Models and strategies for factor mixture analysis:

Two examples concerning the structure underlying psychological disorders

Shaunna L. Clark and Bengt Muthén

University of California, Los Angeles

Jaakko Kaprio

University of Helsinki, Helsinki, Finland

Brian M. D'Onofrio, Richard Viken & Richard J. Rose

Indiana University, Bloomington

Susan L. Smalley

University of California, Los Angeles

Author Note

Correspondence concerning this article should be addressed to Shaunna L. Clark, Department of Education, University of California, Los Angeles, CA 90025. Telephone: (310) 382-0296, Email: shclark@ucla.edu

Abstract

The factor mixture model (FMM) uses a hybrid of both categorical and continuous latent variables. The FMM is a good model for the underlying structure of psychopathology because the use of both categorical and continuous latent variables allows the structure to be simultaneously categorical and dimensional. While the conceptualization of the FMM has been explained in the literature, the use of the FMM is still not prevalent. One reason is that there is little research about how such models should be applied in practice and, once a well fitting model is obtained, how it should be interpreted. In this paper, the FMM will be explored by studying two real data examples: conduct disorder and attention-deficit hyperactivity disorder. By exploring these examples, this paper aims to explain the different formulations of the FMM, the various steps in building a FMM, as well as how to decide between a FMM and alternative models.

Models and strategies for factor mixture analysis:

Two examples concerning the structure underlying psychological disorders

In recent years there has been a debate in the psychological literature about whether the underlying structure of psychological disorders, such as conduct disorder and attention-deficit hyperactivity disorder, is categorical or dimensional. In the categorical view, psychological disorders are represented by diagnostic categories that indicate whether a person is affected or unaffected by a disorder or whether a person has a specific subtype. This has been the predominant view of psychopathology because it has the advantage of meeting clinical needs and the needs of reporting for health care planners and insurance companies (Muthén, 2006). Alternatively, psychological disorders are considered dimensional in nature and are represented as a continuous distribution, with each individual having some amount of the disorder. The advantage of the dimensional view is that each disorder can be represented as a quantitative score, or scores, which provides a more precise measure of functioning and more power for further statistical analyses than categorical outcomes (Muthén, 2006).

In the psychometric literature, each of these views has a counterpart. The categorical view can be represented by latent class analysis, which models unobserved heterogeneity in a sample through the use of categorical latent variables called latent classes. In this analysis, individuals are grouped into their most likely class based on their observed symptoms so that latent classes can then be interpreted as diagnostic categories or subtypes. The problem with latent class analysis, and the categorical approach to psychopathology, is that the categories do not consider the range in severity and impairment within and across diagnostic classes.

The dimensional view of psychological disorders has its counterpart in factor analysis. Here, continuous latent variables, called factors, are used to model the correlations among the

symptoms. Each of these factors represents an underlying dimension of the disorder. One drawback of this approach is that there is, generally, no easy way to classify individuals into groups, which, as stated earlier, is a clinical necessity and required by insurance companies and other reporting agencies.

One solution to the debate, proposed by Muthén (2006), is the factor mixture model (FMM). The FMM is a hybrid of both categorical and continuous latent variables, which allows the underlying structure to be simultaneously categorical and dimensional. The structure is considered categorical because the model allows for the classification of individuals into diagnostic groups through the use of latent class variables. The structure is also considered dimensional because once individuals are classified into groups, the FMM allows for variation in the severity of the disorder through the use of continuous latent variables. This approach is useful because it does not have the limitations of the two conventional representations of psychopathology.

While the studies that have introduced the factor mixture model to the psychological literature have explained the conceptualization of the FMM, the use of the FMM is still not prevalent. One reason for this is that even though the concept of the model has been explained, there is little research about how these models should be applied in practice and, once a well fitting model is obtained, how it should be interpreted. This paper seeks to remedy these shortcomings.

This paper aims to explain the different variations of the FMM. Also, this paper will elucidate the various steps in building a FMM, as well as, how to decide between a FMM and alternative models. The FMM will be explored in detail by studying two real data examples: conduct disorder and attention-deficit hyperactivity disorder.

