Available online at www.sciencedirect.com

SCIENCE DIRECT®

Addictive Behaviors 31 (2006) 1050–1066

ADDICTIVE BEHAVIORS

Item response mixture modeling: Application to tobacco dependence criteria

Bengt Muthen [a,*], Tihomir Asparouhov [b]

[a] UCLA, United States
[b] Muthen & Muthen, United States

Abstract

This paper illustrates new hybrid latent variable models that are promising for phenotypical analyses. The hybrid models combine features of dimensional and categorical analyses seen in the conventional techniques of factor analysis and latent class analysis. The paper focuses on the analysis of categorical items, which presents especially challenging analyses with hybrid models and has recently been made practical in the Mplus program. The hybrid models are typically seen to fit data better than conventional models of factor analysis (IRT) and latent class analysis. An illustration is given in the form of analysis of tobacco dependence in a general population survey.

Keywords: Item response mixture modeling; Tobacco dependence; Latent class analysis

1. Introduction

Recent years have seen a rapid development of new latent variable methods that show promise for clarifying drug abuse phenotypes. An overview of latent variable modeling developments is given in Muthén (2002). In particular, hybrid models that simultaneously allow for both dimensional and categorical representation of latent variable constructs have proven useful. The current paper attempts to provide further understanding of the usefulness of the new hybrid models in the context of DSM-IV

* Corresponding author.
E-mail address: bmuthen@ucla.edu (B. Muthen).

0306-4603/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.addbeh.2006.03.026

diagnostic criteria for tobacco dependence. The focus is on developing hybrid models for later use in genetic analysis as well as for diagnosis. The main question is which measurement model is most suitable for understanding tobacco dependence: factor analysis, latent class analysis, or hybrid models.

Factor analysis can provide a continuous factor score variable to be used as a phenotype in a genetic analysis. Categories for diagnosis are, however, not produced. For an overview of methods for factor analysis of categorical items in the form of unidimensional traits, see, e.g., the item response theory text of Hambleton and Swaminathan (1985). Muthén (1989a) discusses general multi-factorial FA including the use of covariates. Factor analysis in the form of both unidimensional and multidimensional models has been suggested in mental health applications at many points in time: neuroticism in Duncan-Jones, Grayson, and Moran (1986); depression in Muthén (1989a, 1989b) and Gallo, Anthony, and Muthén (1994); and alcohol in Muthén (1992, 1996), Muthén, Grant, and Hasin (1993), Harford and Muthén (2001), and Krueger et al. (2004).

Latent class analysis is more closely aligned with a diagnostic approach in that it results in categories of individuals. On the other hand, the categorical phenotype results in a less powerful genetic analysis than if the phenotype is continuous. For an overview of latent class analysis methods and applications, see e.g. Hagenaars and McCutcheon (2002). Muthén (2001) put latent class analysis in a broader latent variable modeling framework. Muthén and Muthén (2000) discussed several applications including latent class analysis of antisocial behavior items in the National Longitudinal Survey of Youth (NLSY), a survey of individuals in early adulthood, where in addition to a normative class they found three classes of individuals with clearly different profiles of antisocial acts: property offense, person offense, and drug offense. Rasmussen et al. (2002) applied latent class analysis to DSM-IV ADHD symptoms in Australian twin data and found an 8-class solution where only some classes were congruent with DSM-IV subtypes. Further applications to alcohol data include Bucholz et al. (1996) for COGA data and Muthén (2001) for NLSY data, as well as Nestadt et al. (1994) for schizophrenia data.

The creation of categories as formed by diagnosis or latent class analysis often results in within-category heterogeneity in the form of variation in severity. This presents a violation of the latent class analysis assumption of independent items within class (conditional independence assumption). Allowing for such within-class variation in the form of one or more continuous factors results in a hybrid model. A hybrid model can provide both categorical and continuous information, for example using categories defined on a single factor dimension (measurement invariance across categories). The categories can be used for diagnosis and the dimension for genetic analysis. Factor mixture analysis for continuous variables has been described in McLachlan and Peel (2000). McLachlan, Do, and Ambroise (2004) applied factor mixture analysis to cluster analysis of microarray gene expression data, arguing that factor mixture analysis allows for biologically more meaningful clusters given the allowance for within-class correlation among the items. Lubke and Muthen (2005) presented applications to achievement data. Factor mixture analysis for categorical variables has been developed in Asparouhov and Muthen (2004). Applying factor mixture analysis in the context of dichotomous alcohol dependence and abuse criteria, Muthén and Asparouhov (2005) presented analyses using national survey data and Muthén, Asparouhov, and Rebollo (2006) presented heritability analyses using twin data. It is clear that these new models both fit these data better and produce different conclusions than methods that use only dimensional representations or only categorical representations.

In genetic analysis it is of interest to study a certain addictive behavior not only by itself but also in the context of other related addictive behaviors in order to examine commonalities and differences in genetic pathways. The epidemiological counterpart is comorbidity analysis. As an example, comorbidity of tobacco and alcohol disorders is often studied. Analytically, this raises the question of how to handle individuals who drink/smoke but do not exhibit any aspect of tobacco/alcohol dependence. For example, judging from a recent national survey of the U.S. general population that this paper is based on, over 70% of current drinkers exhibit none of the tobacco dependence criteria. An analysis of such data needs to carefully consider how to include model features that properly account for the large number of individuals exhibiting none of the criteria.

Section 2 presents the tobacco dependence data analyzed in this paper. Sections 3–5 give brief reviews of factor analysis, latent class analysis, and new hybrid models, and apply each technique to the tobacco dependence data. Section 6 concludes. As a caveat, it should be mentioned that although non-technical in nature, the general aim of this paper is to propose a methodology illustrated by an example, not to draw strong substantive conclusions.

2. Tobacco dependence data

To understand tobacco dependence, the natural heterogeneity exhibited in a general population sample is important. Such a sample, however, needs to be very large in order to produce sufficient numbers of individuals endorsing some of the criteria. The analyses in this paper use data on tobacco dependence from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC; Grant, Hasin, Chou, Stinson, & Dawson, in press). NESARC is a nationally representative face-to-face survey of 43,093 respondents of age 18 and older, carried out in 2001–2002. NESARC uses a complex survey design with stratification, 435 primary sampling units (PSUs), and oversampling of Black and Hispanic households. Within each household, one person was randomly selected for interview with young adults (18 to 24) oversampled at the rate of 2.25. The analyses to be presented concern three subsamples in order to represent the subsamples that might be considered in a comorbidity analysis. The first is the subsample of 26,946 past-year (current) drinkers (respondents who reported drinking five or more drinks on a single occasion one or more times in the past year). The second is the subsample of 11,118 past-year (current) smokers. The third is the subsample of 8552 individuals who are both current drinkers and current smokers.

