

1. Examples and Tutorials

1.1 Historical Overview

Latent class (LC) modeling was initially introduced by Lazarsfeld and Henry (1968) as a way of formulating latent attitudinal variables from dichotomous survey items. In contrast to factor analysis, which posits continuous latent variables, LC models assume that the latent variable is categorical, and areas of application are more wide-ranging. The initial methodology was formalized and extended to nominal variables by Goodman (1974a, 1974b) who also developed the maximum likelihood (ML) algorithm that serves as the basis for many of today's LC software programs. In recent years, LC models have been extended to include observable variables of mixed scale type (nominal, ordinal, continuous and counts), covariates, and other extensions to deal with sparse data, boundary solutions, and other problem areas. Latent GOLD® 3.0 incorporates many of these extensions in a Windows-friendly environment.

Traditional LC analysis (i.e., Goodman, 1974a) assumes that each observation is a member of one and only one of T latent (unobservable) classes, and that *local independence* exists between the manifest variables. That is, conditional on latent class membership, the manifest variables are mutually independent of each other. This model can be expressed using (unconditional) probabilities of belonging to each latent class, and conditional response probabilities as parameters. Traditional LC analysis dealt almost exclusively with a small number of nominal variables. For example, in the case of 4 nominal manifest variables A , B , C , and D we have:

$$P_{ijklt} = P_t^X P_{it}^{A|X} P_{jt}^{B|X} P_{kt}^{C|X} P_{lt}^{D|X} \quad (1)$$

where P_t^X denotes the probability of being in latent class $t=1,2,\dots,T$ of latent variable X ; $P_{it}^{A|X}$ denotes the conditional probability of obtaining the i th response to item A , from members of class t , $i=1,2,\dots,I$; and $P_{jt}^{B|X}$, $P_{kt}^{C|X}$, $P_{lt}^{D|X}$, $j=1,2,\dots,J$ $k=1,2,\dots,K$ $l=1,2,\dots,L$, denote the corresponding conditional probabilities for items B , C and D respectively.

The latent variable is assumed to explain all of the associations among the manifest variables. A goal of traditional LC analysis is to determine the smallest number of latent classes T that is sufficient to explain away (account for) the associations (relationships) observed among the manifest variables.

The technical Appendix for Latent GOLD presents a generalized notation that allows the extension to many variables. In particular, the 4 manifest variables (indicators) A , B , C and D are denoted Y_1 , Y_2 , Y_3 and Y_4 .

The analysis typically begins by fitting the $T=1$ class baseline model (H_0), which specifies mutual independence among the variables. Model H_0 :

$$P_{ijkl} = P_i^A P_j^B P_k^C P_l^D$$

Assuming that this *null* model does not provide an adequate fit to the data, a 1-dimensional LC model with $T=2$ classes is then fitted to the data. This process continues by fitting successive LC models to the data, each time adding another dimension by incrementing the number of classes by 1, until the simplest model is found that provides an adequate fit.

1.2 Assessing Model Fit

Several complimentary approaches are available for assessing the fit of LC models. The most widely used approach utilizes the likelihood ratio chi-squared statistic L^2 , to assess the extent to which ML estimates for the expected cell frequencies, \hat{F}_{ijkl} , differ from the corresponding observed frequencies, f_{ijkl} :

$$L^2 = \sum_{ijkl} f_{ijkl} \ln(\hat{F}_{ijkl} / f_{ijkl})$$

A model fits the data if the value of L^2 is sufficiently low to be attributable to chance (within normal statistical error limits --generally, the .05 level).

The \hat{F}_{ijkl} are obtained using the following 2-step process. First, maximum likelihood (ML) estimates for the model parameters are obtained and substituted into the right side of Equation (1) to obtain ML estimates of the probabilities \hat{p}_{ijkl} . These probability estimates are then summed over the latent classes to obtain estimated probabilities for each cell in the observed table and multiplied by the sample size N to obtain the ML estimates for the expected frequencies:

$$\hat{F}_{ijkl} = N \sum_{t=1}^T \hat{p}_{ijklt}$$

In the case that $\hat{F}_{ijkl} = f_{ijkl}$ for each cell (i,j,k,l) , the model fit will be perfect and L^2 equals zero. To the extent that the value for L^2 exceeds 0, the L^2 measures lack of model fit, quantifying the amount of association (non-independence) that remains unexplained by that model. When N is sufficiently large, L^2 follows a chi-square distribution, and as a general rule¹, the number of degrees of freedom (DF) equals the number of cells in the full multi-way table minus the number of distinct parameters M minus 1. For example, in the case of 4 categorical variables, the number of cells equals $IJKL$ and the number of parameters is:

$$M = T-1 + T[(I-1) + (J-1) + (K-1) + (L-1)]$$

M is obtained by counting the $T-1$ distinct LC probabilities, and for each latent class, the $I-1$ distinct conditional probabilities associated with the categories of variable A , the $J-1$ distinct conditional probabilities associated with B , etc. Since probabilities sum to 1, the probability associated with one category of each variable is redundant (and hence not counted as a *distinct* parameter): it can be obtained as one minus the sum of the others.

In situations involving sparse data, the chi-squared distribution should not be used to compute the p-value because L^2 would not be well approximated. Instead, the bootstrap approach can be used to estimate p (Langeheine, Pannekoek, and Van de Pol, 1996). Sparse data often occurs when the number of observed variables or the number of categories of these variables is large. In such cases the total number of cells in the resulting multi-way frequency table will be large relative to the sample size, resulting in

¹ According to the general rule, if it turns out that $DF < 0$, the model is not identifiable, which means that unique estimates are not available for all parameters. For example, for $I = J = K = L = 2$, $DF = -4$ for $T = 4$, which means that the 4-class model is not identifiable. In some cases however, this general counting rule may yield $DF > 0$, yet the model may still not be identifiable. For example, Goodman (1974a) shows that in this situation of 4 dichotomous variables, the 3-class model is not identifiable despite the fact that the counting rule yields $DF = 1$. See also note 3.

many empty cells. This situation is illustrated below with a data example. Sparse data also results when LC models are extended to include continuous variables, which is illustrated in the last section.

An alternative approach to assessing model fit in the case of sparse data utilizes an information criterion weighting both model fit and parsimony. Such measures, like AIC and BIC, are especially useful in comparing models. The most widely used is the BIC statistic which is defined as: $BIC = L^2 - \ln(N) DF$ (Raftery, 1986). A model with a lower BIC value is preferred over a model with a higher BIC value.²

If the baseline model (H_0) provides an adequate fit to the data, no LC analysis is needed, since there is no association among the variables to be explained. In most cases however, H_0 will not fit the data in which case $L^2(H_0)$ can serve as a baseline measure of the total amount of association in the data. This suggests a 3rd approach for assessing the fit of LC models by comparing the L^2 associated with LC models for which $T > 1$ with the baseline value $L^2(H_0)$ to determine the percent reduction in L^2 . Since the total association in the data may be quantified by $L^2(H_0)$, the percent reduction measure represents the total association explained by the model. This less formal approach can complement the more statistically precise L^2 and BIC approaches.