The first section of this paper explains the latent class, factor analytic, and factor mixture model in more technical detail. The second section focuses on the model building process and on how to compare among the different types of models. The third section presents the real data examples. The final section discusses the utility and feasibility of the factor mixture model in practical settings. All analyses in this paper were carried out using M*plus* V5.1 (Muthén & Muthén, 1998-2008). In order to elucidate the factor mixture model, sample inputs for each model variation are available in the Appendix.

Background

Latent Class Analysis

The latent class analysis (LCA) model, introduced by Lazarsfeld and Henry (1968), is used to identify subgroups, or classes, of a study population. A diagram of an example of a latent class analysis model is shown in Figure 1. There are two major concepts depicted in Figure 1, the observed outcomes or items that define the class and latent class itself. These can be seen in Figure 1 as u_1-u_r, and c, respectively. The boxes, u_1 to u_r, represent the observed response items or outcomes. The outcomes in an LCA model can be categorical, continuous, count, censored, or nominal, though this paper will specifically focus on dichotomous, categorical items. The circle with the letter c in the middle is the unordered, categorical latent class variable with K classes. The arrows pointing from the latent class variable to the boxes above indicate that those items are measuring the latent class variable. This means that class membership is based on the observed response pattern of items. An important assumption, called the conditional or local independence assumption, implies that the correlation among the observed outcomes is explained by the latent class variable, c. Because of this, there is no residual correlation between the items.

For an LCA model with categorical outcomes, there are two types of model parameters: conditional item probabilities and class probabilities. The conditional item probabilities are specific to a given class and provide information about the probability that an individual in that class will endorse a specific item. The class probabilities specify the relative size of each class, or the proportion of the population that is in a particular class.

Consider an LCA model with r observed binary items, u, and a categorical latent variable c with K classes ($c = k$; $k = 1, 2, \ldots$, K). The marginal item probability for item $u_j = 1$ is

$$P(u_j = 1) = \sum_{k=1}^{K} P(c = k)P(u_j = 1 \mid c = k). \tag{1}$$

Assuming conditional independence, the joint probability of all the r observed items is

$$P(u_1, u_2, \ldots, u_r) = \sum_{k=1}^{K} P(c = k)P(u_1 \mid c = k)P(u_2 \mid c = k)\ldots P(u_r \mid c = k). \tag{2}$$

A product of LCA is the estimated class probabilities for each individual, called posterior probabilities, which are analogous to factor scores in a factor analysis (Muthén, 2001). These are estimates of

$$P(c = k \mid u_1, u_2, \ldots, u_r) = \frac{P(c = k)P(u_1 \mid c = k)P(u_2 \mid c = k)\ldots P(u_r \mid c = k)}{P(u_1, u_2, \ldots, u_r)}. \tag{3}$$

Note that each individual is allowed fractional class membership and may have non-zero values for many classes. It is from these probabilities that class membership is assigned. An individual is assigned to be a member of a class based on their highest probability of being in a given class, even though an individual may have several classes to which they are a partial member. Once assigned to a class, an individual is assumed to be a part of that class 100%.

One potential shortcoming to using LCA to represent psychological disorders is the assumption of conditional independence. This implies that all the symptom items are statistically

independent within a latent class. One way to satisfy the conditional independence assumption is to add additional latent classes until the assumption is met. There are two possible scenarios that can occur as a result of adding more classes. The first is that the added latent classes are reflecting only residual correlations between a small number of items, beyond what the latent classes explain. In this sense, the added latent classes are not genuine subgroups. FMM is a good alternative for this situation because the addition of a factor would account for the residual item correlations and reduce the need for a larger number of classes. Using the FMM in this scenario may lead to the factor not having all loadings be non-zero. For instance, if there is an extra correlation, beyond what the classes explain, between only two items, then only those two item's loadings would be non-zero. This highlights the flexibility of the factor approach. The second scenario is that the additional classes are genuine subgroups. FMM might be useful in reducing the number of classes in this scenario if the classes are modeling differences in severity. In this case the factor will account for this variation in severity.