Table 1
Tobacco dependence criteria

DSM-IV criteria	Sample prevalence (weighted means)		
	Current drinkers ($n=26{,}946$)	Current smokers ($n=11{,}118$)	Current drinkers and smokers ($n=8552$)
1. Tolerance	0.05	0.15	0.15
2. Withdrawal	0.22	0.67	0.67
3. Use in larger amounts/over longer period than intended	0.08	0.23	0.24
4. Persistent desire/unsuccessful efforts to cut down or quit	0.20	0.60	0.61
5. Great deal of time using/recovering	0.07	0.21	0.22
6. Important activities given up	0.03	0.08	0.08
7. Continued use despite emotional/physical problems	0.16	0.48	0.48
Proportion with no criteria fulfilled	0.73	0.18	0.18

The analyses focus on the 7 tobacco dependence criteria which were derived from a set of past year symptom item questions designed to operationalize DSM-IV. The instrument used to generate the symptom items and information about reliability are discussed in Grant et al. (2003, in press). In line with DSM-IV, a diagnosis of tobacco dependence is obtained when at least 3 out of the 7 criteria are fulfilled. The criteria are listed in Table 1 together with the prevalence of each criterion in each of the subsamples. The prevalence is almost the same for current smokers as for current drinkers and smokers because current smokers who are not current drinkers have prevalence similar to current smokers who are current drinkers.

All analyses in this paper are carried out using maximum-likelihood estimation in the Mplus computer program (Muthén & Muthén, 1998–2006; www.statmodel.com). The complex survey features of stratification, clustering, and sampling weights (oversampling) are taken into account in the parameter estimation and the standard error calculations (see Asparouhov, 2005).

3. Factor analysis (Item Response Theory Modeling)

Fig. 1 describes a unidimensional factor analysis (FA) model for a set of four items. This type of model uses a dimensional representation in the form of a continuous latent variable (factor). With categorical items, the analysis is often referred to as latent trait analysis, or item response theory (IRT) modeling, particularly when a single factor is used. For this situation, Fig. 1a shows how the probability of endorsing an item increases as a function of the factor *f*. Different items have different functions, typically represented by logistic regressions with different intercepts and slopes. Below the *f* axis is shown the distribution of the factor, typically assumed to follow a normal distribution. Fig. 1b shows the corresponding model diagram. The boxes at the top represent the four observed items and the circle represents the factor *f*, assumed to describe all the correlations among the items.

The FA model has a certain degree of alignment with the DSM-IV dependence classification of requiring at least 3 out of the 7 criteria in the sense that the factor values are generally strongly correlated

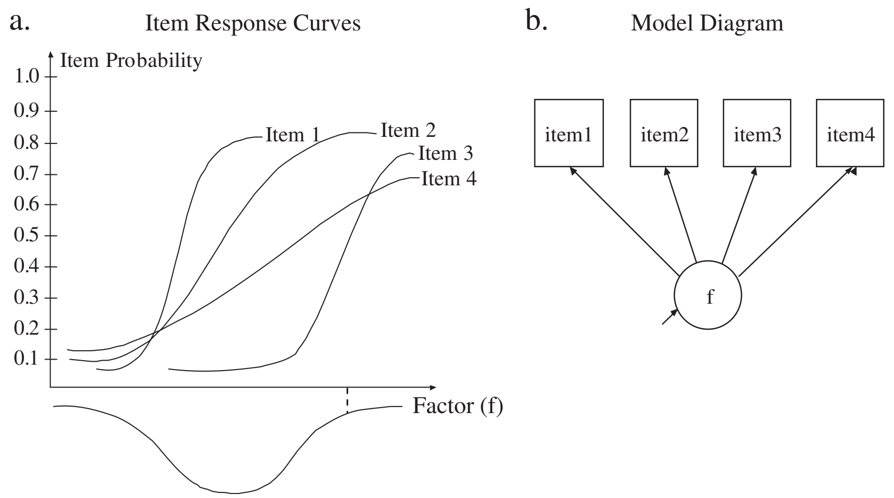

Fig. 1. Uni-dimensional factor analysis.

with the sum of the criteria fulfilled. Using the sum variable implicitly assumes a single dimension as in IRT. In using both the sum and the factor, a further important implicit assumption is that the same interpretation and metric is attached to this dimension in all parts of its range. This is an assumption of homogeneity of the population. Using the sum, however, is different than using the factor. Using the sum treats the criteria as equivalent in the sense that it does not matter which ones are fulfilled as long as at least 3 are fulfilled. In contrast, using the factor, the fulfillment of criteria that are less prevalent implies a higher factor value than the fulfillment of the same number of criteria that are more prevalent. Another key difference between FA and the DSM use of the sum is the DSM notion that at least 3 out of 7 criteria should be fulfilled for tobacco dependence. This notion of a threshold has no counterpart in FA. Fig. 1a shows a threshold on the factor distribution such that individuals above this threshold are likely to be classified as dependent. Such a threshold, however, cannot be determined by the factor model and is typically not discernable as a natural breakpoint when studying the estimated individual factor scores. Muthén (1996) discussed dimensional reporting using IRT models as alternatives to DSM-type classification.

3.1. Tobacco dependence results using FA

Table 2 shows model fit results for FA of the seven tobacco criteria in the three subsamples. Model 1 is the conventional, unidimensional IRT model. The estimated loadings are all significant for this model in all three subsamples. Model 2 uses two factors with the minimal number of restrictions applied in line with exploratory factor analysis. The results for models 1 and 2 clearly indicate that unidimensionality does not hold, but that two factors are needed to explain the correlations among the criteria. This is seen