As an example of how these measures are used, suppose that the L^2 suggests that a 3-class model falls short of providing an adequate fit to some data (say $p = .04$) but explains 90% of the total association. Moreover, suppose a 4-class is the simplest model that fits according to the L^2 statistic but that this model only explains 91% of the association. In this case, it may be that on practical grounds the 3-class model is preferable since it explains almost as much of the total association.

1.3 Example: survey respondent types

We will now consider a first example that illustrates how these tools are used in practice. It is based on the analysis of 4 variables from the 1982 General Social Survey given by McCutcheon (1987) to illustrate how traditional LC modeling can be used to study the different types of survey respondents. Two of the variables ascertain the respondent's opinion regarding (A) the purpose of surveys and (B) how accurate they are, and the others are evaluations made by the interviewer of (C) the respondent's levels of understanding of the survey questions and (D) cooperation shown in answering the questions. McCutcheon initially assumed the existence of 2 latent classes corresponding to 'ideal' and 'less than ideal' types.

The study included separate samples of white and black respondents. Beginning with an analysis of the white sample, McCutcheon later included data from the black sample to illustrate a 2-group LC analysis. We will use these data to introduce the basics of traditional LC modeling and to illustrate several new features that are included in Latent GOLD 3.0. The topics addressed in this section include allowing for specific local dependencies, the usage of LC factor models, and the inclusion of covariates as well as the methodology for making multi-group comparisons. Section 7 contains a detailed tutorial that uses these data to introduce LC Cluster and LC Factor models.

The data used to begin the tutorial are presented in Table 1.

² A more general definition of BIC is based on the log-likelihood (LL) and the number of parameters (npar): $BIC = -2LL + \ln(N) npar$. Both definitions yield the same results when comparing models.

Table 1: Cross-tabulation of 4 Survey Items for Sample of White Respondents

(Source: 1982 U.S. General Social Survey)

A	B	C	D	Observed Frequencies
PURPOSE	ACCURACY	UNDERSTANDING	COOPERATION	Whites
Good	Mostly true	Good	Interested	419
			Cooperative	35
			Impatient/Hostile	2
		Fair, poor	Interested	71
			Cooperative	25
			Impatient/Hostile	5
	Not true	Good	Interested	270
			Cooperative	25
			Impatient/Hostile	4
		Fair, poor	Interested	42
			Cooperative	16
			Impatient/Hostile	5
Depends	Mostly true	Good	Interested	23
			Cooperative	4
			Impatient/Hostile	1
		Fair, poor	Interested	6
			Cooperative	2
			Impatient/Hostile	0
	Not true	Good	Interested	43
			Cooperative	9
			Impatient/Hostile	2
		Fair, poor	Interested	9
			Cooperative	3
			Impatient/Hostile	2
Waste	Mostly true	Good	Interested	26
			Cooperative	3
			Impatient/Hostile	0
		Fair, poor	Interested	1
			Cooperative	2
			Impatient/Hostile	0
	Not true	Good	Interested	85
			Cooperative	23
			Impatient/Hostile	6
		Fair, poor	Interested	13
			Cooperative	12
			Impatient/Hostile	8

Table 2: Results from Various LC Models Fit to Data in Table 1

Model		BIC	L²	DF	p-value	% Reduction in L²(H₀)
<u>Traditional</u>						
H ₀	1-class	51.6	257.3	29	2.0x10 ⁻³⁸	0 %
H _{1C}	2-class	-76.5	79.5	22	2.0x10 ⁻⁸	69.1%
H _{2C}	3-class	-84.3	22.1	15	0.11	91.4%
H _{3C}	4-class	-50.1	6.6	8	0.58	97.4%
<u>Nontraditional</u>						
H _{1C+}	2-class + {CD} direct effect	-129.3	12.6	20	0.89	95.1%
H _{2F}	basic 2-factor	-95.3	11.1	15	0.75	95.7%

Traditional exploratory LC analysis begins by fitting the null model H₀ to the sample of white respondents. Since L²(H₀) = 257.3 with DF = 29 (see Table 2), the amount of association (non-independence) that exists in these data is too large to be explained by chance, so the null model must be rejected (p < .001) in favor of T>1 classes.

Next, we consider McCutcheon's 2-class model (H₁). For this model, the L² is reduced to 79.5³, a 69.1% reduction from the baseline model, but still much too large to be acceptable with DF = 22. Thus, we increment T by 1 and estimate model H_{2C}, the 3-class model. This model provides a further substantial reduction in L² to 22.1 (a 91.5% reduction over the baseline) and also provides an adequate overall fit (p>.05). Table 2 shows that the 4-class LC model provides some further improvement. However, the BIC statistic, which takes parsimony into account, suggests that the 3-class model is preferred over the 4-class model (see Table 2).

Table 3: Parameter Estimates for the 3-Class LC Model by Sample

	Class 1 Ideal	Whites		Class 1	Blacks	
		Class 2 Believers	Class 3 Skeptics		Class 2	Class 3
LC Probabilities	0.62	0.20	0.18	0.49	0.33	0.18
Conditional Probabilities						
(A) PURPOSE						
Good	0.89	0.92	0.16	0.87	0.91	0.19
Depends	0.05	0.07	0.22	0.08	0.04	0.17
Waste	0.06	0.01	0.62	0.05	0.05	0.65
(B) ACCURACY						
Mostly True	0.61	0.65	0.04	0.54	0.65	0.01
Not True	0.39	0.35	0.96	0.46	0.35	0.99
(C) UNDERSTAND						
Good	1.00	0.32	0.75	0.95	0.37	0.68
Fair, poor	0.00	0.68	0.25	0.05	0.63	0.32
(D) COOPERATE						
Interested	0.95	0.69	0.64	0.98	0.56	0.64
Cooperative	0.05	0.26	0.26	0.01	0.37	0.25
Impatient/ Hostile	0.00	0.05	0.10	0.00	0.07	0.11

³ This value differs from the value 79.3 reported in McCutcheon (1984) because our models include a Bayes constant set equal to 1. For further information on Bayes constants see the Technical Appendix.

The parameter estimates obtained from the 3-class model are given in the left-most portion of Table 3. The classes are ordered from largest to smallest. Overall, 62% are estimated to be in class 1, 20% in class 2 and the remaining 18% in class 3. Analogous to factor analysis where names are assigned to the factors based upon an examination of the ‘factor loadings’, names may be assigned to the latent classes based upon the estimated conditional probabilities. Like factor loadings, the conditional probabilities provide the measurement *structure* that defines the latent classes.

McCutcheon assigned the name ‘Ideal’ to latent class 1, reasoning as follows: The first class corresponds most closely to our anticipated ideal respondents.

