Structural Equation Modeling, 18:397–418, 2011

ISSN: 1070-5511 print/1532-8007 online
DOI: 10.1080/10705511.2011.582016

Psychology Press
Taylor & Francis Group

Two Studies of Specification Error in Models for Categorical Latent Variables

David Kaplan and Sarah Depaoli
Department of Educational Psychology, University of Wisconsin–Madison

This article examines the problem of specification error in 2 models for categorical latent variables; the latent class model and the latent Markov model. Specification error in the latent class model focuses on the impact of incorrectly specifying the number of latent classes of the categorical latent variable on measures of model adequacy as well as sample reallocation to latent classes. The results show that the clarity of remaining latent classes, as measured by the entropy statistic depends on the number of observations in the omitted latent class—but this statistic is not reliable. Specification error in the latent Markov model focuses on the transition probabilities when a longitudinal Guttman process is incorrectly specified. The findings show that specifying a longitudinal Guttman process that is not true in the population impacts other transition probabilities through the covariance matrix of the logit parameters used to calculate those probabilities.

Keywords: entropy, information matrix, latent class analysis, latent Markov model, specification error

Central to the goal of empirically testing psychological theories is attempting to capture as closely as possible the data generating process. In recent years there has been significant progress in the development of statistical methods designed to model categorical latent variables, including latent class models and latent Markov models. Categorical latent variables are applied to problems in which the goal is to determine the underlying latent classes to which individuals belong. The fundamental model is latent class analysis (e.g., Clogg, 1995). When applied to longitudinal data, the extension of latent class analysis is referred to as latent Markov modeling, in which the focus is addressing transitions over time in a developmental process.

Correspondence should be addressed to David Kaplan, Department of Educational Psychology, University of Wisconsin–Madison, Educational Sciences, Room 1082B, Madison, WI 53706, USA. E-mail: dkaplan@education.wisc.edu

Both procedures have been successfully used in numerous substantive applications such as the onset and development of substance abuse in early adolescence (e.g., Collins, Hyatt, & Graham, 2000) and stage-sequential models for reading development (Chall, 1965; Kaplan & Walpole, 2005). However, little is known about the behavior of these methodologies under various types of specification error. Two forms of misspecification are examined in this study. The first form of specification error concerns misspecifying the number of latent classes of the categorical latent variable. The second form of misspecification focuses on the latent Markov model and concerns misspecifying the structure of the transition probabilities over time—for example, assuming that a Markov model follows a longitudinal Guttman process when that assumption is false. Thus, the purpose of this article is to study these two forms of misspecification error to examine their effects with respect to inferences drawn from application of the method.

The organization of this article is as follows. In the next section, we describe the latent class model and the latent Markov model, focusing on a variant of the latent Markov model, referred to as latent transition analysis. We also present issues of model estimation and testing. Our presentation of these models follows recent developments that show how models for categorical latent variables can be parameterized as finite mixture models (see, e.g., McLachlan & Peel, 2000). Next, we present the design of simulation Study 1 that examines misspecification of number of latent classes of the categorical latent variable, followed by the design of simulation Study 2 that examines misspecification of the structure of the transition probability matrix in the latent Markov model. Results of both studies follow. We then conclude with a summary of the findings and recommendations for practice.

MODEL SPECIFICATION, ESTIMATION, AND TESTING

In this section, we provide the specifications of the latent class model, followed by the manifest Markov model that sets the stage for the latent Markov model, and its special case, the latent transition model. We also discuss estimation via a finite mixture modeling perspective, and then turn to the problem of model testing.

The Latent Class Model

The latent class model can be written as follows. Let

$$P_{defg} = \sum_{c=1}^{C} \delta_c \rho_{d|c} \rho_{e|c} \rho_{f|c} \rho_{g|c}, \tag{1}$$

where P_{defg} is the probability of giving a particular response to items d, e, f, and g. The parameter δ_c is the proportion of individuals in latent class c. The parameters $\rho_{d|c}$, $\rho_{e|c}$, $\rho_{f|c}$, and $\rho_{g|c}$ are the response probabilities for items *d, e, f,* and *g*, respectively, conditional on membership in latent class *c*.

The application of latent class analysis is quite similar to the application of factor analysis. That is, an investigator would hypothesize a priori a categorical latent variable $\mathbf{\Omega}$ with C latent classes. Under the hypothesis of C latent classes, the model is fit to the observed categorical data. The patterns of response probabilities are used to name the latent classes. Various measures of model fit (described later) can be used to test the hypothesis that the model based on C latent classes reproduces the observed categorical responses.

The Latent Markov Model

Extending the notation of the latent class model, here let $\mathbf{\Omega}_c$ represent a categorical variable, where here C represents the initial number of latent states. Because our interest is in modeling movement across latent states over time, the process starts with individuals assigned to latent state c with probability $\rho^1_{d|c}$. The latent distribution of responses at $t = 1$ is given by δ^1_c. Next, the transition to latent state u given membership in the initial latent state c is governed by $\tau^{21}_{u|c}$, and so on. Thus, the latent Markov model can be written as

$$P_{defg} = \sum_{c=1}^{C}\sum_{u=1}^{U}\sum_{v=1}^{V}\sum_{w=1}^{W} \delta^1_c \rho^1_{d|c} \tau^{21}_{u|c} \rho^2_{e|u} \tau^{32}_{v|u} \rho^3_{f|v} \tau^{43}_{w|v} \rho^4_{g|w}. \tag{2}$$

Note that Equation 2 reveals that when the response probabilities are all 1.0 (indicating perfect measurement of the latent variable), then Equation 2 reduces to a manifest Markov model.

Calculation of transition probabilities is accomplished as follows. First, let $\mathbf{\Omega}^t_c$ represent the categorical latent variable containing C latent classes measured at time t. For simplicity, let the categorical latent variable have $C = 2$ classes measured at $t = 2$ time points. The transition from $\mathbf{\Omega}_1$ to $\mathbf{\Omega}_2$ can be estimated via a logistic regression of $\mathbf{\Omega}_2$ on $\mathbf{\Omega}_1$, yielding a logit intercept α and logit slope γ. Then, in line with Asparouhov and Muthén (2007), the transition probability from Time 1 to Time 2 can be written as

$$\tau^{21} \equiv P(\mathbf{\Omega}_2 = 1|\mathbf{\Omega}_1) = \frac{\exp(\alpha_2 + \gamma I(\mathbf{\Omega}_1))}{\exp(\alpha_2 + \gamma I(\mathbf{\Omega}_1)) + 1}, \tag{3}$$

where $I(\mathbf{\Omega}_1)$ is an indicator variable for the latent class variable $\mathbf{\Omega}_1$, where $I(\mathbf{\Omega}_1) = 1$ if $\mathbf{\Omega}_1 = 1$ and $I(\mathbf{\Omega}_1) = 0$ if $\mathbf{\Omega}_1 = 2$.

Extension to Latent Transition Analysis

The combination of multiple indicator categorical latent variable models and Markov models provides the foundation for the latent transition analysis of stage-sequential dynamic latent variables. In the example given in Kaplan and Walpole (2005), they considered the problem of change over time in discrete reading skills. The data they analyzed provided information on the mastery of five different reading skills. At any given point in time, a child either mastered or did not master one or more of these skills. It was deemed reasonable to postulate a model that specified that these reading skills were related in such a way that mastery of a later skill implied mastery of all preceding skills. At each time point, the child's latent class membership defined his or her latent status. The model specified a particular type of change in latent status over time. This is defined by Collins and Flaherty (2002) as a "model of stage-sequential development, and the skill acquisition process is a stage-sequential dynamic latent variable" (p. 289). It is important to point out that there is no fundamental difference between latent transition analysis and latent Markov modeling. The difference is practical, with latent transition analysis being perhaps better suited conceptually for the study of change in developmental status. The model form for latent transition analysis utilizes Equation 2 except that model estimation is undertaken with multiple indicators of the categorical latent variable.