Table 2
Model fit results

Model	Current drinkers ($n=26,946$)			Current smokers ($n=11,118$)			Current drinkers and smokers ($n=8552$)		
	logL	# par's	BIC	logL	# par's	BIC	logL	# par's	BIC
Factor analysis (IRT)									
1. FA 1f (IRT)	−40,949	14	82,042	−35,986	14	72,102	−27,671	14	55,470
2. FA 2f	**−40,759**	**20**	**81,721**	**−35,818**	**20**	**71,822**	**−27,550**	**20**	**55,280**
Latent class analysis									
3. LCA 2c	−41,428	15	83,008	−36,616	15	73,371	−28,139	15	56,415
4. LCA 3c	−40,430	23	81,094	−35,910	23	72,035	−27,609	23	55,427
5. LCA 4c	**−40,364**	**31**	**81,044**	**−35,809**	**31**	**71,908**	**−27,540**	**31**	**55,361**
6. LCA 5c	−40,322	39	81,042	−35,774	39	71,912	−27,514	39	55,380
7. LCA 6c	−40,290	47	81,060	−35,740	47	71,919	−27,486	47	55,398
Factor mixture analysis (IRT mixture)									
8. FMA 1f, zero class	−40,459	15	81,072	−35,957	15	72,055	−27,650	15	55,435
9. FMA 2f, zero class	−40,344	21	80,903	−35,804	21	71,804	−27,539	21	55,269
10. FMA 1f, 2c+zero class	**−40,316**	**30**	**80,937**	**−35,771**	**30**	**71,821**	**−27,512**	**30**	**55,295**

L—likelihood, # par—number of parameters, BIC—Bayesian information criterion, FA—factor analysis, LCA—latent class analysis, FMA—factor mixture analysis, f—factor, c—class.

Table 3
Factor analysis with two factors

	Current drinkers (n=26,946)		Current smokers (n=11,118)		Current drinkers and smokers (n=8552)	
	F1	F2	F1	F2	F1	F2
Estimated factor loadings						
1. Tolerance	**0.93**	0	**0.86**	0	**0.84**	0
2. Withdrawal	**0.41**	**0.58**	**0.47**	**0.38**	**0.52**	0.34
3. Larger amounts	**0.67**	**0.28**	**0.63**	**0.24**	**0.68**	0.18
4. Cut down	**0.17**	**0.78**	0.09	**0.61**	0.21	0.51
5. Time using	**0.76**	0.16	**0.66**	0.10	**0.69**	0.07
6. Give up	0.12	**0.70**	**0.15**	**0.46**	**0.22**	0.41
7. Continued use	0	**0.96**	0	**0.87**	0	**0.90**
Estimated factor variances and correlation						
F1	**21.11**		**9.58**		**7.83**	
F2	**0.89**	**35.92**	**0.67**	**9.79**	**0.65**	13.44

in the considerably larger (better) loglikelihood value for model 2 compared to model 1, where model 2 uses only 6 more parameters. It is also seen in model 2 having a considerably smaller (better) BIC value. BIC is a measure that combines the loglikelihood value, which we want to maximize, with the number of parameters, which we want to keep at a minimum.

The model 2 FA estimates for the three subsamples are shown in Table 3. The three subsamples have remarkably similar factor loading patterns. Bolded values indicate parameter estimates significant at the 5% level. A consistent pattern across the three subsamples is that the second factor appears to mainly relate to criteria 4 and 7, which have large loadings for factor 2 but not for factor 1. Criterion 4 is “Persistent desired/unsuccessful efforts to cut down or quit” and criterion 7 is “Continued use despite emotional/physical problems”. This loading pattern might imply that factor 2 has to do with unsuccessful attempts at smoking cessation. The factor model therefore appears interpretable and stable across samples, although the factor loadings for factor 2 are decreasing in size when going from

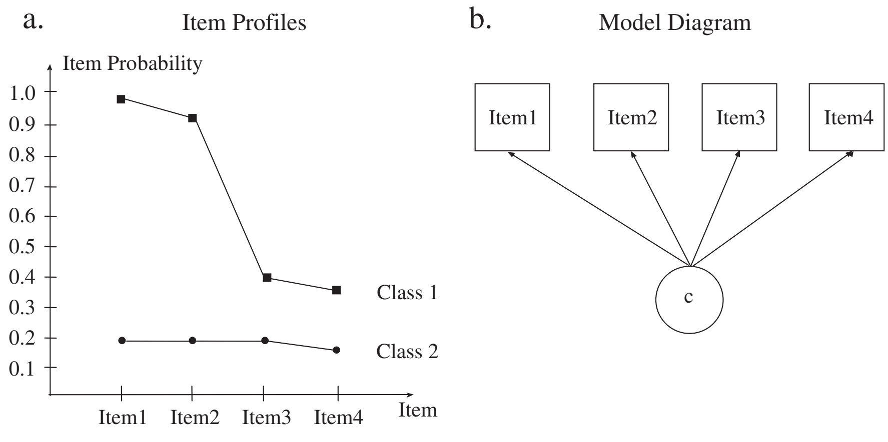

Fig. 2. Latent class analysis.

current drinkers to current smokers to current drinkers and smokers. Before putting trust in this model and discussing it in the DSM tobacco dependence context, however, alternative models need to be considered, in particular modifications using hybrid modeling that allow heterogeneity among subjects.

4. Latent class analysis

Latent class analysis (LCA) is used to uncover heterogeneous groups of individuals, thereby having the same goal as cluster analysis. Fig. 2 describes a latent class analysis model. Fig. 2a considers analysis results in terms of item profiles for the four items listed along the x-axis. The picture shows two latent classes (unobserved groups) of individuals who are homogeneous within classes and different across classes. Class 2 has low endorsement probabilities for all four items and the classes are further differentiated by class 1 having considerably higher endorsement probabilities for items 1 and 2. It is this type of class differentiation that lends an interpretation to the behavior that characterizes individuals in the two classes. In a general population sample, the prevalence is the largest for the normative class 2.

Fig. 2b shows a corresponding model diagram for LCA. The circle contains the latent class variable, which in this example has two categories. Analogous to the FA model of Fig. 1, the latent variable explains the dependence among the items. Within each class the items are independent (this is referred to as the conditional independence assumption).