Nearly 9 of 10 in this class believed that surveys ‘usually serve a good purpose;’ 3 of 5 expressed a belief that surveys are either ‘almost always right’ or ‘right most of the time’; 19 of 20 were evaluated by the interviewer as ‘friendly and interested’ during the interview; and nearly all were evaluated by the interviewer as having a good understanding of the survey questions.

He named the other classes ‘Believers’ and ‘Skeptics’ based on the interpretations of the corresponding conditional probabilities for those classes.

1.4 Testing the Significance of Effects

The next step in a traditional LC analysis is to delete from the model any variable that does not exhibit a significant difference between the classes. For example, to test whether to delete variable *A* from a *T*-class model, one would test the null hypothesis that the distribution over the *I* categories of *A* is identical within each class *t*:

$$p_{i1}^{AX} = p_{i2}^{AX} = \dots = p_{iT}^{AX} \text{ for } i=1,2,\dots,I$$

In order to implement this test we make use of the relationship between the conditional response probabilities and the log-linear parameters (see, e.g., Haberman 1979; Formann, 1992; or Heinen, 1996):

$$p_{it}^{AX} = \frac{\exp(I_i^A + I_{it}^{AX})}{\sum_{i'=1}^I \exp(I_{i'}^A + I_{i't}^{AX})}$$

Standard log-linear modeling techniques can then be used to test the null hypothesis, re-expressed in terms of the log-linear parameters associated with the *AX* relationship:

$$I_{i1}^{AX} = I_{i2}^{AX} = \dots = I_{iT}^{AX} = 0 \text{ for } i=1,2,\dots,I$$

Table 4: Comparison of L² difference and Wald statistic for testing significance of variables in the 3-class LC model

Variable	DF	L ² difference		Wald test	
		ΔL^2	p-value	Wald	p-value
(A) PURPOSE	4	145.3	2.1×10^{-30}	29.6	6.0×10^{-6}
(B) ACCURACY	2	125.4	5.8×10^{-28}	8.4	0.015
(C) UNDERSTAND	2	61.3	4.9×10^{-14}	7.4	0.024
(D) COOPERATE	4	101.1	5.6×10^{-21}	19.0	0.0008

Table 4 provides the results of the significance test for each of the 4 variables in our 3-class model. The left-most portion of the table provides the results based on the L^2 difference test, where ΔL^2 is computed as the difference between the L^2 statistics obtained under the *restricted* and *unrestricted* 3-class models respectively. The right-most portion of the table gives the corresponding results obtained using the Wald statistic which is easier to compute but uniformly less powerful than ΔL^2 statistic. Under the assumption that the unrestricted model is true, both statistics are distributed asymptotically as chi-square with $DF = (I-1)*(T-1)$, where I denotes the number of categories in the nominal variable. Table 4 shows that each of the 4 variables included in the model is significant.

1.5 Classification

The final step in a traditional LC analysis is to use the results of the model to classify cases into the appropriate latent classes. For any given response pattern (i,j,k,l) , estimates for the posterior membership probabilities can be obtained using Bayes theorem as follows:

$$\hat{p}_{ijkl}^{X|ABCD} = \frac{\hat{P}_{ijkl}^{ABCDX}}{\sum_{t=1}^T \hat{P}_{ijkl}^{ABCDX}}, t=1,2,\dots, T \quad (2)$$

where the numerator and denominator in Equation (2) are obtained by substituting the model parameter estimates in place of the corresponding parameters in Equation (1).

Magidson and Vermunt (2001) refer to this kind of model as a LC *cluster* model because the goal of classification into T homogeneous groups is identical to that of cluster analysis. In contrast to an ad hoc measure of distance used in cluster analysis to define homogeneity, LC analysis defines homogeneity in terms of probabilities. As indicated by Equation (1), cases in the same latent class are similar to each other because their responses are generated by the same probability distribution.

Table 5: ML Estimates for the Posterior Membership Probabilities under model H_{2C}

A	B	C	D	Posterior Probabilities		
				Ideal	Believers	Skeptics
PURPOSE	ACCURACY	UNDERSTANDING	COOPERATION			
Good	Mostly true	Good	Interested	0.92	0.08	0.00
			Cooperative	0.64	0.35	0.01
			Impatient/Hostile	0.02	0.94	0.04
		Fair, poor	Interested	0.02	0.97	0.00
			Cooperative	0.00	0.99	0.00
			Impatient/Hostile	0.00	0.99	0.01
	Not true	Good	Interested	0.88	0.06	0.06
			Cooperative	0.52	0.24	0.24
			Impatient/Hostile	0.01	0.35	0.64
		Fair, poor	Interested	0.02	0.85	0.12
			Cooperative	0.00	0.86	0.13
			Impatient/Hostile	0.00	0.78	0.22
Depends	Mostly true	Good	Interested	0.87	0.10	0.04
			Cooperative	0.49	0.36	0.15
			Impatient/Hostile	0.01	0.56	0.44
		Fair, poor	Interested	0.02	0.93	0.06
			Cooperative	0.00	0.94	0.06
			Impatient/Hostile	0.00	0.89	0.11
	Not true	Good	Interested	0.37	0.04	0.59
			Cooperative	0.08	0.05	0.87

			Impatient/Hostile	0.00	0.03	0.97	
		Fair, poor	Interested	0.01	0.28	0.72	
			Cooperative	0.00	0.27	0.73	
Waste	Mostly true	Good	Impatient/Hostile	0.00	0.16	0.84	
			Interested	0.88	0.02	0.10	
			Cooperative	0.53	0.07	0.41	
			Fair, poor	Impatient/Hostile	0.01	0.08	0.91
				Interested	0.05	0.50	0.44
				Cooperative	0.01	0.51	0.48
		Not true	Good	Impatient/Hostile	0.00	0.36	0.64
	Interested			0.20	0.00	0.80	
	Cooperative			0.03	0.00	0.96	
			Fair, poor	Impatient/Hostile	0.00	0.00	1.00
				Interested	0.00	0.03	0.97
				Cooperative	0.00	0.02	0.97
			Impatient/Hostile	0.00	0.01	0.99	

Cases are then assigned to the class for which the posterior probability is highest (i.e., the modal class). The posterior membership probabilities for our 3-class LC model are given in Table 5, with the modal probabilities highlighted.

1.6 Graphical Displays

Since for any given response pattern (i,j,k,l) the T class membership probabilities sum to 1, only $T-1$ such probabilities are required as the probability of belonging to the remaining class can be obtained from the others. Hence, the class membership probabilities $\hat{p}_{ijkl}^{X|ABCD}$ can be used to position each response pattern in $T-1$ dimensional space, and for $T=3$, various 2-dimensional barycentric coordinate displays can be produced.