Estimation

For this article the latent class and latent Markov models are estimated via maximum likelihood (ML) using the EM algorithm (Dempster, Laird, & Rubin, 1977) under a finite mixture modeling perspective (McLachlan & Krishnan, 1997; McLachlan & Peel, 2000). Drawing on Everitt (1984) and McLachlan and Peel (2000) let $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2 \ldots \mathbf{y}_n)'$ be a set of binary variables obtained on a random sample of size n. In our motivating example, these variables represent mastery/nonmastery scores on a set of reading subtests. The density of an observation $\mathbf{y}_i$, $(i = 1, 2, \ldots n)$ can be written as

$$f(\mathbf{y}_i|\boldsymbol{\Psi}) = \sum_{c=1}^{C} \pi_c f(\mathbf{y}_i|\boldsymbol{\theta}_c), \tag{4}$$

where

$$\boldsymbol{\Psi} = (\boldsymbol{\pi}, \boldsymbol{\Theta}')', \tag{5}$$

is a vector of unknown parameters containing the mixing proportions $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_c)$ and the model parameters $\boldsymbol{\Theta} = (\boldsymbol{\theta}'_1, \ldots, \boldsymbol{\theta}'_C)$. The C densities are assumed to follow a finite mixture multivariate Bernoulli distribution with mixing proportions π_c.

The likelihood function can be written as

$$L = \prod_{i=1}^{n} f(\mathbf{y}_i|\boldsymbol{\Psi}), \tag{6}$$

and the observed data log-likelihood is given by

$$\log L = \sum_{i=1}^{n} \log\left\{\sum_{c=1}^{C} \pi_c f(\mathbf{y}_i|\boldsymbol{\theta}_j)\right\}. \tag{7}$$

Differentiating Equation 7 with respect to the unknown parameters yields

$$\hat{\pi}_c = \frac{1}{n}\sum_{i=1}^{n} \tau_c(\mathbf{y}_i|\hat{\boldsymbol{\theta}}_i), \tag{8}$$

and

$$\sum_{c=1}^{C}\sum_{i=1}^{n} \tau_c(\mathbf{y}_i|\hat{\boldsymbol{\Psi}}_i)\partial \log f(\mathbf{y}_i|\boldsymbol{\theta}_c)/\partial\boldsymbol{\Theta} = \mathbf{0}, \tag{9}$$

where

$$\tau_c(\mathbf{y}_i|\boldsymbol{\Psi}) = \frac{\pi_c f(\mathbf{y}_i|\boldsymbol{\theta}_c)}{\sum_{c=1}^{C} \pi_c f(\mathbf{y}_i|\boldsymbol{\Theta})} \tag{10}$$

is the posterior probability that $\mathbf{y}_i$ belongs to the cth mixture (latent) class.

The EM algorithm proceeds by obtaining initial starting values of $\boldsymbol{\pi}$ and $\boldsymbol{\theta}_c$, which are then inserted into Equation 10 to obtain initial posterior probabilities. These initial posterior probabilities are then inserted into Equation 8 to obtain revised estimates of $\boldsymbol{\pi}$ and $\boldsymbol{\theta}$. This iterative process continues until a convergence criteria is met.

Measures of Model Adequacy

For this article, we examine several measures of model adequacy. The measures we chose are by no means exhaustive, but are representative of the measures of model adequacy used in practical settings. The first two measures are classical measures of model fit, including the Pearson chi-square test and the likelihood ratio chi-square test. The Pearson chi-square test can be obtained as follows. Let F_{defg} be observed frequencies of the *DEFG* contingency table and let f_{defg} be the expected cell counts. The Pearson chi-square test is written as

$$\chi^2 = \sum_{defg} \frac{(F_{defg} - f_{defg})^2}{f_{defg}}, \tag{11}$$

and the likelihood ratio chi-square test can be written as

$$LR = 2\sum_{defg} F_{defg} \ln(F_{defg}/f_{defg}) \tag{12}$$

where the degrees of freedom are obtained by subtracting the number of parameters to be estimated from the total number of cells of the contingency table that are free to vary. In cases where there are sizable disagreements between the Pearson chi-square test and the likelihood ratio chi-square test, it is likely due to the occurrence of sparse cells.

The next two indexes provided methods of model comparison and selection. These include Akaike's (1985) Information Criterion (AIC) and Schwarz's (1978) criteria, more commonly referred to as the Bayesian Information Criterion (BIC). The AIC can be written as

$$AIC = \chi^2 - 2df, \tag{13}$$

and the BIC can be written as

$$BIC = \chi^2 - q[ln(N)], \tag{14}$$

where q is the number of parameters in the model and N represents the sample size. These indexes can be used to assess model adequacy from a predictive point of view, where the lowest value of an information criterion indicates the best fitting model.

In addition to model selection measures, an important criterion for judging model adequacy is the extent to which latent profiles can be clearly distinguished. One approach is to examine the estimated posterior probabilities of latent profile assignment for each individual. A summary of measure of classification adequacy is given by the so-called entropy measure (Ramaswamy, Desarbo, Reibstein, & Robinson, 1993), and written as

$$Entropy = 1 - \frac{\sum_i \sum_c (-\hat{p}_{ic} \ln \hat{p}_{ic})}{n \ln C}, \tag{15}$$

where $\hat{P}_{ic}$ is the estimated conditional probability of teacher i begin in profile c. Note that these values range from 0 to 1, where one denotes perfect clear classification.

Asymptotic Covariance Matrix of the Estimates

In the context of Study 2, interest centers on the calculation of the asymptotic covariance matrix of the estimates as this is believed to govern the manner in which specification errors propagate through the equations of the system. Let the observed data log-likelihood be written as

$$\log L = \sum_{i=1}^{n} \log L_i. \tag{16}$$

Three alternative methods can be used to obtain the asymptotic covariance matrix of the estimates. Let the observed data log-likelihood be written as

$$\log L = \sum_{i=1}^{n} \log L_i. \tag{17}$$

The first approach, referred to as *MLF*, approximates Fisher's information matrix via

$$I_{MLF} = \sum_{i=1}^{n} \frac{\partial \log L_i}{\partial \boldsymbol{\Psi}} \times \frac{\partial \log L_i}{\partial \boldsymbol{\Psi}'}. \tag{18}$$

Standard ML estimation approximates Fisher's information matrix as

$$I_{ML} = \sum_{i=1}^{n} \frac{\partial^2 \log L_i}{\partial \boldsymbol{\Psi} \partial \boldsymbol{\Psi}'}. \tag{19}$$

Finally, the covariance matrix of the maximum likelihood estimator with robust standard errors (MLR) can be obtained as

$$I_{MLR}^{-1} = I_{ML}^{-1} I_{MLF} I_{ML}^{-1}. \tag{20}$$

Equation 20 is also referred to as the *Huber sandwich estimator* (Huber, 1967), and the square root of the diagonal elements of Equation 20 are the *Huber–White standard errors* (Huber, 1967; White, 1980; White, 1994).