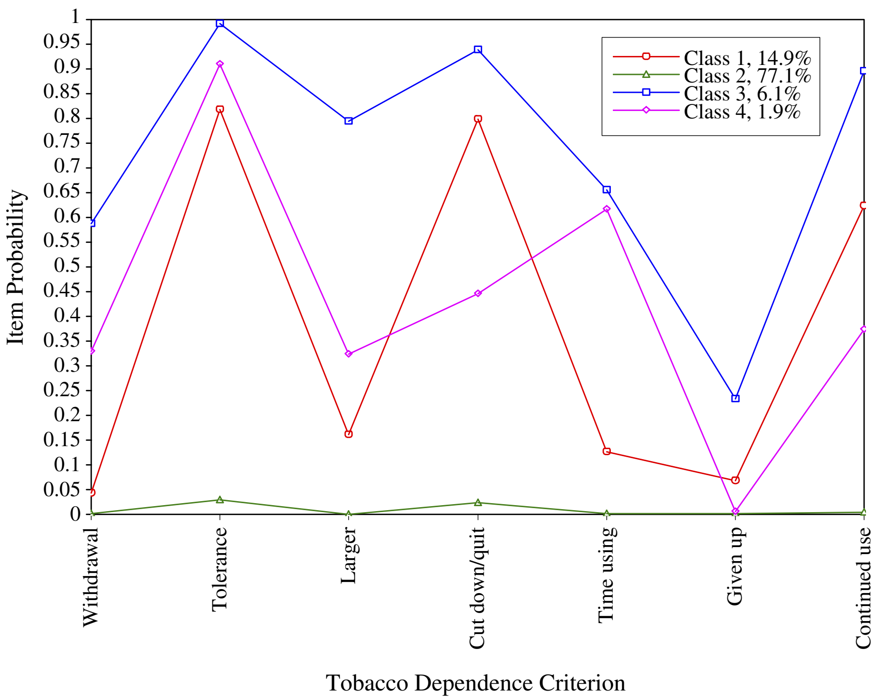

Fig. 3. Item profiles for latent class model for current drinkers.

LCA has the advantage over FA that a categorical representation is obtained. In this sense, LCA is well aligned with the DSM diagnostic approach. The latent class analysis may not, however, lead to the same diagnostic algorithm as DSM (at least 3 out of 7 criteria fulfilled). In contrast to such an algorithm, LCA places different importance on different items in the classification. This is similar to FA giving different weights to different items in estimating factor values.

4.1. Tobacco dependence results using LCA

Table 2 shows model fitting results for LCA with 2–6 classes (models 3–7). Judging from both the improvement in loglikelihood value as the number of classes is increased and from the BIC value, the current smoker and current drinker and smoker subsample analyses both suggest model 5 as the best model. Model 5 is a reasonable choice also for the subsample of current drinkers.

A comparison of the best LCA model 5 with the best FA model 2 shows that model 5 gives an improvement in the loglikelihood value. Given the 11 added parameters of model 5, however, the loglikelihood improvement is only large for the current drinker sample and the BIC values for the two smaller subsamples favor model 2. Because of these findings, LCA does not provide a superior representation of these data than obtained by FA.

Fig. 3 shows the model 5 LCA estimates of the item profiles. The highest class (class 3) contains 6% of the current drinker sample. The normative class (class 2) contains 77%. Class 2 includes the 73% who have a response pattern of zeros for all the 7 criteria. The two largest non-normative classes, class 1 and class 3, have item profiles that are similar in shape but different in height, indicating an ordering of these two classes in terms of severity. The ordering of classes suggests that it might be useful to add a dimensional aspect to the model. As will be shown, such hybrid models will clearly supersede model 5, and conventional LCA will not be further discussed here.

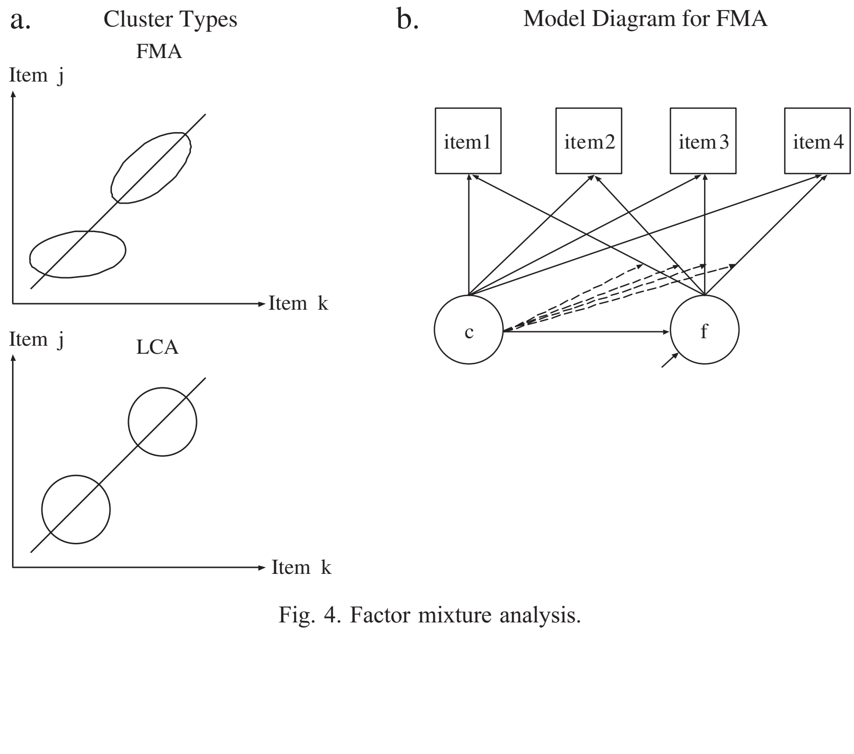

Fig. 4. Factor mixture analysis.

5. Factor mixture analysis (IRT mixture modeling)

Fig. 4 shows a hybrid model in the form of a factor mixture analysis (FMA) model. Fig. 4a shows two items that are strongly related as indicated by the positive slope of the lines. Ellipsoids and circles show the shapes of the within-class distributions of the two classes for FMA and LCA, respectively. In contrast with FMA, LCA shows independence among the items within the classes. In line with FA, the FMA dependence among the items arises due to the items being influenced by a factor within each class. The factor represents within-class variation among individuals that in a tobacco dependence context might be thought of as within-class variation in tobacco dependence severity. Because of the factor influence on the items, the items are correlated within class. In contrast, LCA allows for no such within-class variation and forms clusters of individuals defined as having independence among the items (the conditional independence assumption). Substantively more meaningful clusters might be found when allowing within-class correlations among items and this will lead to a different latent class formation. A more limited form of relaxing the conditional independence assumption in LCA for only some items in only one class was proposed by Qu, Tan, and Kutner (1996).

The arrows from the latent class variable *c* in Fig. 4b indicate that the item probabilities in FMA vary as a function of latent class membership as in latent class analysis. In addition, as indicated by broken arrows, the item factor loadings are allowed to vary across classes. In this way, different items may be more representative of the severity dimension in the different classes. Furthermore, the factor variances are allowed to vary across classes indicating class differences in heterogeneity with respect to the severity. FMA can be seen as a generalization of either FA or LCA. Generalizing FA, it is natural to specify class-invariance of the measurement parameters of the conditional item probabilities and factor loadings, letting item probabilities vary across classes only as a function of class-varying factor means and variances. Generalizing LCA, it is natural to allow fully class-varying measurement parameters. In the current application, the latter approach was found to fit the data better.