Rather than plotting every one of the many response patterns, instructive plots of the kind used in correspondence analysis can be produced, where points are plotted for each category of each variable as well as other meaningful aggregations of these probabilities (Magidson and Vermunt, 2001) (See Figure 1)

Figure 1: Barycentric coordinate display for 3-class model

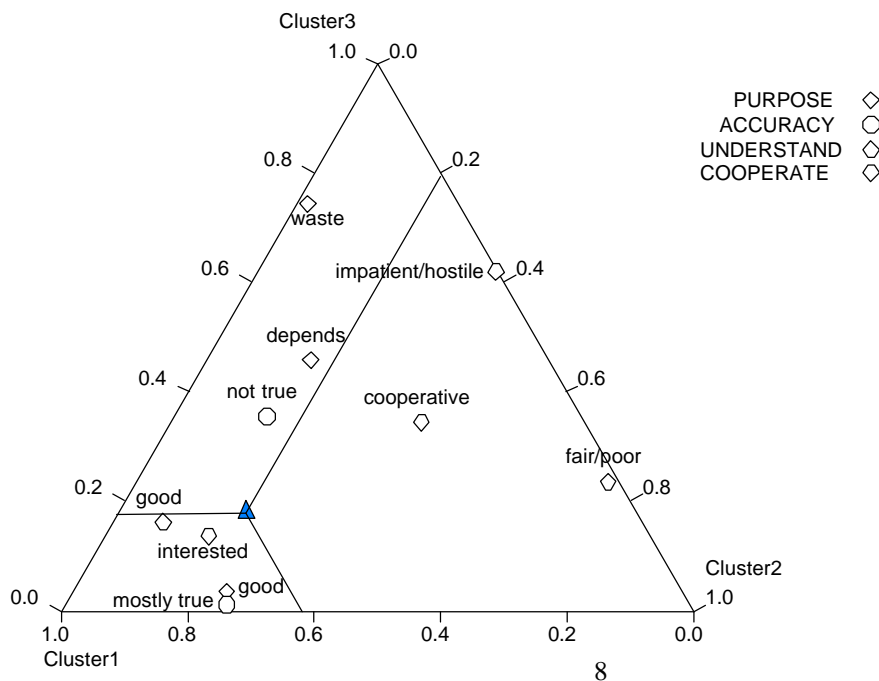


Figure 1 depicts the corresponding barycentric coordinate display under the 3-class LC model. Points are plotted for each category of each of the 4 variables in our example. Since these points contain information equivalent to the LC parameter estimates (Van der Heijden, Gilula and Van der Ark, 1999), this type of plot provides a graphical alternative to the traditional tabular display of parameter estimates and can yield new insights into data. Also displayed in Figure 2 are 2 additional aggregations associated with the response categories UNDERSTANDING = ‘good’ and ‘fair,poor’ ($k=1,2$) among those for whom COOPERATION = ‘Hostile/Impatient’ ($l=3$).

The horizontal dimension of the plot corresponds to differences between McCutcheon’s ‘ideal’ and ‘believer’ types (latent classes 1 and 2). We see that the categories of the variable *C* tend to spread out along this dimension. Respondents showing ‘good’ understanding are most likely to belong to the ideal class (the corresponding symbol is plotted closest to the lower left vertex that represents class 1) while those showing only ‘fair or poor’ understanding are plotted closest to the lower right vertex which represents class 2.

Differences along the vertical dimension of the plot are best shown by the categories of *A* and *B*. For example, respondents agreeing that the purpose of surveys is ‘good’ are plotted close to the lower left (class 1) vertex. Those who say ‘it depends’ are plotted somewhat midway between the class 1 and class 3 (top) vertex. Those who say ‘it’s a waste of time and \$’ are most likely to be in class 3 and are positioned near the top vertex. The fact that the positioning of categories for both *A* and *B* spread out over the vertical dimension suggests a high degree of association between these variables. In contrast, the categories of *C* are spread over the horizontal dimension, suggesting that the association between *C* and the 2 variables *A* and *B* is close to nil.

The categories of the variable *D* form an interesting diagonal pattern. Respondents showing they are ‘interested’ in the questions are most likely to be in class 1 (‘ideal’) while those who are only ‘cooperative’ or exhibit ‘impatience/hostility’ are plotted closer to classes 2 and 3. This suggests the hypothesis that impatience and hostility may be come about for either of 2 distinct reasons – 1) disagreement that surveys are accurate and serve a good purpose (indicated by the vertical dimension of the plot) and/or 2) lack of understanding (indicated by the horizontal dimension).

The additional points plotted deal with the relationship between variables *C* and *D*. The positioning of these points suggest that among impatient/hostile respondents, those who show good understanding of the questions tends to be more in class 3 while those whose understanding is Fair/Poor tend to be about equally likely to be in classes 2 or 3.

2. NONTRADITIONAL LC MODELING

Rejection of a traditional T -class LC model for lack of fit means that the local independence assumption does not hold with T classes. In such cases the traditional LC model fitting strategy is to fit a $T+1$ class model to the data. In both of our examples, theory supported a 2-class model but since this model failed to provide an adequate fit we formulated a 3-class model. In this section we consider some alternative strategies for modifying a model. In both cases we will see the nontraditional alternatives lead to models that are more parsimonious than traditional models as well as models that are more congruent with our initial hypotheses. The alternatives considered are:

- 1] adding 1 or more direct effects
- 2] deleting 1 or more items
- 3] increasing the number of latent variables

Alternative #1 is to include ‘direct effect’ parameters in the model (Hagenaars, 1988) that account for the residual association between the observed variables that is responsible for the local dependence. This approach is particularly useful when some external factor, unrelated to the latent variable, creates irrelevant association between two variables. Examples of such external factors include similar question wording used in two survey items, as well as two raters using the same incorrect criterion in evaluating slides.

Alternative #2 also deals with the situation where 2 variables are responsible for some local dependency. In such cases rather than add a direct effect between two variables, it may make more sense to eliminate the dependency by simply deleting one of the two items. This variable reduction strategy is especially useful in situations where there are *many* redundant variables.

Alternative #3 is especially useful when a group of several variables account for a local dependency. Magidson and Vermunt (2001) show that by increasing the dimensionality through the addition of latent variables rather than latent classes, the resulting *LC factor* model often fits data substantially better than the traditional LC cluster models having the same number of parameters. In addition, LC factor models are identified in some situations where the traditional LC model is not⁴.

In the next section we introduce a diagnostic statistic called the *bivariate residual* (BVR) and illustrate its use to develop some nontraditional alternative models for our two data examples. The BVR helps pinpoint those *bivariate* relationships⁵ that fail to be adequately explained by the LC model and can help determine which of the 3 alternative strategies to employ. We will see that even in situations where the L^2 statistic reports that the model provides an adequate *overall* fit, the fit in one or more 2-way tables may not be adequate and may indicate a flaw or weakness in the model.

2.1 Bivariate Residuals and Direct effects

A formal measure of the extent to which the observed association between 2 variables is reproduced by a model is given by the BVR statistic (Vermunt and Magidson, 2001). Each BVR corresponds to a Pearson X^2 statistic (divided by the degrees of freedom) where the observed frequencies in a 2-way cross-tabulation of the variables are contrasted with those expected counts estimated under the corresponding LC model⁶. A BVR value substantially larger than 1 suggests that the model falls somewhat short of explaining the association in the corresponding 2-way table.