A feature of the asymptotic covariance matrix of the estimates under ML estimation is that for complex models, this matrix often contains patterns of zero and nonzero elements which have been shown to govern the propagation of specification errors in high-dimensional models. Specifically, in several early papers on specification errors in structural equation models (Kaplan, 1998, 1989; Saris, Satorra, & Sörbom, 1987; Satorra, 1989) it was noted that specification errors in measurement and structural equations did not propagate their effects throughout the entire system of equations, as might be expected by full information ML estimation. Kaplan (1998) speculated that the reason for this finding resided in the pattern of zero and nonzero elements in the asymptotic covariance matrix of the estimates. Further investigation into this issue led back to the classical results of Aitchison (1962) on the problem of asymptotically independent parameters and separable hypotheses. A paper by Kaplan and Wenger (1993) showed, among other things, how this feature of I_{ML}^{-1} governs the propagation of specification errors throughout systems of structural equations. In this article, we show that specification errors in transition probabilities are governed by the pattern of zero and nonzero values of the asymptotic covariance matrix of the MLR estimator.

DESIGN

The design of Study 1 assesses misspecification in latent class analysis where the true number of latent classes in the population are misspecified in the analysis model. The concern here is the manner in which class proportions are reallocated to remaining latent classes. Study 1 also serves as a partial replication of Nylund, Asparouhov, and Muthén (2007).

The design of Study 2 represents specification error in the latent Markov model by misspecifying the structural part of the model. Specifically, we assume that the structure of the transition probability matrix in the population allows for a form of "backtracking" or "forgetting." In the context of the Kaplan and Walpole (2005) reading example, this might not be possible in that once mastery is gained in a lower order skill, one cannot lose that skill and still proceed to a higher order skill. Other substantive contexts might, however, give rise to this specification. This transition structure is then misspecified to represent a strict Guttman process in the analysis model, namely, a process where there is no backtracking or forgetting. It should be noted, however, that one could start with a fully specified transition matrix in the population and examine the effect of incorrectly fixing a number of transition probabilities in the analysis sample. It was decided that for simplicity, and to more clearly monitor the effect of specification error propagation, that focus would be on only one incorrectly fixed parameter.

Finally, it is not possible to examine the role that misspecification of the measurement structure plays in the latent Markov model because misspecifying the number of latent classes in the latent Markov modeling context changes the dimension of the transition probability matrix.

Design of Study 1

Latent class misspecification is assessed by first specifying a true four-class model in the population and then analyzing the model under a three-class specification. In the context of finite mixture modeling, this misspecification is akin to inadvertently omitting a component of the finite mixture distribution. Logit thresholds for eight binary outcomes measuring the four latent classes (C1–C4) are chosen to yield class proportions of approximately 10%, 53%, 18%, and 20%, respectively. These proportions are not unusual in real data examples.

Two different versions of the three-class model are studied. In the first version, the latent class corresponding to the fewest number of cases in the four-class population model is not considered in the three-class analysis model. In the second version, the latent class corresponding to the most number of cases in the four-class population model is not considered in the three-class analysis model. Specifically, the first condition omits C1 (accounting for approximately 10% of the cases) and thus sets C2 as the first class; the class-ordering in this analysis is C2, C3, and C4. The second condition omits C2 (accounting for approximately 53% of the cases) and, as a result, retains C1 as the first class; the class-ordering is C1, C3, and C4. Both of these misspecified conditions are studied under sample sizes of n = 500, 1,000, and 3,000. These sample sizes are admittedly large, but allow us to ensure that the ($2^8 = 256$) cells have a sufficient number of observations. The omitted classes conditions and sample size conditions are completely crossed. There are 220 degrees of freedom for the true model and 229 degrees of freedom for the misspecified model.

Of interest in Study 1 is the effect of omitting a latent class on the measures of model fit and predictive accuracy discussed earlier. In addition, a set of class membership statistics are examined, including the class proportions based on the estimated model, and the classification of likely class membership. Finally, we present the entropy index, which provides an assessment of the clarity of the classification.

Design of Study 2

For this study, we focus on the lag–1 Markov model insofar as this is the most common form of the Markov model used in social and behavioral science research. Misspecification in the latent Markov model is accomplished by specifying a non-Guttman process in the population and then analyzing the model as if it were a strict Guttman process. Specifically, the population logit parameter values corresponding to the transition from Class 2 at Time 1 back to Class 1 at Time 2 was chosen to yield transition probability values of .2, .4, and .6. These transition probabilities are typical of those found in substantive literature (see, e.g., Kaplan & Walpole, 2005). When imposing the restriction of a strict Guttman response pattern, these transition probability values are fixed to zero. For Study 2, we examine specification error only for a sample size of 3,000. Preliminary studies using sample sizes of 500 and 1,000 reveal virtually no parameter estimate bias, but bias in the standard errors relative to the empirical variability of the estimates are observed, which, as expected, improves with increasing sample sizes.[1] Three latent classes are specified for all of the population models in this study, and 1,000 replications were used for each analysis.

RESULTS

All analyses utilized the Monte Carlo simulation feature of the M*plus* software program (Muthén & Muthén, 1998–2007). For each analysis across both studies, 1,000 replications

[1]Tables for the 500 and 1,000 sample size conditions are available from the authors on request.

TABLE 1
Study 1 Model Fit Indexes and Class Membership Proportions

Model	*Entropy*	*AIC*	*BIC*	*Pearson* χ^2	*Likelihood Ratio* χ^2	*Proportions: Est. Model*	*Proportions: Most Likely Class*
500 cases							
True model	.724	4,445.387	4,592.899	222.745	223.471	C1 = .095	.104
						C2 = .530	.558
						C3 = .170	.153
						C4 = .205	.187
Class 1 removed	.688	4,457.480	4,567.059	267.025	253.465	C2 = .532	.559
						C3 = .227	.226
						C4 = .241	.215
Class 2 removed	.841	4,463.653	4,573.233	266.177	259.631	C1 = .107	.115
						C3 = .231	.207
						C4 = .663	.678
1,000 cases							
True model	.771	8,862.065	9,033.836	222.165	246.097	C1 = .088	.097
						C2 = .532	.563
						C3 = .181	.162
						C4 = .199	.179
Class 1 removed	.750	8,897.055	9,024.657	294.236	299.195	C2 = .531	.564
						C3 = .230	.229
						C4 = .239	.207
Class 2 removed	.737	8,908.087	9,035.688	298.105	310.438	C1 = .102	.110
						C3 = .236	.213
						C4 = .661	.677
3,000 cases							
True model	.766	26,513.740	26,723.963	221.564	240.310	C1 = .080	.091
						C2 = .530	.564
						C3 = .189	.163
						C4 = .201	.182
Class 1 removed	.764	26,643.358	26,799.524	397.893	388.013	C2 = .516	.559
						C3 = .236	.236
						C4 = .247	.205
Class 2 removed	.841	26,680.232	26,836.398	437.843	424.888	C1 = .090	.097
						C3 = .233	.212
						C4 = .677	.691

Note. C1–C4 = the four latent classes; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion.

were used. No fewer than 997 out of 1,000 replications converged for Study 1. All 1,000 replications converged for each model in Study 2.[2]

Results of Study 1

Classification results for Study 1 are presented in Table 1. The clarity of the classification results can be assessed through the entropy measure along with the proportion of cases categorized into each of the classes. Results for the sample size conditions of 500 and 3,000 indicated that

[2] The number of random starts for the models used in this study were 100. The number of final stage optimization steps was 10. These settings were used to ensure that estimates were not the result of problems with local maxima.

removal of Class 2, the class accounting for the largest proportion of cases (approximately 53% of the cases), resulted in classes that were ultimately more distinct from one another. For example, by removing Class 2 for $n = 500$, the entropy value was .841. The latent class proportions based on the estimated model for this condition were .107, .231, and .663. These proportions indicate that the classes are relatively distinct from one another and the entropy value is therefore relatively high.