Fig. 5 illustrates a particular form of FMA that allows for a non-parametric representation of the factor distribution. This can be a useful alternative to the assumption of normality in some applications. Fig. 5a shows a non-normal factor distribution discretized into four points. Fig. 5b shows that this distribution can be achieved by the use of latent classes that have different factor means. The factor means represent the *x*-axis values in Fig. 5a and the latent class probabilities represent the heights of the bars in Fig. 5a.

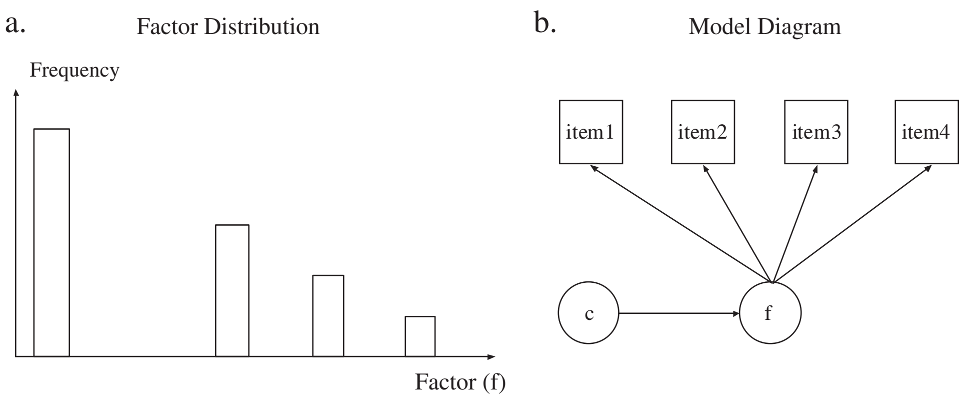

Fig. 5. Non-parametric representation of the factor distribution.

Fig. 5b indicates that this modeling is used for class-invariant item parameters, but a combination of the class-varying item parameter modeling of Fig. 4 is also possible in the non-parametric framework. This type of modeling will not be presented in this paper but did give results that were similar to those of FMA with normal factors.

5.1. Tobacco dependence results using FMA

Because LCA did not give a real improvement over factor analysis, it is of interest to consider if the fit improves by using hybrid modeling. The flexible allowance of heterogeneity in the sample that FMA provides is an attractive avenue. Table 2 shows model fit results for different versions of factor mixture models (models 8–10). Both mixture models introducing a zero class as well as more general FMA models are considered. In the general FMA model 10, the two non-zero classes are allowed fully class-varying measurement parameters.

As mentioned in Introduction, the percentage having an all zero response pattern is 73% for the current drinkers and 18% for the current smokers and current drinkers and smokers. The addition of a zero class, that is a class in which the probability is one of not fulfilling any of the criteria, adds only 1 parameter; the probability of being in the zero class. This is similar to, but not the same as, deleting individuals that have the all zero response pattern. Deleting individuals with the all zero pattern is an inferior approach because the estimated probability of being in the zero class may be different from the sample proportion having a response pattern of all zeros. For example, fitting the model to data may result in some response patterns with only a few criteria fulfilled being viewed as sufficiently similar to the all zero pattern, classifying such patterns into the zero class. Also, including individuals with the all zero response pattern is important when including covariates, making it possible to study the prediction of membership in a zero class. This is useful in comorbidity analyses, where the correlates of smoking and drinking are of interest.

Returning to the unidimensional FA model 1 in Table 2, model 8 investigates whether a zero class improves on the model. This is clearly the case for all three subsamples. Across all three subsamples,

Table 4
Factor analysis with two factors and a zero class

	Current drinkers ($n=26{,}946$)		Current smokers ($n=11{,}118$)		Current drinkers and smokers ($n=8552$)	
	F1	F2	F1	F2	F1	F2
Estimated factor loadings						
1. Tolerance	**0.83**	0	**0.85**	0	**0.83**	0
2. Withdrawal	**0.54**	0.27	**0.50**	**0.31**	**0.55**	0.26
3. Larger amounts	**0.70**	0.15	**0.63**	0.22	**0.70**	0.14
4. Cut down	0.23	0.42	0.09	0.54	0.24	0.41
5. Time using	**0.69**	0.06	**0.65**	0.09	**0.69**	0.06
6. Give up	**0.25**	0.36	**0.17**	**0.43**	**0.25**	0.36
7. Continued use	0	**0.92**	0	**0.86**	0	**0.93**
Estimated factor variances and correlation						
F1	**7.16**		**8.82**		**7.13**	
F2	0.59	18.28	**0.64**	9.36	0.59	21.04

model 8 is not, however, uniformly better than the 2-factor model 2. As an alternative, model 9 uses both two factors and a zero class and is clearly much better than both model 2 and model 8. The model 9 estimates are shown in Table 4.

Table 4 shows that factor 2 loses its significance when a zero class is introduced. It is questionable whether or not the second factor is needed. At the same time, the unidimensional model with a zero class, model 8, fits considerably worse. This dilemma can be resolved by using a more general factor mixture model, model 10, which includes not only a zero class but also allows for two different non-zero classes. Model 10 adds 9 parameters to model 9 and although BIC is not better, the loglikelihood improvement is substantial. In addition, model 10 avoids the model 9 problem of an ill-defined second factor.

The item profiles for model 10 are shown in Fig. 6 for the subsample of current drinkers. The item profile plots for the other subsamples are very similar. The profiles for the two non-zero classes are distinguished by particularly large differences for the 4th and 7th criteria. This was also found in the LCA. In line with this, the high non-zero class contains individuals for whom unsuccessful attempts to quit smoking and continued use despite problems are key characteristics. Among the subsample of current smokers the class percentages for the high non-zero class (class 1), low non-zero class (class 2), and zero class (class 3) are 58%, 30%, and 12%, respectively. These percentages differ across the three subsamples as would be expected given the differences in percentages of individuals with the all zero response pattern, but the percentage of individuals in the high non-zero class among those who are in the non-zero classes is rather stable: 68% for the current drinkers, 66% for the current smokers, and 68% for the current drinkers and smokers.