Another implication of the conditional independence assumption is that for each item in a class, individuals have the same conditional item probability. This implies that all individuals in a class will have the same probability of symptom presence. This is problematic because there is often within-class heterogeneity in the form of variation in severity which violates the conditional independence assumption (Muthén & Asparouhov, 2006). This implies that individuals do not have the same conditional item probability in a class. The FMM avoids this through the use of the latent factor. Within-class, individuals have different factor values which allows for people to vary in their conditional item probabilities.

Class enumeration in mixture models. A well-known problem with mixture models is that they tend to converge on local solutions, rather than a global solution (McLachlan & Peel,

2000). This is problematic because a local solution may differ dramatically from the global

solution. The use of multiple sets of randomly generated starting values is one way to avoid

converging on a local solution (McLachlan & Peel, 2000). Observing the same maximum

likelihood for multiple sets of starting values makes it more likely that a global solution has been

found. One of the advantages of using *Mplus* V5.1 (Muthén & Muthén, 1998-2008) is that it has

a random starts feature that generates a different number of random starting value sets to

facilitate in finding the global solution.

Another known issue with mixture models is the difficulty in deciding on the appropriate

number of classes, called class enumeration. Despite the various suggestions, there is no

commonly accepted methodology on how to compare models with differing numbers of classes

(Nylund, et al. 2007). When considering a plausible set of models it is wise to utilize a

combination of statistical and substantive model checking (Muthén, 2003), which is the strategy

that is used in this paper. There are two common ways to compare models. One is statistical tests

of model fit, such as the likelihood ratio test of neighboring models. Another is using statistical

indicators, such as information criteria. The standard chi-square difference test (likelihood ratio

test, LRT) cannot be used with mixture models because the more restrictive model has its

probability parameter for one class at zero, which is on the border of the admissible parameter

space (Machlan & Peel, 2000). Two alternatives to using the chi-square difference test are the

Vuong-Lo-Mendell-Rubin test (LMR-LRT; Lo, Mendell, & Rubin, 2001) and the parametric

bootstrapped LRT (BLRT; McLachlan, 1987). Both of these tests approximate the difference in

likelihood values between a k- and $(k-1)$-class model. Each of these tests provides a p-value,

which indicates whether a $(k-1)$-class model can be rejected in favor of a k-class model. When

comparing the *p*-values of a plausible set of models, one is looking for the first time the *p*-value is non-significant, or greater than 0.05.

Another way of choosing the number of classes is to use information criteria. Information criteria, such as the Bayesian Information Criterion (BIC; Schwarz, 1978), are based on the best maximum likelihood of a model and different criteria apply different penalties for the number of parameters and sample size. Because of the different penalties that are applied by each information criterion, they will often point to differing numbers of classes. For mixture models with different numbers of classes, the model with lowest value of an information criterion is deemed the best fitting model. Nylund et al. (2007), which considers the performance of statistics and indices described here, recommend using the BIC and LMR to narrow the number of models under consideration and then include the BLRT for a smaller set of model comparisons due to the computational demands of the BLRT.

Besides looking at the statistical comparisons of the models, it is important to also understand whether the models make substantive sense and to examine the residuals. One way to explore the interpretation of the models is to examine the mean class profile for different models. It is also worthwhile to examine the class size and proportion since an over-extraction of classes can result in small and non-distinct classes (Masyn et al., in press). Another way to explore the interpretation of an FMM is through a pattern response table that displays the item response pattern, its frequency, the total number of items endorsed, class probabilities, and factor scores broken down by each class. The residuals, or the difference between the observed and estimated values, provide further insight into whether the model fits the data well. A large number of significant residuals indicates that the model does not fit the data well.