These general results can be compared to the condition when Class 1 (the class accounting for the smallest proportion) was removed. In this case, only about 10% of the cases were reallocated to different classes. The class proportions based on the estimated model are noticeably closer in value to one another for this condition. The classes respectively accounted for .532, .227, and .241 of the overall proportion. As a result, the entropy value indicates less clarity of the classification results yielding a value of .688.

Results for the sample size conditions of $n = 500$ and 3,000 indicate that the classes are becoming more distinct from one another when the largest proportion are reallocated across the remaining three latent classes. However, results from $n = 1{,}000$ differ from these findings in that there is greater distinction between classes when the smaller class is removed; likewise, there is less distinction between the latent classes when the larger class size is reallocated. In this case, the entropy statistic did not follow the expected pattern.

Two indexes of model adequacy are also presented in Table 1 for the different conditions of this study. The AIC indexes are consistently higher than the true models. However, the BIC values are consistently higher only for the $n = 3{,}000$ misspecified conditions; the other sample size conditions yield BIC values inconsistent with this general finding. This result partially replicated the findings from Nylund et al. (2007) in that the BIC was more accurate for larger sample sizes. As expected, the Pearson chi-square and likelihood ratio chi-square tests both indicate that there is greater misfit for the misspecified models across all sample size conditions. Furthermore, for both chi-square tests, reallocating the largest proportion of cases (by removing Class 2) produces greater misfit than reallocating the cases from the smallest latent class (by removing Class 1) for most conditions; the Pearson chi-square value for $n = 500$ was the exception to this finding.

The AIC index is consistently higher than for the true models and the two chi-square tests consistently indicate model misfit. However, the BIC is slightly less systematic in its values relative to the true model. Overall, the findings indicate that entropy does not consistently measure the relative distinctiveness of the class proportions across the misspecified models. We conclude that when determining classification adequacy, the entropy measure must be used with caution, and in combination with substantive guidance regarding the classification proportions.

Results of Study 2

Table 2 shows the population transition probabilities for the .2, .4, and .6 conditions. Specifically, the transition from Class 2 at Time 1 back to Class 1 at Time 2 represents a non-Guttman process and was specified as one of the transition probabilities in the population model. This table of population transition probabilities corresponds to Tables 3 through 5, which present the population logit parameter information for each of the transition probability conditions when $n = 3{,}000$. Each of these population models produced logit parameter estimates comparable to the population logit parameter values specified. For these logit estimates, the average standard

TABLE 2
Population Transition Probabilities

.2 Condition				*.4 Condition*				*.6 Condition*			
	T2 Classes				*T2 Classes*				*T2 Classes*		
T1 Classes	*C1*	*C2*	*C3*	*T1 Classes*	*C1*	*C2*	*C3*	*T1 Classes*	*C1*	*C2*	*C3*
C1	.301	.687	.012	C1	.301	.687	.012	C1	.301	.687	.012
C2	.200	.525	.275	C2	.400	.394	.207	C2	.600	.264	.137
C3	.000	.000	1.000	C3	.000	.000	1.000	C3	.000	.000	1.000
	T3 Classes				*T3 Classes*				*T3 Classes*		
T2 Classes	*C1*	*C2*	*C3*	*T2 Classes*	*C1*	*C2*	*C3*	*T2 Classes*	*C1*	*C2*	*C3*
C1	.633	.318	.049	C1	.632	.318	.050	C1	.633	.318	.049
C2	.000	.823	.177	C2	.000	.823	.177	C2	.000	.823	.177
C3	.000	.000	1.000	C3	.000	.000	1.000	C3	.000	.000	1.000
	T4 Classes				*T4 Classes*				*T4 Classes*		
T3 Classes	*C1*	*C2*	*C3*	*T3 Classes*	*C1*	*C2*	*C3*	*T3 Classes*	*C1*	*C2*	*C3*
C1	.302	.477	.220	C1	.301	.479	.220	C1	.301	.479	.220
C2	.000	.134	.866	C2	.000	.134	.866	C2	.000	.134	.866
C3	.000	.000	1.000	C3	.000	.000	1.000	C3	.000	.000	1.000

Note. T1–T4 = the four time points; C1, C2, and C3 = the three latent classes.

deviation across all 1,000 replications closely matches its corresponding estimated standard error. Likewise, the mean squared error values are consistently very small across all conditions. The coverage statistic for the logit parameter estimates is close to the expected 95% range for all of the true models. Coverage represents the ability of the misspecified model to reproduce the population parameters within each condition of the study. Specifically, the coverage is the percentage of replications that produced confidence intervals containing the true population parameter (Nylund et al., 2007). Likewise, power is at 100%.

Table 6 presents the model adequacy results for the population and misspecified models for the three transition probability conditions. The entropy statistic does not appear to be related to the size of the misspecification. The AIC and BIC both yield higher values for the misspecified models. Likewise, the Pearson chi-square and likelihood ratio chi-square indexes indicate relative misfit in the misspecified models. The bias measure indicates the percentage difference between the population and misspecified parameters. As the magnitude of the misspecification increases, bias also increases as expected.

Tables 7 through 9 present the transition probabilities and logit parameter estimates for the .2, .4, and .6 misspecified models. The transition probability matrices shown in the top half of the tables illustrate that the transition from Class 2 at Time 1 back to Class 1 at Time 2 was specified as zero, representing a strict Guttman process. The bottom half of the table

TABLE 3
Study 2 Population Model Logit Parameters: .2 Condition

Logit Parameter Estimates							
Regression	*Population*	*Estimate*	*SD*	*SE*	*MSE*	*95% Coverage*	*Power*
C1_T2 ON							
C2_T1	14.700	14.697	.101	.106	.010	.959	1.000
C1_T1	18.267	18.292	.300	.303	.091	.972	1.000
C2_T2 ON							
C1_T1	19.089	19.118	.300	.301	.091	.970	1.000
C2_T1	15.649	15.647	.083	.083	.007	.947	1.000
C1_T3 ON							
C1_T2	17.543	17.574	.222	.216	.050	.958	1.000
C2_T3 ON							
C1_T2	16.856	16.882	.233	.223	.055	.954	1.000
C2_T2	16.540	16.538	.066	.068	.004	.957	1.000
C1_T4 ON							
C1_T3	15.306	15.317	.144	.148	.021	.957	1.000
C2_T4 ON							
C1_T3	15.784	15.776	.138	.137	.019	.945	1.000
C2_T3	13.136	13.133	.081	.080	.007	.944	1.000
Means							
Regression	*Population*	*Estimate*	*SD*	*SE*	*MSE*	*95% Coverage*	*Power*
C1_T1	3.206	3.212	.121	.122	.015	.961	1.000
C2_T1	2.528	2.529	.126	.125	.016	.950	1.000

Note. *SD* = the average standard deviation across replications; *SE* = the average estimated standard error across all replications; *MSE* = the mean squared error for the parameter.

presents results describing the accuracy of the logit parameter estimates that were obtained for the misspecified models. Several different indicators are presented to assess the effect of the logit parameter estimates. The information provided for each regression includes the population logit parameter value, the average of the estimated logit parameter from the misspecified model across all 1,000 replications, the standard deviation of the estimates across all replications, the average estimated standard error across all replications, and the mean squared error for the parameter. The 95% coverage statistic is also presented. The final index presented is the power associated with rejecting the null hypothesis when it is, in fact, false.[3]

Table 7 shows the results for the .2 misspecification condition. The affect of this misspecification on the subsequent transition probabilities can be assessed by comparing these results to the population results presented in Table 2. For example, there are some transition probabilities appearing to be substantially affected by this misspecification. One such probability represents the transition from Class 1 at Time 1 to Class 2 at Time 2, which changed from .687 in the true model to .659 in the misspecified model. There were also several transitions virtually

[3]In the M*plus* program, this column is labeled as "% Sig" giving the proportion over the number of replications in which the null hypothesis is rejected when it is true (Muthén & Muthén, 1998–2007).