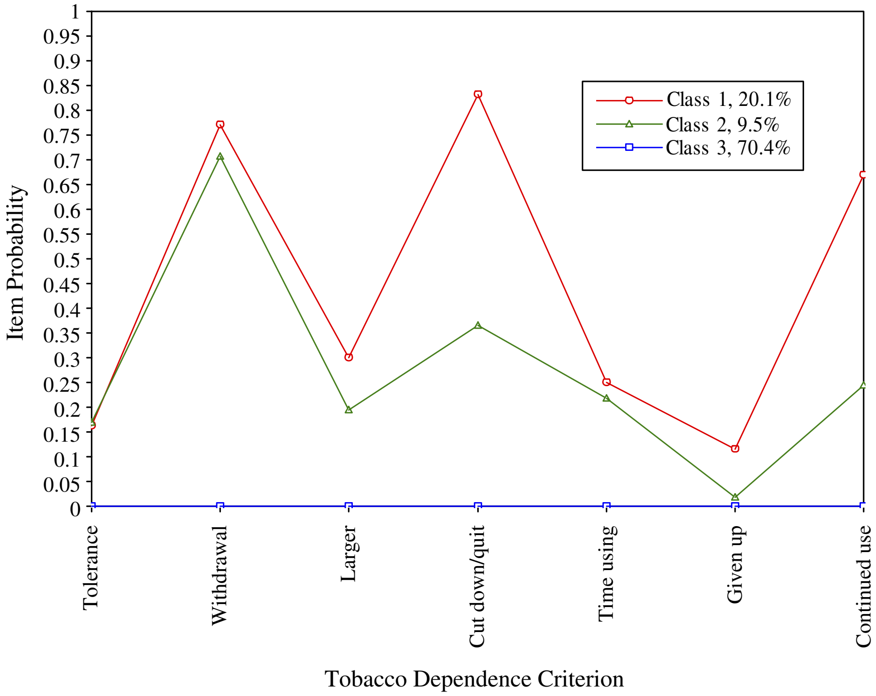

Fig. 6. Item profiles for factor mixture model for current drinkers.

As an aside, it may be noted that for all three subsamples the estimated class probability for the zero class is not the same as the percentage having the all zero response pattern, but is a few percentage points lower. This is because for these data the individuals with the all zero response pattern do not have probability 1.0 of being in the zero class, but have a small probability of being in the other two classes (see also the pattern classification in Table 6 below). This is an illustration of the fact that the classification in mixture models is not 100% clearcut.

It should be noted that the model 10 item profiles of Fig. 6 correspond to the marginal probability of fulfilling a criterion. The two non-zero classes have within-class variation in the endorsement probabilities due to the variation in the factor for these classes. For example, an individual with factor value above the mean has a higher endorsement probability than shown in the figure. As mentioned earlier, the factor variation may be seen as a severity variation. In line with Fig. 4b, model 10 allows class differences in the two non-zero classes for both intercepts/threshold parameters and for loadings and these differences are significant. This implies that it is not the same single dimension that underlies the items for all individuals. The estimated loadings for the two non-zero classes are given in Table 5. The loading patterns show strong similarity across the three subsamples and clear differences between the two non-zero classes. For the high non-zero class, the 4th criterion (“Persistent desired/unsuccessful efforts to cut down or quit”) has an insignificant loading. For the low non-zero class, the 6th criterion (“Important activities given up”) has an insignificant loading and the loading for the 2nd criterion (“Withdrawal”) is relatively low. Taken together with the latent class profile differences across the two non-zero classes, this loading pattern difference may be helpful to substantive experts in better understanding the tobacco dependence phenotype.

To further shed light on the FMA results (model 10), estimated class membership probabilities and estimated factor scores can also be studied for each of the response patterns. This also makes it possible to connect the FMA results with the DSM-IV classification based on having at least 3 out of the 7 criteria fulfilled. Table 6 gives this information for current drinkers and response patterns that have frequency of at least 10. The results are very similar for the other two subsamples. The table has three parts, one for each latent class. Within each class, the response patterns are ordered with respect to their estimated factor score value. Probability estimates for class membership are also printed (“cprob”). Such estimates are typically referred to as posterior probability estimates in the mixture literature.

The listed response patterns of the high class (class 1) all have a markedly higher-class probability estimates for class 1 than for the other two classes. The pattern with all 1’s suggests that class 1 is the

Table 5
Factor loadings for the factor mixture model

	Current drinkers (n=26,946)		Current smokers (n=11,118)		Current drinkers and smokers (n=8552)	
	High class	Low class	High class	Low class	High class	Low class
1. Tolerance	**0.86**	**0.78**	**0.88**	**0.78**	**0.86**	**0.78**
2. Withdrawal	**0.83**	**0.43**	**0.81**	**0.43**	**0.83**	**0.43**
3. Larger amounts	**0.72**	**0.88**	**0.72**	**0.86**	**0.72**	**0.88**
4. Cut down	0.30	**0.70**	0.21	**0.67**	0.30	**0.70**
5. Time using	**0.71**	**0.66**	**0.70**	**0.62**	**0.71**	**0.66**
6. Give up	**0.49**	−0.13	**0.45**	0.08	**0.49**	−0.13
7. Continued use	**0.68**	**0.43**	**0.72**	**0.36**	**0.68**	**0.43**

Table 6
Response pattern classification using factor mixture analysis for current drinkers