Table 9: Values for Bivariate Residuals obtained Under Various Models for the Sample of White Respondents

	Model					
2-way table	H ₀	H ₁	H _{2C}	H _{3C}	H _{2C+}	H _{2F}
{AB}	61.6	0.1	0.1	0.0	0.0	0.0
{AC}	0.5	0.7	0.1	0.0	0.2	0.0
{AD}	10.6	0.0	0.1	0.0	0.2	0.1
{BC}	0.3	1.1	0.0	0.0	0.0	0.0

⁴ For example, with 4 dichotomous variables, a LC 2-factor model (comprised of 4 latent classes) is identified whereas a traditional 3-class model is not (Goodman, 1974a).

⁵ Traditional factor analysis, through the assumption of multivariate normality, limits its focus to bivariate relationships (i.e., the correlations), since higher-order relationships are assumed not to exist. In contrast, LC models do not make strict distributional assumptions, and hence attempt to explain higher-order associations as well. Nevertheless, the 2-way (bivariate) associations are generally the most prominent, and the ability to pinpoint specific 2-way tables where lack of fit may be concentrated can be useful in suggesting alternative models.

⁶ These residuals are similar to Lagrangian statistics. A difference is that they are limited information fit measures: dependencies with parameters corresponding to other items are not taken into account.

{BD}	8.6	0.4	0.3	0.2	0.2	0.4
{CD}	43.4	32.3	2.4	0.0	0.0	0.2

Table 9 reports BVRs for each variable pair under each of several models. Since model H_0 corresponds to the model of mutual independence, each BVR for this model provides a measure of the overall association in the corresponding observed 2-way table; that is, each BVR equals the usual Pearson X^2 statistic used to test for independence in the corresponding 2-way table divided by the degrees of freedom. The results show that except for the non-significant relationships in the {AC} and {BC} tables, all of the remaining BVRs are quite large, attesting to several significant associations (local dependencies) that exist among these variables. The BVR is especially large for {AB} and for {CD}. For example, in Table {CD}, a Pearson chi-square test confirms that the observed relationship is highly significant ($X^2 = 86.8$, 2df, $p < .001$; $BVR = 86.8/2 = 43.4$).

Under the 2-class model (H_1), note that the BVRs are all near or less than 1 except for one very large value of 32.3 for {CD}. This suggests that the overall lack of fit for this model can be traced to this single large BVR. The traditional way to account for the lack of fit is by adding another latent class. However, Table 9 shows that even after the addition of a 3rd class, the BVR for {CD} under the 3-class model H_{2C} remains unacceptably high ($BVR = 2.4$). Although the inclusion of the 3rd class does add a second dimension that causes the *overall* fit to be adequate, it is not until we add a 4th class (model H_{3C}) that *all* BVRs are at acceptable levels.

We will now consider the alternative approach of adding a ‘direct effect’ to the model to account for the residual correlation. Then, we consider use of the 2-factor LC model and further explore the differences between the 3 and 4-class models.

Table 2 shows that the inclusion of the direct effect parameters, I_{kl}^{CD} , in the 2-class model provides a good fit to the data. However, under this model, the parameter measuring the contribution of C to the latent classes is no longer significant and therefore C can be deleted from the LC model completely. As this amounts to deleting an association simply because it could not be explained by 2 latent classes, alternative #1 does not provide a desirable solution here. In other situations, with other data the inclusion of direct effects has been shown to be extremely beneficial in improving the model (for example, see the diabetes example in Vermunt and Magidson 2002a).

2.2 LC Factor Models

Next we consider alternative approach #3 where we utilize LC factor models to include more than one latent variable in the model. LC factor models were proposed as a general alternative to the traditional exploratory LC modeling by Magidson and Vermunt (2001). For both examples, the results (given in Table 2 and Table 7) show that a 2-factor model is preferable to the other models. We shall see that the 2-factor model is actually a restricted 4-class model. In both cases the fit is almost as good as the (unrestricted) 4-class solution, but is more parsimonious and parameterized in a manner that allows easier interpretation of the results.

LC factor models were initially proposed by Goodman (Goodman, 1974b) in the context of confirmatory latent class analysis. Certain traditional LC models containing 4 or more classes can be interpreted in terms of 2 or more component latent variables by treating those components as a joint variable (see e.g., McCutcheon 1987; Hagenaars 1990). For example, a latent variable X consisting of $T=4$ classes can be re-expressed in terms of 2 dichotomous latent variables $V = \{1,2\}$ and $W = \{1, 2\}$ using the following correspondence:

	W=1	W=2
V=1	X = 1	X = 2
V=2	X = 3	X = 4

Thus, $X=1$ corresponds with $V=1$ and $W=1$, $X=2$ with $V=1$ and $W=2$, $X=3$ with $V=2$ and $W=1$, and $X=4$ with $V=2$ and $W=2$.

Formally, for 4 nominal variables, the 4-class LC model can be re-parameterized as a LC factor model with two dichotomous latent variables as follows:

$$p_{ijklrs} = p_{rs}^{VW} p_{ijklrs}^{ABCD|VW} = p_{rs}^{VW} p_{irs}^{A|VW} p_{jrs}^{B|VW} p_{krs}^{C|VW} p_{lrs}^{D|VW}$$

In Magidson and Vermunt (2001), we discussed various restricted factor models. We used the term *basic* LC factor models to refer to certain LC models that contain 2 or more dichotomous latent variables that are mutually independent of each other and that exclude higher-order interactions from the conditional response probabilities. Such a model is analogous to the approach of traditional factor analysis where multiple latent variables are used to model multidimensional relationships among manifest variables. This type of LC factor model is the default factor model in Latent GOLD.

It turns out that by formulating the model in terms of R mutually independent, dichotomous latent factors, the basic LC factor model has the same number of distinct parameters as a traditional LC model with $R+1$ classes. That is, the LC factor parameterization allows specification of a 2^R -class model with the same number of parameters as a traditional LC model with only $R+1$ classes! This offers a great advantage in parsimony over the traditional T -class model as the number of parameters is greatly reduced by natural restrictions.

Table 2 shows that the basic 2-factor model (model H_{2F}) provides an excellent fit and is preferred over any of the LC cluster models according to the BIC statistic. In addition, this model explains all bivariate relationships in the data (see Table 9). Below we will interpret the results from this model in conjunction with a more extensive multi-group analysis that includes both the white and black sample.

2.3 Multi-group Models

Multi-group LC models can be used to compare models across groups. A completely unrestricted multi-group LC model, referred to by Clogg and Goodman (1984) as the model of complete heterogeneity, is equivalent to the estimation of a separate T -class LC model for each group. The fit of such a model can be obtained by simply summing the L^2 values (and corresponding degrees of freedom) for the corresponding models in each group.