TABLE 4
Study 2 Population Model Logit Parameters: .4 Condition

Logit Parameter Estimates							
Regression	*Population*	*Estimate*	*SD*	*SE*	*MSE*	*95% Coverage*	*Power*
C1_T2 ON							
C2_T1	15.670	15.670	.091	.094	.008	.954	1.000
C1_T1	18.267	18.291	.285	.291	.082	.968	1.000
C2_T2 ON							
C1_T1	19.089	19.115	.285	.290	.082	.969	1.000
C2_T1	15.649	15.648	.097	.097	.009	.950	1.000
C1_T3 ON							
C1_T2	17.543	17.554	.186	.184	.035	.960	1.000
C2_T3 ON							
C1_T2	16.856	16.865	.195	.191	.038	.952	1.000
C2_T2	16.540	16.539	.069	.071	.005	.954	1.000
C1_T4 ON							
C1_T3	15.306	15.316	.125	.133	.016	.963	1.000
C2_T4 ON							
C1_T3	15.784	15.781	.121	.122	.015	.955	1.000
C2_T3	13.136	13.133	.082	.081	.007	.945	1.000
Means							
Regression	*Population*	*Estimate*	*SD*	*SE*	*MSE*	*95% Coverage*	*Power*
C1_T1	3.206	3.212	.121	.121	.015	.957	1.000
C2_T1	2.528	2.530	.126	.125	.016	.948	1.000

Note. *SD* = the average standard deviation across replications; *SE* = the average estimated standard error across all replications; *MSE* = the mean squared error for the parameter.

unaffected by this misspecification such as the transition from Class 1 at Time 3 to Class 1 at Time 4, which was .302 in the true model and .307 in this misspecified condition.

The transition probabilities that were most affected were associated, as expected, with logit parameter estimates that were not as precise. The standard deviation, standard error, and mean squared errors show an increase for some of the logit parameters affected by the misspecification. One example is the regression of Class 1 at Time 2 on Class 1 at Time 1 where the standard deviation, standard error, and mean squared error all increased in the misspecified model when compared to the true model results presented in Table 3. Likewise, the coverage statistic decreased from .959 to .721 for this particular regression. This contrasts with the results corresponding to the unaffected transition probabilities. For instance, the regression of Class 1 at Time 4 on Class 1 at Time 3 did not show any notable changes in the logit parameter estimate.

Table 8 presents the results for the .4 transition probability misspecification condition and Table 9 presents the misspecified results for the .6 condition. A similar pattern of results emerged for these conditions as for the .2 condition. Some of the transition probabilities were more affected by the misspecification and others were less affected. Overall, it appears that the misspecification only had an effect on some of the logit parameter estimates whereas other estimates were left unaffected.

TABLE 5
Study 2 Population Model Logit Parameters: .6 Condition

Regression	*Population*	*Estimate*	*SD*	*SE*	*MSE*	*95% Coverage*	*Power*
Logit Parameter Estimates							
C1_T2 ON							
C2_T1	16.480	16.486	.105	.103	.011	.947	1.000
C1_T1	18.267	18.288	.268	.277	.072	.971	1.000
C2_T2 ON							
C1_T1	19.089	19.113	.267	.276	.072	.971	1.000
C2_T1	15.649	15.656	.122	.121	.015	.952	1.000
C1_T3 ON							
C1_T2	17.543	17.562	.168	.163	.029	.948	1.000
C2_T3 ON							
C1_T2	16.856	16.876	.171	.170	.030	.952	1.000
C2_T2	16.540	16.535	.071	.073	.005	.955	1.000
C1_T4 ON							
C1_T3	15.306	15.316	.118	.121	.014	.956	1.000
C2_T4 ON							
C1_T3	15.784	15.782	.116	.112	.013	.939	1.000
C2_T3	13.136	13.135	.081	.082	.007	.949	1.000
Means							
Regression	*Population*	*Estimate*	*SD*	*SE*	*MSE*	*95% Coverage*	*Power*
C1_T1	3.206	3.212	.121	.120	.015	.953	1.000
C2_T1	2.528	2.530	.126	.124	.016	.948	1.000

Note. *SD* = the average standard deviation across replications; *SE* = the average estimated standard error across all replications; *MSE* = the mean squared error for the parameter.

TABLE 6
Study 2 Model Fit Indexes

Model	*Cases*	*Entropy*	*AIC*	*BIC*	*Pearson* χ^2	*Likelihood Ratio* χ^2
Trans prob = .2						
True	3,000	.930	38,443.378	38,875.836	4,380.079	2,008.469
Misspecified	3,000	.929	38,985.448	39,411.900	6,744.826	2,469.144
Bias		−.108	1.410	1.379	53.989	22.937
Trans prob = .4						
True	3,000	.926	38,674.848	39,107.306	4,385.532	2,018.565
Misspecified	3,000	.928	39,752.026	40,178.478	7,087.422	2,948.035
Bias		.216	2.785	2.739	61.609	46.046
Trans prob = .6						
True	3,000	.926	38,634.383	39,066.841	4,349.764	1,997.364
Misspecified	3,000	.932	40,232.652	40,659.104	7,566.672	3,406.744
Bias		.648	4.137	4.076	73.956	70.562

Note. AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion. Bias = 100*[(Estimate − Population)/Population].

TABLE 7
Specification Error Results: .2 Condition

Transition Probabilities

T2 Classes

T1 Classes	*C1*	*C2*	*C3*
C1	.324	.659	.017
C2	.000	.631	.369
C3	.000	.000	1.000

T3 Classes

T2 Classes	*C1*	*C2*	*C3*
C1	.670	.286	.043
C2	.000	.825	.175
C3	.000	.000	1.000

T4 Classes

T3 Classes	*C1*	*C2*	*C3*
C1	.307	.479	.214
C2	.000	.137	.863
C3	.000	.000	1.000

Logit Parameter Estimates

Regression	*Population*	*Estimate*	*SD*	*SE*	*MSE*	*95% Coverage*	*Power*
C1_T2 ON							
C1_T1	18.267	17.995	.226	.233	.125	.721	1.000
C2_T2 ON							
C1_T1	19.089	18.705	.225	.231	.198	.552	1.000
C2_T1	15.649	15.539	.092	.091	.021	.756	1.000
C1_T3 ON							
C1_T2	17.543	17.766	.253	.242	.114	.910	1.000
C2_T3 ON							
C1_T2	16.856	16.913	.267	.254	.074	.961	1.000
C2_T2	16.540	16.549	.065	.067	.004	.950	1.000
C1_T4 ON							
C1_T3	15.306	15.362	.147	.150	.025	.943	1.000
C2_T4 ON							
C1_T3	15.784	15.809	.140	.140	.020	.956	1.000
C2_T3	13.136	13.155	.081	.079	.007	.937	1.000

Means

Regression	*Population*	*Estimate*	*SD*	*SE*	*MSE*	*95% Coverage*	*Power*
C1_T1	3.206	3.379	.124	.124	.045	.745	1.000
C2_T1	3.528	2.216	.134	.133	.115	.354	1.000

Note. *SD* = the average standard deviation across replications; *SE* = the average estimated standard error across all replications; *MSE* = the mean squared error for the parameter.