Frequency	Tobacco items							Factor score	Class probabilities			
	Tolerance	Withdrawal	Larger amounts	Cut down	Time using	Give up	Continued use		Cprob1	Cprob2	Cprob3	Total
Class one												
115	1	1	1	1	1	1	1	5.808	0.994	0.006	0	7
309	1	1	1	1	1	0	1	4.725	0.78	0.22	0	6
43	1	1	1	1	0	1	1	4.212	0.99	0.01	0	6
21	1	1	0	1	1	1	1	4.172	0.993	0.007	0	6
25	1	1	1	0	1	0	1	4.046	0.697	0.303	0	5
156	1	1	1	1	0	0	1	3.456	0.799	0.201	0	5
79	1	1	0	1	1	0	1	3.323	0.865	0.135	0	5
40	0	1	1	1	1	1	1	3.296	0.992	0.008	0	6
15	1	1	0	1	0	1	1	2.896	0.986	0.014	0	5
19	1	1	1	0	0	0	1	2.893	0.634	0.366	0	4
219	0	1	1	1	1	0	1	2.653	0.878	0.122	0	5
24	1	1	0	0	1	0	1	2.551	0.602	0.398	0	4
24	0	1	1	0	1	0	1	2.227	0.746	0.254	0	4
92	1	1	0	1	0	0	1	2.153	0.833	0.167	0	4
48	0	1	1	1	0	1	1	2.06	0.993	0.007	0	5
40	0	1	0	1	1	1	1	2.018	0.99	0.01	0	5
363	0	1	1	1	0	0	1	1.393	0.929	0.071	0	4
256	0	1	0	1	1	0	1	1.31	0.91	0.09	0	4
53	0	1	1	0	0	0	1	1.039	0.8	0.2	0	3
11	0	1	0	1	1	1	0	0.812	0.879	0.121	0	4
55	0	1	0	0	1	0	1	0.796	0.617	0.383	0	3
91	0	1	0	1	0	1	1	0.686	0.988	0.012	0	4
132	0	1	1	1	0	0	0	0.669	0.666	0.334	0	3
110	0	1	0	1	1	0	0	0.526	0.535	0.465	0	3
34	0	1	0	0	0	1	1	0.135	0.868	0.132	0	3
952	0	1	0	1	0	0	1	−0.069	0.936	0.064	0	3
26	0	0	0	1	1	0	1	−0.412	0.78	0.22	0	3
28	0	0	1	1	0	0	1	−0.419	0.858	0.142	0	3
279	0	1	0	0	0	0	1	−0.506	0.596	0.404	0	2
24	0	1	0	1	0	1	0	−0.6	0.904	0.096	0	3
24	0	0	0	1	1	0	0	−0.896	0.533	0.467	0	2
698	0	1	0	1	0	0	0	−1.043	0.728	0.272	0	2
34	0	0	1	1	0	0	0	−1.199	0.726	0.274	0	2
15	0	0	0	1	0	1	1	−1.478	0.973	0.027	0	3
143	0	0	0	0	0	0	1	−2.15	0.556	0.444	0	1
271	0	0	0	1	0	0	1	−2.23	0.931	0.069	0	2
15	0	0	0	1	0	1	0	−2.818	0.915	0.085	0	2
611	0	0	0	1	0	0	0	−3.63	0.88	0.12	0	1
Class two												
52	1	1	1	1	1	0	0	3.698	0.139	0.861	0	5
14	1	1	1	0	1	0	0	2.865	0.083	0.917	0	4
50	1	1	1	1	0	0	0	2.802	0.206	0.794	0	4

Table 6 (*continued*)

Frequency	Tobacco items							Factor score	Class probabilities			
	Tolerance	Withdrawal	Larger amounts	Cut down	Time using	Give up	Continued use		Cprob1	Cprob2	Cprob3	Total
Class two												
28	1	1	0	1	1	0	0	2.183	0.239	0.761	0	4
14	1	1	1	0	0	0	0	2.153	0.094	0.906	0	3
64	0	1	1	1	1	0	0	2.139	0.406	0.594	0	4
20	0	1	1	0	1	0	0	1.715	0.206	0.794	0	3
20	1	1	0	0	1	0	0	1.421	0.064	0.936	0	3
20	1	1	0	0	0	0	1	1.404	0.463	0.537	0	3
44	1	1	0	1	0	0	0	1.344	0.273	0.727	0	3
11	1	0	0	1	0	0	0	0.745	0.143	0.857	0	2
34	0	1	1	0	0	0	0	0.598	0.365	0.635	0	2
21	1	1	0	0	0	0	0	0.568	0.055	0.945	0	2
86	0	1	0	0	1	0	0	0.057	0.139	0.861	0	2
13	1	0	0	0	0	0	0	−0.04	0.026	0.974	0.001	1
25	0	0	0	0	1	0	0	−0.828	0.136	0.864	0	1
15	0	0	1	0	0	0	0	−0.859	0.457	0.542	0	1
743	0	1	0	0	0	0	0	−1.328	0.173	0.827	0	1
16	0	1	0	0	0	1	0	−1.329	0.349	0.651	0	2
10	0	0	0	0	0	1	0	−2.833	0.327	0.673	0	1
Class three												
19998	0	0	0	0	0	0	0	−0.099	0.011	0.022	0.967	0

likely class membership with almost certainty (class probability 0.994). This pattern also has the highest factor score estimate. The least well-classified pattern for class 1 is the last pattern listed, but it still discriminates rather well between the classes in that it has probability 0.88 for class 1 as compared to 0.12 for class 2. It is interesting to note that the sum of the criteria fulfilled varies from 1 to 7 for the patterns in class 1. For class 2, this range is 1–5. This reflects the key fact that it is not the number of criteria fulfilled that determines class membership, but rather the type of item profile (see Fig. 6). Within class, however, the factor score estimates correlate rather well with the number of criteria fulfilled. Note that the factor score estimates are not comparable across classes because they represent different dimensions due to measurement parameter differences.

6. Conclusions

This paper has illustrated new hybrid latent variable models that are promising for phenotypical analyses. The hybrid models combine features of dimensional and categorical analyses seen in the conventional techniques of factor analysis and latent class analysis. The hybrid models are typically seen to fit data better than conventional models of factor analysis (IRT) and latent class analysis. The paper focused on the analysis of categorical items, which presents especially challenging analyses with hybrid models and has recently been made practical in the Mplus program. The methods are, however, of general applicability in that outcomes that are continuous, censored, counts, or combinations thereof can also be handled. Although this paper focused on cross-sectional data, hybrid modeling is also available for longitudinal data where even richer analysis possibilities are presented (for growth mixture modeling applications to alcohol, externalizing behavior, and mathematics achievement, see Muthén, 2004; Muthén et al., 2002; Muthén & Shedden, 1999). There are several implications of the availability of these new techniques. One question is how this might affect attempts at revising DSM-IV diagnoses. Another question is how this might affect investigations of new measures of different phenotypes.

Table 6 response pattern classification for current drinkers shows the difference between the classification reached by using a factor mixture model (IRT mixture model) and the DSM-IV classification. Consider the most severe class of the factor mixture model. For the current drinkers of Table 6, this class contains 20%. For current smokers and current drinkers and smokers, this class contains 58% and 61%, respectively. Table 6 shows that many individuals who do not fulfill at least 3 out of the 7 criteria are in class 1. At the same time, several individuals classified into class 2 do fulfill at least 3 out of the 7 criteria. The factor mixture analysis does not suggest that the high class 1 should necessarily be seen as the critical group to be reported as "tobacco dependent". Indeed, 58% of current smokers would seem a high number (the DSM-IV prevalence of tobacco dependence is, however, 45% among current smokers). An alternative is to use a severity score in the form of the estimated factor scores within the classes. One can take either a categorical approach and report the percentage above a certain percentile on the factor score distribution or take a dimensional approach and report percentiles in line with suggestions in Muthén (1996). The important point is that factor mixture modeling uncovers a heterogeneous latent variable structure that fits the data well and that sheds more light on the tobacco dependence phenotype. This is useful for guiding substantive experts in their understanding and decisions.