Let G denote a categorical variable representing membership in group g . The model of *complete heterogeneity* is expressed as (model M_{2C}):

$$p_{ijkl|g}^{ABCDX|G} = p_{i|g}^{X|G} p_{i|g}^{A|X,G} p_{j|g}^{B|X,G} p_{k|g}^{C|X,G} p_{l|g}^{D|X,G}$$

TABLE 11: Results from Various LC Models Fit to Data in Table 2 – Black Respondents

Model	BIC	L²	DF	p-value	% Reduction in L²(H₀)
Traditional					
H' ₀ 1-class	-64.9	112.1	29	1.0x10 ⁻¹¹	0.0%
H' ₁ 2-class	-77.4	56.9	22	.00006	49.2%
H' _{2C} 3-class	-73.2	18.3	15	.25	83.7%
H' _{3C} 4-class	-39.4	9.4	8	.31	91.6%
Nontraditional					
H' _{1C+} 2-class + {CD} direct effect	-106.8	15.2	20	.77	86.4%
H' _{2F} basic 2-factor	-80.0	11.5	15	.72	89.7%

Table 11 provides the results of repeating our example 1 analyses for the sample of *black* respondents. These results turn out to be very similar to those obtained for the white respondents (recall Table 2). As in our analysis for the white sample we again reject the 1 and 2-class models in favor of 3 classes in order to obtain a model that provides an overall fit to the data that is adequate. The right-most portion of Table 3 presents the parameter estimates obtained from the 3-class model (model H'_{2C}) as applied to the sample of blacks. As in our earlier analysis the classes are ordered from largest to smallest.

In comparing results across these two groups, it is important to be able to interpret the 3 classes obtained from the black respondents as representing the same latent constructs ('ideal', 'believers' and 'skeptics') as in our analysis of the white respondents. Otherwise, any between-group comparisons would be like comparing 'apples' with 'oranges'. While it is tempting to interpret class 1 for both samples as representing the 'ideal' respondents, this is not appropriate without first restricting the measurement portion of the models (the conditional probabilities) to be equal. These restrictions are accomplished using the model of *partial* homogeneity (model M_{2CR}):

$$p_{ijkl|g}^{ABCDX|G} = p_{i|g}^{X|G} p_{it}^{A|X} p_{jt}^{B|X} p_{kt}^{C|X} p_{lt}^{D|X} \quad (3)$$

Table 12: Parameter Estimates for the 3-Class LC model of partial homogeneity (model M_{2CR}) and the corresponding LC 2-factor model M_{2FR}

	Class 1	Class 2	Class 3	Factor V		Factor W	
LC Probabilities	Ideal	Believers	Skeptics	Level 1	Level 2	Level 1	Level 2
Whites	0.68	0.15	0.17	0.81	0.19	0.85	0.15
Blacks	0.51	0.30	0.19	0.79	0.21	0.70	0.30
Conditional Probabilities							
PURPOSE							
Good	0.89	0.90	0.16	0.90	0.20	0.82	0.83
Depends	0.06	0.06	0.21	0.06	0.21	0.09	0.07
Waste	0.05	0.04	0.63	0.05	0.59	0.09	0.10
ACCURACY							
Mostly True	0.60	0.64	0.01	0.63	0.02	0.41	0.45
Not True	0.40	0.36	0.99	0.37	0.98	0.59	0.55
UNDERSTAND							
good	0.94	0.32	0.74	0.85	0.80	0.91	0.24
Fair, poor	0.06	0.68	0.26	0.15	0.20	0.09	0.76
COOPERATE							
Interested	0.95	0.57	0.65	0.90	0.68	0.91	0.48
Cooperative	0.05	0.35	0.25	0.09	0.25	0.08	0.41
Impatient/ Hostile	0.00	0.08	0.10	0.01	0.07	0.01	0.11

Estimates from this model are given in the left-most portion of Table 12. Table 13 compares the fit of the unrestricted model M_{2C} and restricted model M_{2CR} . The L^2 statistic can be used to test the restrictions made under model M_{2CR} . Since $L^2 = 9.0$ with 18 DF is *not* significant, we are free to use this restricted model for our group comparisons.

Table 13: Results from Various Multi-Group LC Cluster and Factor Models

Model		BIC	L^2	DF	p-value	% Reduction in $L^2(H_0)$
Traditional						
M_0	1-class	-74.1	400.0	64	4.3×10^{-50}	0%
M_1	2-class	-245.4	169.5	56	2.4×10^{-13}	57.6%
M_{2C}	3-class unrestricted (complete heterogeneity)	-181.8	40.4	30	.10	89.9%
M_{2CR}	3-class restricted (partial homogeneity)	-306.2	49.4	48	.42	87.7%
M_{2CRR}	3-class restricted (complete homogeneity)	-297.1	73.3	50	.02	81.7%
M_{3CR}	4-class restricted (partial homogeneity)	-269.4	27.0	40	.94	93.3%
Nontraditional						
M_{2F}	basic 2-factor unrestricted	-199.6	22.6	30	.83	94.4%
M_{2FR}	basic 2-factor restricted	-324.3	31.3	48	.97	92.2%

The model of complete homogeneity (model M_{2CRR}) imposes the further restriction that the latent class probabilities across the groups are identical: $p_{t|1}^{X|G} = p_{t|2}^{X|G}$, for $t = 1, 2, 3$. Since these restrictions yield a significant increase in L^2 , we reject the model of complete homogeneity in favor of the model of partial homogeneity and conclude that there are significant differences in latent class membership between the white and black samples.

Table 13 also includes results obtained from the LC factor model counterparts to the models of complete heterogeneity and partial heterogeneity. Since these models contain 2 dichotomous and independent factors, they contain exactly the same number of parameters as the 3-class models M_{2C} and M_{2CR} . Table 13 shows that these models fit better than the corresponding LC cluster models according to the BIC criteria. Table 14 shows that the factor model also contains smaller BVRs than the LC cluster counterpart.

Table 14: Comparison of BVR's for multi-group cluster and factor models

2-way table	Model	
	M_{2CR}	M_{2FR}
{CD}	2.8	0.0
{GB}	1.6	1.5
{GC}	5.0	0.0
{GD}	1.2	0.0

The parameter estimates from the 2-factor model M_{2FR} are presented in the right-most portion of Table 12. Note that variable D is strongly related to both factors V and W . That is, respondents at level 1 of each factor have a higher probability (.90 or .91) of being 'interested' than those at level 2. Variables A and B relate only to factor V and variable C relates only to factor W . That is, for factor V , those at level 1 are substantially more likely to agree that surveys serve a good purpose and are accurate than those at level 2, but the 2 levels are about equal in showing a good understanding of the questions. For factor W , level 1 shows good understanding while level 2 does not.

Moreover, Table 12 shows that group differences exist primarily with respect to factor 2 (observed group differences on factor V are not significant). Black respondents are twice as likely as whites to be at level 2 of factor 2 (30% vs. 15%). These results allow us to formulate a more rigorous test of our earlier hypothesis that cooperation may be due to two separate factors – one associated with the belief that surveys serve a good purpose and are accurate (as assessed by LC factor 1), the second related to understanding the questions (as assessed by LC factor 2).