Factor Analysis

In factor analysis (FA), the goal is to investigate common content among the items (Lubke & Muthén, 2005) by seeing if items group together on continuous latent variables called factors. An example of an FA model can be seen in Figure 2. As with the latent class model, there is a circle, this time with the letter *f* in the middle, representing the factor. But, the circle in an FA is not an unordered, categorical latent variable like in LCA, but instead, a continuous one. Because the latent variable is continuous, there is no assumption of different subpopulations of individuals, like in LCA. Instead, it is assumed that all individuals in the sample are from the same homogenous population and that differences among individuals arise because of differences on the factor. Similar to LCA, there are arrows emanating from the factor that point towards the items. This indicates that the factors are measured by those combinations of items that people tend to endorse together. For example, in the context of ADHD diagnosis, one factor might be measured by those symptom items relating Inattentiveness and another might be measured by those symptom items relating to Hyperactivity.

The factor analytic model with dichotomous items *y* takes on the following form for individual *i*:

$$\boldsymbol{y^*}_i = \boldsymbol{\tau} + \boldsymbol{\Lambda}\,\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i\,,$$

$$\boldsymbol{\eta}_i = \boldsymbol{\alpha} + \boldsymbol{\zeta}_i\,,$$

and

$$y_{ij} = \begin{cases} 1 & \text{if } y^*_{ij} > 0 \\ 0 & \text{otherwise.} \end{cases}$$

(4)

where $\boldsymbol{y}_i$ are individual *i*'s responses, which is a *p* vector of observed outcomes; $\boldsymbol{y^*}_i$ is individual *i*'s latent response vector; *j* is one of the items in a *p* vector; $\boldsymbol{\tau}$ is a *p* vector of measurement

thresholds; Λ is a $p \times m$ matrix of factor loadings, where m is the number of factors; $\boldsymbol{\eta}_i$ is an m vector of factor scores; $\boldsymbol{\varepsilon}$ is p vector of residuals; $\boldsymbol{\alpha}$ is an m vector of the intercepts of the factors; $\boldsymbol{\zeta}_i$ is a p vector of residuals which is assumed to be normally distributed with mean zero and variance $\boldsymbol{\Psi}$.

An important piece of information that results from a factor analysis is the estimated factor scores, $\boldsymbol{\eta}_i$. The factor scores are estimated from the factor analysis equation given above, after the analysis has been conducted. The factor scores of all the individuals in a sample can be thought of as forming an approximation to the sample distribution of the factor. The factor score of an individual can be used as indication of where that individual is located on the factor, relative to the other individuals in the sample.

One disadvantage of FA is that it does not give a model-based classification of individuals, which can be problematic if this is the goal of the analysis. One option would be to plot the factor scores and see if there are any natural cut points or thresholds for classification. Natural cut points, however, can be hard to find in practice (Muthén, 2006), especially if there is more than one factor.

One advantage of using FA is that it is a widely used technique and so, there are many guidelines for how to do factor analysis in practice. For a technical description of FA with continuous items see Joreskog (1969) and for a more applied description and a guideline for how to do FA in practice see Brown (2006). For information on FA with categorical items see Bartholomew and Knott (1999) and Muthén (1989).

Factor Mixture Model

The factor mixture model (FMM) is a hybrid of latent class and factor analysis (Muthén, 2008) and an example of an FMM can be seen in Figure 3. The figure shows that the FMM is a

combination of LCA and FA because there is both a latent class variable, the circle with the c in the middle, and a latent continuous factor, the circle with the f in the middle. Through the use of these two types of latent variables the FMM simultaneously provides both a categorical and dimensional view of psychology. The latent class variable allows for the classification of individuals into groups while the factor models the severity of the disorder. The factor also yields quantitative scores of the disorder in the form of the factor scores.

Figure 3 shows an added flexibility of the FMM. The solid lines starting from the latent variables and ending at the symptom items indicate that the latent variables are measured by those symptoms. Additionally, there are dashed lines coming from the latent class variable and ending on the line that goes from the factor to the symptom items. This indicates that the factor structure of the model can be different in each class.