TABLE 8
Specification Error Results: .4 Condition

Transition Probabilities			
	T2 Classes		
T1 Classes	*C1*	*C2*	*C3*
C1	.366	.615	.018
C2	.000	.600	.400
C3	.000	.000	1.000
	T3 Classes		
T2 Classes	*C1*	*C2*	*C3*
C1	.670	.286	.044
C2	.000	.825	.175
C3	.000	.000	1.000
	T4 Classes		
T3 Classes	*C1*	*C2*	*C3*
C1	.305	.481	.214
C2	.000	.137	.863
C3	.000	.000	1.000

Logit Parameter Estimates

Regression	*Population*	*Estimate*	*SD*	*SE*	*MSE*	*95% Coverage*	*Power*
C1_T2 ON							
C1_T1	18.267	18.020	.197	.204	.100	.724	1.000
C2_T2 ON							
C1_T1	19.089	18.539	.196	.204	.340	.254	1.000
C2_T1	15.649	15.409	.119	.116	.072	.450	1.000
C1_T3 ON							
C1_T2	17.543	17.753	.224	.212	.094	.884	1.000
C2_T3 ON							
C1_T2	16.856	16.899	.231	.221	.055	.952	1.000
C2_T2	16.540	16.550	.067	.069	.005	.952	1.000
C1_T4 ON							
C1_T3	15.306	15.354	.127	.135	.018	.941	1.000
C2_T4 ON							
C1_T3	15.784	15.813	.122	.125	.016	.955	1.000
C2_T3	13.136	13.157	.081	.080	.007	.937	1.000

Means

Regression	*Population*	*Estimate*	*SD*	*SE*	*MSE*	*95% Coverage*	*Power*
C1_T1	3.206	3.511	.128	.127	.109	.316	1.000
C2_T1	2.528	1.831	.145	.142	.507	.009	1.000

Note. *SD* = the average standard deviation across replications; *SE* = the average estimated standard error across all replications; *MSE* = the mean squared error for the parameter.

TABLE 9
Specification Error Results: .6 Condition

Transition Probabilities

T2 Classes

T1 Classes	*C1*	*C2*	*C3*
C1	.414	.568	.018
C2	.000	.574	.426
C3	.000	.000	1.000

T3 Classes

T2 Classes	*C1*	*C2*	*C3*
C1	.663	.292	.044
C2	.000	.824	.176
C3	.000	.000	1.000

T4 Classes

T3 Classes	*C1*	*C2*	*C3*
C1	.303	.482	.215
C2	.000	.137	.863
C3	.000	.000	1.000

Logit Parameter Estimates

Regression	*Population*	*Estimate*	*SD*	*SE*	*MSE*	*95% Coverage*	*Power*
C1_T2 ON							
C1_T1	18.267	18.172	.194	.201	.047	.904	1.000
C2_T2 ON							
C1_T1	19.089	18.487	.193	.201	.399	.184	1.000
C2_T1	15.649	15.302	.172	.171	.150	.467	1.000
C1_T3 ON							
C1_T2	17.543	17.721	.194	.186	.693	.891	1.000
C2_T3 ON							
C1_T2	16.856	16.900	.197	.194	.041	.957	1.000
C2_T2	16.540	16.547	.070	.072	.005	.955	1.000
C1_T4 ON							
C1_T3	15.306	15.342	.120	.123	.016	.940	1.000
C2_T4 ON							
C1_T3	15.784	15.807	.118	.113	.014	.936	1.000
C2_T3	13.136	13.159	.080	.081	.007	.935	1.000

Means

Regression	*Population*	*Estimate*	*SD*	*SE*	*MSE*	*95% Coverage*	*Power*
C1_T1	3.206	3.615	.132	.131	.185	.082	1.000
C2_T1	2.528	1.334	.163	.160	1.453	.000	1.000

Note. *SD* = the average standard deviation across replications; *SE* = the average estimated standard error across all replications; *MSE* = the mean squared error for the parameter.

Patterns in I_{MLR}^{-1}

Tables 10 through 12 present a portion of the correlation matrix of the estimates for a subset of the logit parameters under the .2, .4, and .6 misspecification conditions. These correlations were produced from a single replication with a sample size of 500,000, thus approaching population values. In the context of our true model, the parameter $\gamma_{1|2}^{21}$ is the one associated with the non-Guttman transition probability. Similar to the results described for Tables 7 through 9, this parameter appears to relate to some of the subsequent parameters and not to others. Table 10 indicates that some of the correlations with $\gamma_{1|2}^{21}$ are relatively higher than others. The higher correlations correspond to the logit parameters that were more affected by the misspecification. Likewise, the lower correlation values correspond to the logit parameters that were largely unaffected by the misspecification. These results hold true for the .4 and .6 conditions as well.

To see the role played by the pattern in I_{MLR}^{-1} consider the logit parameter $\gamma_{2|2}^{21}$. The correlation between this parameter and the true parameter is .465 for the .2 transition probability condition. This differs from the results for $\gamma_{1|1}^{43}$, which corresponded to a relatively unaffected transition probability. The correlation between this parameter and the true parameter was −.003 for the .2 condition. A similar pattern holds for the remaining conditions presented in Tables 11 and 12. However, the magnitude of nonzero and near-zero elements became more extreme as the misspecified transition probability value increased. These results are in line with the findings of Kaplan and Wenger (1993) regarding the role of the asymptotic covariance matrix of the estimates in the propagation of specification errors.

TABLE 10
Asymptotic Correlation Matrix for the Estimates: .2 Condition

Parameter	α_1	α_2	γ_{11}^{21}	$\gamma_{2\mid 1}^{21}$	$\gamma_{1\mid 2}^{21}$	$\gamma_{2\mid 2}^{21}$	$\gamma_{1\mid 1}^{32}$	$\gamma_{2\mid 1}^{32}$	$\gamma_{2\mid 2}^{32}$	$\gamma_{1\mid 1}^{43}$	$\gamma_{2\mid 1}^{43}$	$\gamma_{2\mid 2}^{43}$
α_1	1.000	.920	−.023	−.022	**−.043**	−.051	.000	.000	−.002	.000	.000	.000
α_2	.892	1.000	.043	.043	**−.004**	−.006	.001	.001	−.002	.000	.000	.000
$\gamma_{1\mid 1}^{21}$	−.019	.065	1.000	.978	**−.037**	−.055	−.017	.000	−.025	−.003	−.002	−.002
$\gamma_{2\mid 1}^{21}$	−.022	.075	.971	1.000	**−.041**	−.055	.010	.005	−.023	.001	.000	.000
$\gamma_{1\mid 2}^{21}$						**.465**	**−.052**	**−.004**	**−.018**	**−.003**	**−.004**	**−.005**
$\gamma_{2\mid 2}^{21}$	−.061	.026	−.102	−.081		1.000	.029	.008	−.025	.002	.002	.002
$\gamma_{1\mid 1}^{32}$	−.004	.014	−.021	.013		.030	1.000	.884	−.077	−.012	−.012	−.008
$\gamma_{2\mid 1}^{32}$	.000	.001	.000	.004		.005	.894	1.000	−.094	.033	.030	.003
$\gamma_{2\mid 2}^{32}$	−.004	.006	−.024	−.020		−.008	−.076	−.094	1.000	.000	.000	−.014
$\gamma_{1\mid 1}^{43}$	−.003	.003	−.005	.001		.009	−.010	.031	.002	1.000	.544	.047
$\gamma_{2\mid 1}^{43}$	−.001	.006	−.004	.002		.011	−.009	.029	.002	.552	1.000	.015
$\gamma_{2\mid 2}^{43}$	−.001	.015	−.004	.003		.023	−.002	.004	−.011	.048	.016	1.000