2.4 Ordinal Indicators

Thus far we have treated the trichotomous variables COOPERATE (A) and PURPOSE (C) as nominal. Alternatively, they can be treated as ordinal, which serves to simplify the model by reducing the number of parameters. The most straightforward approach is to restrict the log-linear parameters by using uniform scores v_i^A and v_k^C for the categories of A and C , implying the following constraints: $I_{ir}^{AV} = I_r^{AV} v_i^A$ and $I_{is}^{AW} = I_s^{AW} v_i^A$ (see, e.g., Formann, 1992; or Heinen, 1996).

The use of these restrictions in our example increased the L^2 by very little, indicating that variables A and C may in fact be treated as ordinal. In the next section, we present the results of a modified 2-factor model where variables A and C are treated as ordinal.

2.5 Covariates

The parameters in the traditional LC model consist of unconditional and conditional probabilities. The conditional probabilities comprise the measurement portion of the model. They characterize the distribution among the observed variables (indicators)

conditional on the latent classes. The *unconditional* probabilities describe the distribution of the latent variable(s). In order to obtain improved description/ prediction of the latent variable(s), a multinomial logit model is used to express these probabilities as a function of one or more exogenous variables \mathbf{Z} called covariates (Dayton and McReady, 1988).

The multi-group model described in the previous section is an example of the use of a single nominal covariate ($\mathbf{Z} = G$). For example, the term $p_{tg}^{X|G}$ appearing Equation (3) can be expressed as:

$$p_{tg}^{X|G} = \frac{\exp(\mathbf{g}_t^X + \mathbf{g}_{tg}^{XG})}{\sum_{t=1}^T \exp(\mathbf{g}_t^X + \mathbf{g}_{tg}^{XG})}$$

While the latent variable(s) explain all of the associations among the indicators, associations between the covariates are *not* explained by the latent variables. This is what distinguishes the indicators from the covariates.

Regarding the interpretation of the 3-class solution, McCutcheon questioned whether some of the difference in latent class membership between black and white respondents might be explained by education, a question that falls outside the scope of traditional LC modeling. We address this question below by including E :EDUCATION as a second covariate in the 2-factor model -- $\mathbf{Z} = (G,E)$.

Table 15: Parameter Estimates for the 2-Factor Restricted Multi-Group LC Model with Covariates

	gammas	
Intercept	V	W
1	0.56	0.61
2	-0.56	-0.61
Covariates	V	W
G: Group		
WHITE	0	-0.20
BLACK	0	0.20
E:Years of Education		
<8	0	2.19
8-10	0	0.97
11	0	0.08
12	0	-0.34
13-15	0	-1.01
16-20	0	-1.89
Indicator Variables	lambdas	
	V	W
A: PURPOSE	2.26	0
B: ACCURACY		
mostly true	-1.34	0.00
not true	1.34	0.00
C:UNDERSTANDING		
Good	0	-5.14

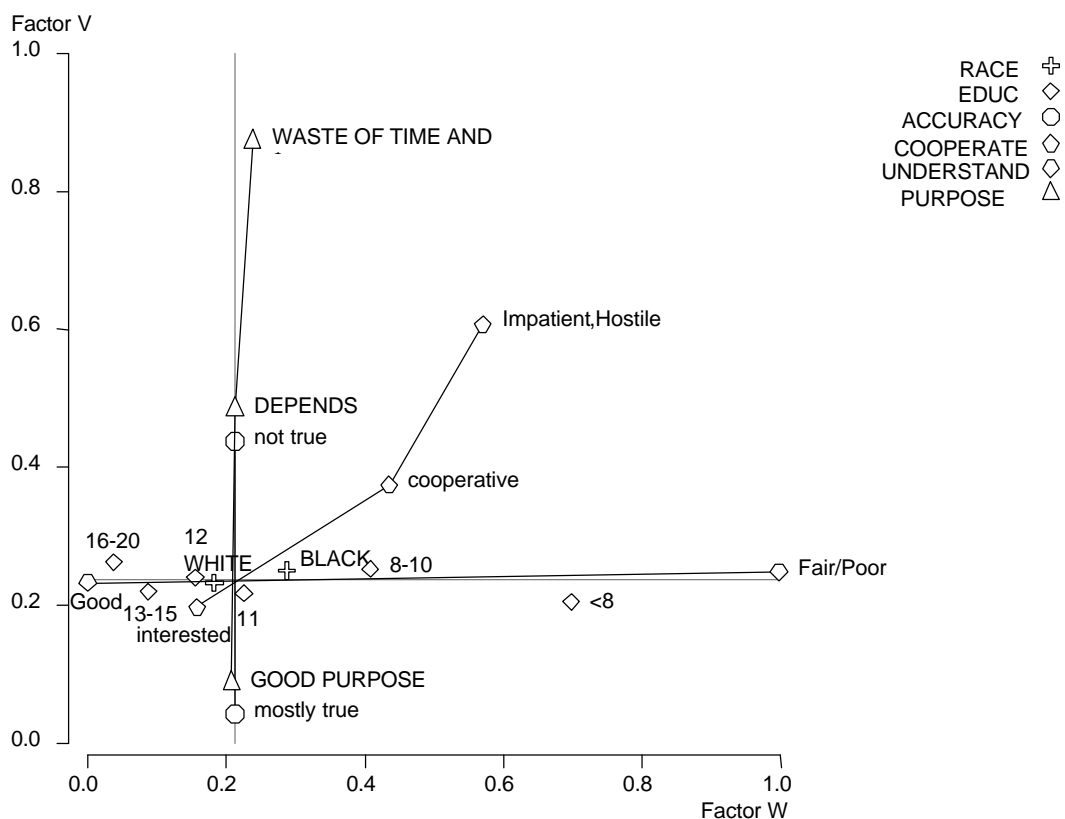
Fair/Poor 0 5.14

D: COOPERATION 0.98 1.26

The model provides a good fit to the data. The results indicate that the effect of education *does* explain most, but not all, of the group effect on factor *W*. The log-linear parameter estimates are given in Table 15. Nonsignificant estimates were set to zero. The multinomial model used for the covariates was:

$$p_{rsge}^{VW|GE} = \frac{\exp(\mathbf{g}_r^V + \mathbf{g}_s^W + \mathbf{g}_{gs}^{GW} + \mathbf{g}_{es}^{EW})}{\sum_{r=1,s=1}^{R,S} \exp(\mathbf{g}_r^V + \mathbf{g}_s^W + \mathbf{g}_{gs}^{GW} + \mathbf{g}_{es}^{EW})}$$

Figure 2: Bi-plot for 2-factor model with covariates



The results for this 2-factor restricted multi-group model are also displayed in the bi-plot display (Magidson and Vermunt, 2001) given in Figure 2. Like the barycentric coordinate display in Figure 2 we see that the horizontal axis, corresponding to factor *W*, is associated with UNDERSTANDING. Overall, respondents having a good understanding are highly likely to be at level 1 of factor *W* while those with a Fair/Poor understanding are highly likely to be in level 2. The figure makes it clear that education is much more related to this factor than race. The vertical dimension is highly related to PURPOSE. Figure 2 shows more clearly than Figure 1 that COOPERATION is related to both factors. In particular, those rated as ‘Impatient/hostile’ tend to include 2 different types of respondents – those whose understanding is fair/poor as well as those who view the purpose of surveys as a ‘waste of time and \$’.