The formulation of the factor mixture model is similar to what was seen in the factor analytic model except that all the parameters have the potential to be different across the classes. A factor mixture model for $k = 1, 2, \ldots, K$ latent classes with dichotomous items can be specified as follows:

$$y^*_{ik} = \tau_k + \Lambda_k \eta_{ik} + \varepsilon_{ik},$$

$$\eta_{ik} = \alpha_k + \zeta_{ik},$$

where,

$$\zeta_{ik} \sim N(0, \Psi_k)$$

and

$$y_{ij} = \begin{cases} 1 & \text{if } y^*_{ij} > 0 \\ 0 & otherwise. \end{cases}$$

$$(5)$$

All of the parameters in the equation above have the same interpretation as the factor analytic model presented in equation 4, but now they can be class specific as indicated by the subscript k. By allowing all of the parameters to have the potential to be class-specific, the model becomes more flexible and allows for several variations that differ in the amount of measurement invariance. These variations will be discussed in more detail in the next section.

The FMM has distinct advantages over LCA and FA that may make it a better model to represent psychological disorders. Unlike LCA, the FMM does not need a conditional independence assumption given latent class. This suggests that individual differences in the severity of a disorder can be explored within a given class. These variations within a class are represented by the continuous factors and can be quantified by using the factor score for each individual (Lubke & Muthén, 2005; Muthén et al., 2006). One disadvantage to factor analysis is that there is no model-based classification of individuals. In an FMM, classification is obtained through the latent class variable.

The latent class and factor analytic models can be seen as special cases of the FMM. LCA is a special case of FMM in which the factor covariance matrix is zero. The factor analytic model can be thought of as an FMM with one latent class to which every individual in the sample belongs.

While there has been little research on how to do the FMM in practice, several authors have made contributions toward the development of this model. One of the first articles is by Yung (1997), which considers a FMM where all of the parameters were class-varying. McLachlan and Peel (2000) discuss a FMM where the factor structure is exploratory rather than confirmatory. Muthén (2006) and Muthén and Asparouhov (2006) consider an FMM using dichotomous outcomes. Muthén (2008) provides an overview of the different types factor

mixture models and breaks them down into four branches depending on the amount of measurement invariance present and whether the factor in each model is parametric or not. This paper shows how the FMM variations presented map onto the branches presented in the Muthén (2008) article. Muthén (2008) and Masyn et al. (in press) discuss longitudinal extensions of FMM. Masyn et al. also position the FMM in their dimensional-categorical spectrum which organizes latent variable models based on the types and combinations of latent variables used. Kim and Muthén (2008) explain how to do two-part FMM to be able to model data with strong floor effects.

There have been additional applications of FMMs in genetics and IRT analysis. FMM has been applied to genetic contexts by McLachlan, Do, and Ambroise (2004) to micro-array gene expression data and by Muthén, Asparouhov, and Rebollo (2006) to a twin heritability study. Mislevy and Verhelst (1990), Mislevy and Wilson (1996), Wilson (1998), and Boeck, Wilson, and Acton (2005) applied FMM to IRT studies.

FMM: Model Variations

In this paper, several different variations of the FMM and their interpretations are presented. Model presentation will start with the more restrictive models, since the interpretations tend to be simpler, and move to less restrictive models, where the interpretation can be more complex. The formulas in this section have no link function relating y* to y in order to save space. The link function is the same as the one presented in the preceding section.

The first model variation is the latent class factor analytic (LCFA) model. It is referred to as FMM-1 and can be formulated as:

$$y^*_{ik} = \tau + \Lambda \eta_{ik} + \varepsilon_{ik},$$

$$\eta_{ik} = \alpha_k , \qquad\qquad\qquad (6)$$

In this model, the only parameter that changes across classes is the factor mean, which is indicated by the subscript k on $\boldsymbol{\alpha}$. The item thresholds and factor loadings are held invariant across classes, suggesting that the disorder is being measured the same way across all classes. The factor covariance matrix, $\boldsymbol{\Psi}$, is fixed at zero, indicating that there is no severity in the disorder. Figure 4 shows a diagram of this model on the left and on the right, an example plot of the factor means versus the frequency of each class. On the right, bars are used to mark the factor means instead of a distribution because the factors have no variance, which suggests that this model has a non-parametric factor distribution. The height of the bars in the plot represents the percentage of the sample in that class. This plot suggests that the only difference between individuals arises due to having different amounts of the disorder and that there are only a set number of amounts that individuals can have, which are the number of classes. In the model diagram on the left, the latent class variable points to the factor, and not directly to the items, indicating that class membership is based on each individual's location on the factor. Because this variation has measurement invariance and non-parametric factor distributions, FMM-1 is representative of Branch 2 in Muthén (2008).