Note. Values above the diagonal represent parameter correlations for the true model. Values below the diagonal represent parameter correlations for the misspecified model. The notation, $\gamma_{2|1}^{32}$, represents the logit slope for the probability of being in Class 2 at Time 3 given that you were in Class 1 at Time 2. Bold values correspond to the correlations between the misspecified parameter $\gamma_{1|2}^{21}$ and the other parameters.

TABLE 11
Asymptotic Correlation Matrix for the Estimates: .4 Condition

Parameter	α_1	α_2	γ^{21}_{11}	$\gamma^{21}_{2\|1}$	$\gamma^{21}_{1\|2}$	$\gamma^{21}_{2\|2}$	$\gamma^{32}_{1\|1}$	$\gamma^{32}_{2\|1}$	$\gamma^{32}_{2\|2}$	$\gamma^{43}_{1\|1}$	$\gamma^{43}_{2\|1}$	$\gamma^{43}_{2\|2}$
α_1	1.000	.917	−.017	−.019	**−.045**	−.050	.000	.000	−.002	.000	−.001	.000
α_2	.863	1.000	.039	.041	**−.017**	−.004	.001	.001	−.003	.000	.000	−.001
$\gamma^{21}_{1\|1}$	−.001	.068	1.000	.977	**−.070**	−.063	−.016	.000	−.027	−.003	−.002	−.002
$\gamma^{21}_{2\|1}$	−.006	.093	.968	1.000	**−.069**	−.059	.010	.005	−.025	.000	.001	.000
$\gamma^{21}_{1\|2}$						**.613**	**−.028**	**−.001**	**−.018**	**−.003**	**−.002**	**−.003**
$\gamma^{21}_{2\|2}$	−.066	.042	−.182	−.136		1.000	.033	.009	−.021	.003	.003	.001
$\gamma^{32}_{1\|1}$	−.006	.035	−.024	.015		.067	1.000	.885	−.068	−.012	−.012	−.009
$\gamma^{32}_{2\|1}$	.000	.005	−.001	.005		.011	.893	1.000	−.083	.032	.030	.004
$\gamma^{32}_{2\|2}$	−.004	.011	−.026	−.021		.006	−.076	−.096	1.000	.001	.000	−.013
$\gamma^{43}_{1\|1}$	−.005	.007	−.006	.002		.018	−.007	.030	.003	1.000	.537	.046
$\gamma^{43}_{2\|1}$	.000	.011	−.006	.001		.018	−.007	.029	.002	.545	1.000	.012
$\gamma^{43}_{2\|2}$	.003	.024	−.006	.003		.030	.000	.005	−.010	.047	.013	1.000

Note. Values above the diagonal represent parameter correlations for the true model. Values below the diagonal represent parameter correlations for the misspecified model. The notation, $\gamma^{32}_{2|1}$, represents the logit slope for the probability of being in Class 2 at Time 3 given that you were in Class 1 at Time 2. Bold values correspond to the correlations between the misspecified parameter $\gamma^{21}_{1|2}$ and the other parameters.

TABLE 12
Asymptotic Correlation Matrix for the Estimates: .6 Condition

Parameter	α_1	α_2	γ^{21}_{11}	$\gamma^{21}_{2\|1}$	$\gamma^{21}_{1\|2}$	$\gamma^{21}_{2\|2}$	$\gamma^{32}_{1\|1}$	$\gamma^{32}_{2\|1}$	$\gamma^{32}_{2\|2}$	$\gamma^{43}_{1\|1}$	$\gamma^{43}_{2\|1}$	$\gamma^{43}_{2\|2}$
α_1	1.000	.917	−.010	−.014	**−.045**	−.049	.000	.000	−.001	.000	.000	.000
α_2	.810	1.000	.029	.035	**−.029**	−.005	.001	.001	−.001	.000	.000	.000
$\gamma^{21}_{1\|1}$	.017	.082	1.000	.974	**−.092**	−.071	−.015	−.001	−.028	−.002	−.002	−.002
$\gamma^{21}_{2\|1}$	.012	.120	.970	1.000	**−.088**	−.065	.009	.004	−.026	.001	.000	.000
$\gamma^{21}_{1\|2}$						**.695**	**−.011**	**.001**	**−.015**	**−.001**	**−.001**	**−.002**
$\gamma^{21}_{2\|2}$	−.086	.062	−.267	−.206		1.000	.029	.007	−.015	.004	.002	.003
$\gamma^{32}_{1\|1}$	−.006	.053	−.028	.010		.090	1.000	.887	−.056	−.013	−.013	−.010
$\gamma^{32}_{2\|1}$	.001	.006	.000	.005		.008	.889	1.000	−.069	.032	.031	.005
$\gamma^{32}_{2\|2}$	−.003	.020	−.029	−.022		.024	−.071	−.093	1.000	.000	−.001	−.013
$\gamma^{43}_{1\|1}$	−.005	.011	−.007	.000		.024	−.007	.031	.003	1.000	.528	.045
$\gamma^{43}_{2\|1}$	.001	.019	−.007	.000		.027	−.006	.030	.003	.534	1.000	.008
$\gamma^{43}_{2\|2}$	.010	.036	−.007	.001		.032	−.001	.006	−.010	.046	.011	1.000

Note. Values above the diagonal represent parameter correlations for the true model. Values below the diagonal represent parameter correlations for the misspecified model. The notation, $\gamma^{32}_{2|1}$, represents the logit slope for the probability of being in Class 2 at Time 3 given that you were in Class 1 at Time 2. Bold values correspond to the correlations between the misspecified parameter $\gamma^{21}_{1|2}$ and the other parameters.

SUMMARY AND CONCLUSIONS

The availability of software programs for conducting latent class analysis as well as latent Markov modeling and its variants is now widespread. In this article, we utilized M*plus* for our analyses, but programs such as WinLTA (Collins, Lanza, Schafer, & Flaherty, 2002), SAS PROC LTA (SAS Institute, 2000), Latent GOLD (Vermunt & Magidson, 2000), as well as programs available in the R programming environment (R Development Core Team, 2008), now allow social and behavioral scientists to easily employ these powerful methods to gain insights into substantive problems. As such, and in line with studies of other methodologies, it is important to understand how these models behave when their underlying assumptions are violated.