REFERENCES

- Agresti, A. (2002). *Categorical Data Analysis*. Second Edition. New York: Wiley.
- Andrews, R.L., Ansari, A., and Currim, I.S. (2002). Hierarchical bayes versus finite mixture conjoint analysis models: A comparison of fit, prediction, and partworth recovery, *Journal of Marketing Research*, 39, 87-98.
- Banfield, J.D., and Raftery, A.E. (1993). Model-based Gaussian and non-Gaussian clustering. *Biometrics*, 49, 803-821.
- Böckenholt, U. (2002). Comparison and choice: analyzing discrete preference data by latent class scaling models. J.A. Hagenaars and A.L. McCutcheon (eds.), *Applied Latent Class Analysis*. Cambridge University Press.
- Clogg, C.C., and Goodman, L.A. (1984). Latent structure analysis of a set of multidimensional contingency tables. *Journal of the American Statistical Association*, 79, 762-771.
- Croon, M.A. (2002) Ordering the classes. J.A. Hagenaars and A.L. McCutcheon (eds.), *Applied Latent Class Analysis*. Cambridge University Press.
- Dayton, C.M., and Macready, G.B. (1988). Concomitant-variable latent-class models. *Journal of the American Statistical Association*, 83, 173-178.
- Dayton, C.M. (1998). *Latent Class Scaling Models*. Thousand Oakes: Sage Publications.
- Dillon, W.R., and Kumar, A. (1994). Latent structure and other mixture models in marketing: An integrative survey and overview. R.P. Bagozzi (ed.), *Advanced Methods of Marketing Research*, 352-388, Cambridge: Blackwell Publishers.
- Everitt, B.S., and Hand, D.J. (1981). *Finite Mixture Distributions*. London: Chapman and Hall.
- Fraley, C. and Raftery, A.E. (1998). *How many clusters? Which clustering method? Answers via model-based cluster analysis*. Department of Statistics, University of Washington: Technical Report no. 329.:
- Formann, A.K. (1992). Linear logistic latent class analysis for polytomous data. *Journal of the American Statistical Association*, 87, 476-486.
- Goodman, L.A. (1974a). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, 61, 215-231.
- Goodman, L.A. (1974b). The analysis of systems of qualitative variables when some of the variables are unobservable. Part I: A modified latent structure approach, *American Journal of Sociology*, 79, 1179-1259.
- Haberman, S.J. (1979). *Analysis of Qualitative Data, Vol 2, New Developments*. New York: Academic Press.

- Hagenaars, J.A. (1988). Latent structure models with direct effects between indicators: local dependence models. *Sociological Methods and Research*, 16, 379-405.
- Hagenaars, J.A. (1990). *Categorical Longitudinal Data – Loglinear Analysis of Panel, Trend and Cohort Data*. Newbury Park: Sage.
- Heinen, T. (1996). *Latent Class and Discrete Latent Trait Models: Similarities and Differences*. Thousand Oakes: Sage Publications.
- Hunt, L., and Jorgensen. M. (1999). Mixture model clustering using the MULTIMIX program. *Australian and New Zealand Journal of Statistics*, 41:153-172.
- Jedidi, K., Jagpal, H.S., and DeSarbo, W.S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science*, 16, 39-59.
- Landis, J.R., and Koch, G.G. (1977). The measurement of observer agreement for categorical data, *Biometrics*, 33, 159-174.
- Langeheine, R., and Van de Pol, F. (1994). Discrete-time mixed Markov latent class models. A. Dale and R.B. Davies (eds.), *Analyzing social and political change: a casebook of methods*, 171-197.
- Langeheine, R., Pannekoek, J., Van de Pol, F. (1996). Bootstrapping goodness-of-fit measures in categorical data analysis. *Sociological Methods and Research*, 24, 492-516.
- Lazarsfeld, P.F., and Henry, N.W. (1968). *Latent Structure Analysis*. Boston: Houghton Mill.
- Magidson, J., and Vermunt, J.K. (2001). Latent class factor and cluster models, bi-plots and related graphical displays, *Sociological Methodology*, 31, 223-264.
- Magidson J. and Vermunt, J.K. (2002). Latent class models for clustering: A comparison with K-means. *Canadian Journal of Marketing Research*, 20, 37-44.
- McCutcheon, A.L. (1987). *Latent Class Analysis*, Sage University Paper. Newbury Park: Sage Publications.
- McGrath, K., and Waterton, J. (1986). *British Social Attitudes, 1983-1986 Panel Survey*. London: Social and Community Planning Research, Technical Report.
- McLachlan, G.J., and Peel, D. (2000). *Finite Mixture Models*. New York: Wiley.
- Rindskopf, R., and Rindskopf, W. (1986). The value of latent class analysis in medical diagnosis. *Statistics in Medicine*, 5, 21-27.
- Uebersax JS. (1999) Probit latent class analysis with dichotomous or ordered category measures: conditional independence/dependence models. *Applied Psychological Measurement*, 23, 283-297.
- Uebersax JS, and Grove, WM. (1990). Latent class analysis of diagnostic agreement. *Statistics in Medicine*, 9, 559-572.
- Van der Heijden P.G.M., Gilula, Z., and Van der Ark, L.A. (1999). On a relationship between joint correspondence analysis and latent class analysis. *Sociological Methodology*, 29, 147-186.

- Vermunt, J.K. (1997). *Log-linear Models for Event Histories*. Thousand Oakes: Sage Publications.
- Vermunt, J.K., and Magidson, J. (2000). *Latent GOLD 2.0 User's Guide*. Belmont, MA: Statistical Innovations Inc.
- Vermunt, J.K., and Magidson, J. (2002). Latent class cluster analysis. J.A. Hagenaars and A.L. McCutcheon (eds.), *Applied Latent Class Analysis*, 89-106. Cambridge University Press.
- Vermunt, J.K. and Van Dijk, L. (2001). A nonparametric random-coefficients approach: the latent class regression model. *Multilevel Modelling Newsletter*, 13, 6-13.
- Wedel, M., and DeSarbo, W.S (1994). A review of recent developments in latent class regression models. R.P. Bagozzi (ed.), *Advanced Methods of Marketing Research*, 352-388, Cambridge: Blackwell Publishers.
- Wedel, M., and Kamakura, W.A. (1998). *Market Segmentation: Concepts and Methodological Foundations*. Boston: Kluwer Academic Publishers.
- Yung, Y.F. (1997). Finite mixtures in confirmatory factor-analysis models. *Psychometrika*, 62, 297-330.