical Association, 87, 476-486.

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

Haberman, J. S. (1974). Log-linear models for frequency tables derived by indirect observations: Maximum likelihood equations. The Annals of Statistics, 2, 911-924.

Hoijtink, H. (1990). A latent trait model for dichotomous choice data. Psychometrika, 55, 641-656.

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

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. Journal of the Royal Statistical Society B, 44, 226-223.

Luenberger, D. G. (1984). Linear and nonlinear programming. MA: Addison-Wesley.

Ramsay, J. O. (1980). The joint analysis of direct ratings, pairwise preferences and dissimilarities. Psychometrika, 45, 149-166.

Schönemann, P. H., & Wang, W. M. (1972). On individual differences model for the multidimensional analysis of preference data. Psychometrika, 37, 275-309.

Takane, Y. (1983). Choice model analysis of the 'pick any/n' type of binary data. Paper presented at the European Psychometric and Classification Meeting, Jouy-en-Joas, France.

Zinnes, J. L. , & Griggs, R. A. (1974) Probabilistic multidimensional unfolding analysis. Psychometrika, 39, 327-350.