To fully benefit from the new analysis possibilities offered by factor mixture modeling (IRT mixture modeling), the design of the measurement instrument should be carried out with an eye toward utilizing the new modeling flexibility in substantively meaningful ways. When new measures of a phenotype are explored the choice of such measures should be done while having in mind that both categories and dimensions can be represented. Categories (latent classes) are better captured when measures are designed to capture specific hypothesized item profiles corresponding to distinct behaviors. Dimensions (factors) are better captured when measures are designed to capture severity variations. In both regards, the nature and scale of the new measures are of less importance than the types of latent constructs they attempt to capture. Furthermore, new phenotypical studies should be both exploratory and confirmatory in nature. Despite solid theoretical validity, new measures are often invalid because respondents perceive them differently than intended. New measures need to be investigated and revised in a series of pilot studies that provide information on how respondents actually perceive the measures. This leads to larger studies where hypothesized models can be confirmed. New hybrid models hold great promise but also present a great challenge to using them well.

Acknowledgements

Invited contribution to the special issue "Novel Approaches to Phenotyping Drug Abuse" forthcoming in the Annual Review of Addictions Research. The research of the first author was supported under grant K02 AA 00230 from NIAAA.

References

Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Structural Equation Modeling*, *12*(3), 411–434.

Asparouhov, T., & Muthen, B. (2004). Maximum-likelihood estimation in general latent variable modeling. *Draft*.

Bucholz, K. K., Heath, A. C., Reich, T., Hesselbrock, V. M., Kramer, J. R., Nurnberger, J. I., et al. (1996). Can we subtype alcoholism? A latent class analysis of data from relatives of alcoholics in a multi-center family study of alcoholism. *Alcoholism, Clinical and Experimental Research*, *20*, 1462–1471.

Duncan-Jones, P., Grayson, D. A., & Moran, P. A. P. (1986). The utility of latent trait models in psychiatric epidemiology. *Psychologie und Medizin*, *16*, 391–405.

Gallo, J. J., Anthony, J. C., & Muthén, B. (1994). Age differences in the symptoms of depression: a latent trait analysis. *Journal of Gerontology*, *49*, 251–264.

Grant, B. F., Dawson, D. A., Stinson, F. S., Chou, P. S., Kay, W., & Pickering, R. (2003). The alcohol use disorder and associated disabilities interview schedule-IV (AUDADIS=IV): Reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug and Alcohol Dependence*, *71*, 7–16.

Grant, B. F., Hasin, D. S., Chou, P. S., Stinson, F. S., & Dawson, D. A. (in press). Nicotine dependence and psychiatric disorders in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry*.

Hagenaars, J. A., & McCutcheon, A. (2002). *Applied latent class analysis*. Cambridge: Cambridge UP.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston, MA: Kluwer Nijhoff.

Harford, T., & Muthén, B. (2001). The dimensionality of alcohol abuse and dependence: a multivariate analysis of DSM-IV symptom items in the National Longitudinal Survey of Youth. *Journal of Studies on Alcohol*, *62*, 150–157.

Krueger, R. F., Nichol, P. E., Hicks, B. M., Markon, K. E., Patrick, C. J., Iacono, W., et al. (2004). Using latent trait modeling to conceptualize an alcohol problems continuum. *Psychological Assessment*, *16*, 107–119.

Lubke, G., & Muthen, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, *10*, 21–39.

McLachlan, G. J., Do, K. A., & Ambroise, C. (2004). *Analyzing microarray gene expression data*. New York: Wiley & Sons.

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

Muthén, B. (1989a). Latent variable modeling in heterogeneous populations. Presidential address to the Psychometric Society, July 1989. *Psychometrika*, *54*, 557–585.

Muthén, B. (1989b). Dichotomous factor analysis of symptom data. In W. O. Eaton, & G. W. Bohrnstedt (Eds.), *Latent variable models for dichotomous outcomes: Analysis of data from the Epidemiological Catchment Area Program*. *Sociol Methods Res.*, *vol. 18* (pp. 19–65).

Muthén, B. (1992). Latent variable modeling in epidemiology. *Alcohol Health and Research World*, *16*, 286–292.

Muthén, B. (1996). Psychometric evaluation of diagnostic criteria: Application to a two-dimensional model of alcohol abuse and dependence. *Drug and Alcohol Dependence*, *41*, 101–112.

Muthén, B. (2001). Latent variable mixture modeling. In G. A. Marcoulides, & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 1–33). Mahwah, NJ: Lawrence Erlbaum Associates.

Muthén, B. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, *29*, 81–117.

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Newbury Park, CA: Sage.

Muthén, B., & Asparouhov, T. (in press). Should substance use disorders be considered as categorical or dimensional? Addiction.

Muthén, B., Asparouhov, T., & Rebollo, I. (2006). Advances in behavioral genetics modeling using Mplus: Applications of factor mixture modeling to twin data. Forthcoming in the special issue “Advances in statistical models and methods”. *Twin Research and Human Genetics*.

Muthén, B., Brown, C. H., Masyn, K., Jo, B., Khoo, S. T., Yang, C. C., et al. (2002). General growth mixture modeling for randomized preventive interventions. *Biostatistics*, *3*, 459–475.

Muthén, B., Grant, B., & Hasin, D. (1993). The dimensionality of alcohol abuse and dependence: factor analysis of DSM-III and proposed DSM-IV criteria in the 1988 National Health Interview Survey. *Addiction*, *88*, 1079–1090.

Muthén, B., & Muthén, L. (2000). Integrating person-centered and variable-centered analysis: Growth mixture modeling with latent trajectory classes. *Alcoholism, Clinical and Experimental Research*, *24*, 882–891.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, *55*, 463–469.

Muthén, L. K., & Muthén, B. (1998–2006). *Mplus user's guide. version 4*. Los Angeles, CA: Muthén & Muthén.

Nestadt, G., Hanfelt, J., Liang, K. Y., Lamacz, M., Wolyniec, P., & Pulver, A. E. (1994). An evaluation of the structure of schizophrenia spectrum personality disorders. *Journal of Personality Disorders*, *8*, 288–298.

Qu, Y., Tan, M., & Kutner, M. H. (1996). Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics*, *52*, 797–810.

Rasmussen, E. R., Neuman, R. J., Heath, A. C., Levy, F., Hay, D. A., & Todd, R. D. (2002). Replication of the latent class structure of Attention-Deficit/Hyperactivity Disorder (ADHD) subtypes in a sample of Australian twins. *Journal of Child Psychology and Psychiatry*, *43*, 1018–1028.