This article focuses specifically on one, arguably, major assumption when employing these models—namely, correct model specification. In the context of latent class analysis, specification error manifests itself most obviously in the omission of a latent class. In our investigation, we found that the reallocation of observations to remaining latent classes did not necessarily result in clear separation of the latent classes. Clarity of separation was found to be related, in part, to the size of the omitted latent class, but the entropy measure was not a reliable indicator of clarity of class separation across different sample size conditions. These results are similar to the findings presented in Lubke and Muthén (2007), which in part compared correct class assignment with the entropy measure in the context of factor mixture models. They found that the improvement of correct class assignment did not correspond to an increase in the entropy index under conditions of lower class separation. Although these results are not sufficient to conclude that the entropy index is an unreliable indicator of classification clarity, they do suggest the need for more extensive research on this measure under different modeling conditions in order to determine the accuracy of the entropy index. We conclude that the entropy measure should be used with caution and in conjunction with class proportion measures as well as substantive theory to decide on the number of latent classes to retain.

We also note that only a limited selection of model adequacy measures were included in this study. There are additional measures that have been used in similar modeling situations for determining class enumeration. For example, Nylund, Asparouhov, and Muthén (2007) found the Bootstrap Likelihood Ratio Test (BLRT) to be a consistent indicator of latent classes within the context of latent class analysis. Likewise, the Lo–Mendell–Rubin (LMR) test is another alternative for assessing latent class enumeration. The BLRT and LMR tests are similar in that the difference between two latent class models (e.g., a three-class model vs. a two-class model) is not assumed to be chi-square distributed as is the case with the likelihood ratio test. There are also other information-based criteria, such as the sample size adjusted BIC and the consistent AIC, that are commonly used to assist in assessing class enumeration. For a more detailed study directly focused on the optimal criteria used for assessing class enumeration within the context of latent class analysis, see Nylund, Asparouhov, and Muthén (2007).

In the context of latent Markov modeling, structural relations among logit parameters (and corresponding transition probabilities) can be specified to represent a strict Guttman process or a process that allows for some form of backtracking or forgetting over time. We found, as in studies of structural equation modeling, that specifying a strict Guttman process that is not true in the population will result in a propagation of bias in parameter estimates that is controlled by the asymptotic covariance matrix of the logit parameters. As noted by Kaplan and Wenger (1993), this result is somewhat troubling insofar as the initial specification of the model governs the pattern of values in I^{-1}_{MLR}, and subsequent changes to the model will likely result in new biases that might be difficult, if not impossible, to predict ahead of time. However, as originally suggested by Kaplan (1998), a paper by Yuan, Marshall, and Bentler (2003) provided a Hausman-like specification error test (Hausman, 1978) to gauge the impact of an omitted parameter on other parameters in the model. Yuan, Marshall, and Bentler (2003) noted that the concepts of asymptotic independence and separability discussed in Kaplan and Wenger (1993) might be apropos to their test. We advocate additional research on this topic, extended to latent transition models.

Not every form of specification error was examined in this article. Another form of misspecification relevant to the latent Markov model that is worthy of study concerns the problem of incorrectly specifying homogenous or heterogenous transition probabilities, akin to the problem of invariance testing in structural equation modeling. We also note that only lag-1 models were considered in this study. Although uncommon, higher lagged Markov models can be specified. We expect that there would be differences in the propagation of specification error throughout the logit parameters depending on the exact form of the specification error, but that the mechanism that gives rise to this propagation (i.e., the pattern of zero and nonzero elements in the information matrix) is the same regardless of the lag. Nevertheless, we view this study as the first to explore specification error in both latent class analysis and latent Markov modeling with the hope of providing insights into the behavior of these important analytic methods.

REFERENCES

Aitchison, J. (1962). Large-sample restricted parametric tests. *Journal of the Royal Statistical Society, Series B, 24,* 234–250.

Akaike, H. (1985). Prediction and entropy. In A. C. Atkinson & S. E. Feinberg (Eds.), *A celebration of statistics* (pp. 1–24). New York, NY: Springer-Verlag.

Asparouhov, T., & Muthén, B. O. (2007). Multilevel mixture models. In G. R. Hancock & K. M. Samuleson (Eds.), *Advances in latent variable mixture models* (pp. 27–51). Charlotte, NC: Information Age.

Chall, J. S. (1965). *Stages of reading development* (2nd ed.). Belmont, CA: Wadsworth.

Clogg, C. C. (1995). Latent class models. In G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.), *Handbook of statistical modeling in the social and behavioral sciences* (pp. 81–110). San Francisco, CA: Jossey-Bass.

Collins, L. M., & Flaherty, B. P. (2002). Latent class models for longitudinal data. In J. A. Hagenaars & A. L. McCutcheon (Eds.), *Applied latent class analysis* (pp. 287–303). New York, NY: Cambridge University Press.

Collins, L. M., Hyatt, S. L., & Graham, J. W. (2000). Latent transition analysis as a way of testing models of stage-sequential change in longitudinal data. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 147–161). Mahwah, NJ: Lawrence Erlbaum Associates.

Collins, L. M., Lanza, S. T., Schafer, J. L., & Flaherty, B. P. (2002). *WinLTA User's Guide.* State College, PA: Pennsylvania State University, University Park Campus, Methodology Center.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B, 39,* 1–38.

Everitt, B. S. (1984). *An introduction to latent variable models.* New York, NY: Chapman & Hall.

Hausman, J. A. (1978). Specification error tests in econometrics. *Econometrica, 46,* 1251–1271.

Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1,* 221–223.

Kaplan, D. (1988). The impact of specification error on the estimation, testing, and improvement of structural equation models. *Multivariate Behavioral Research, 23,* 69–86.

Kaplan, D. (1989). A study of the sampling variability and z-values of parameter estimates from misspecified structural equation models. *Multivariate Behavioral Research, 2,* 41–57.

Kaplan, D., & Walpole, S. (2005). A stage-sequential model of reading transitions: Evidence from the early childhood longitudinal study. *Journal of Educational Psychology, 97,* 551–563.

Kaplan, D., & Wenger, R. N. (1993). Asymptotic independence and separability in covariance structure models: Implications for specification error, power, and model modification. *Multivariate Behavioral Research, 28,* 483–498.

Lubke, G., & Muthén, B. O. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling, 14,* 26–47.

McLachlan, G., & Krishnan, T. (1997). *The* EM *algorithm and extensions.* New York, NY: Wiley.

McLachlan, G., & Peel, D. (2000). *Finite mixture models.* New York, NY: Wiley.

Muthén, L. K., & Muthén, B. O. (1998–2007). *M*plus *user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.

Nylund, K., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14,* 535–569.

R Development Core Team. (2008). *R: A language and environment for statistical computing.* Vienna, Austria: Author.

Ramaswamy, V., Desarbo, W. S., Reibstein, D. J., & Robinson, W. T. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science, 12,* 103–124.

Saris, W. E., Satorra, A., & Sörbom, D. (1987). The detection and correction of specification errors in structural equation models. In C. C. Clogg (Eds.), *Sociological methodology 1987* (pp. 105–129). San Francisco, CA: Jossey-Bass.

SAS Institute. (2000). SAS Online Doc, version 8 [Computer software]. Carey, NC: Author.

Satorra, A. (1989). Alternative test criteria in covariance structure analysis: A unified approach. *Psychometrika, 54,* 131–151.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6,* 461–464.

Vermunt, J. K., & Magidson, J. (2000). *Latent gold user's guide: Statistical innovations.* Belmont, MA: Statistical Innovations.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48,* 817–838.

White, H. (1994). *Estimation, inference, and specification analysis.* Cambridge, UK: Cambridge University Press.

Yuan, K.-H., Marshall, L. L., & Bentler, P. M. (2003). Assessing the effect of model misspecifications on parameter estimates in structural equation models. In R. M. Stolzenberg (Ed.), *Sociological methodology* (Vol. 23, pp. 241–265). Boston, MA: Blackwell.