The second model variation is the FMM-2, also called a mixture factor analysis. The FMM-2 is similar to the FMM-1 except that instead of setting the factor variances and covariances to zero, they are now freely estimated in each of the classes. The formulation for the second model is:

$$y^*_{ik} = \boldsymbol{\tau} + \boldsymbol{\Lambda}\,\boldsymbol{\eta}_{ik} + \boldsymbol{\varepsilon}_{ik},$$

$$\boldsymbol{\eta}_{ik} = \boldsymbol{\alpha}_k + \boldsymbol{\zeta}_{ik},$$

where,

$$\boldsymbol{\zeta}_{ik} \sim N(\mathbf{0}, \boldsymbol{\Psi}_k). \tag{7}$$

Figure 5, similar to Figure 4, shows a model diagram on the left and a plot of the factor means versus the frequency on the right. On the right, the classes are now represented as distributions, rather than bars as in FMM-1, because the factors have variance. This indicates that there are many possible amounts of the disorder an individual can have. The model diagram on the left is almost the same as for the FMM-1 except that there is now an arrow pointing into the latent factor which indicates a residual so that the within-class factor variance is now being modeled. One variation of the FMM-2 is to allow the factor variances to be freely estimated, but held equal across classes. Because this model variation has measurement invariance and parametric factor distributions, FMM-2 is representative of Branch 1 in Muthén (2008).

In the third model variation, FMM-3, the factor loadings and factor covariance matrix are invariant, while the item thresholds are allowed to change across classes. The formulation for the FMM-3 is:

$$y^*_{ik} = \tau_k + \Lambda \eta_{ik} + \varepsilon_{ik} ,$$

$$\eta_{ik} = \zeta_i ,$$

where,

$$\zeta_{ik} \sim \mathrm{N}(0, \Psi). \tag{8}$$

Also, the factor mean is set to zero for identification purposes and so does not appear in the equations above. The factor mean will be set to zero in the remaining model variations. Figure 6, which shows the model diagram for the FMM-3, is different than the previous two model diagrams because the arrows starting from the latent class variable now point to the items rather than the factor. This shows that the item thresholds are changing across classes, which implies that the classes are now based on the responses to those items rather than the factor, as was seen in the previous two models. In FMM-3, and the remaining FMM variations, some or all the

measurement parameters are non-invariant and the factor distribution is parametric, so FMM-3

through -5 fit into Branch 3 in the Muthén (2008) paper.

The fourth model variation, FMM-4, is the same as the third model except that the factor

covariance matrix is now allowed to change across the classes. Having different factor variation

across the classes implies that there are different amounts of severity within each class. For

example, a class that can be thought of as "Unaffected" might have less variance because

individuals are showing none to almost no symptoms while a class that is "Affected" might have

more variation because individuals might have a greater range of symptoms. The formulation for

the fourth model is:

$$\boldsymbol{y^*}_{ik} = \boldsymbol{\tau_k} + \boldsymbol{\varLambda}\,\boldsymbol{\eta}_{ik} + \boldsymbol{\varepsilon}_{ik}\ ,$$

$$\boldsymbol{\eta}_{ik} = \boldsymbol{\zeta}_{ik}\ ,$$

where,

$$\boldsymbol{\zeta}_{ik} \sim \mathrm{N}(\boldsymbol{0},\ \boldsymbol{\varPsi}_k). \tag{9}$$

The fifth model, FMM-5, is the least restrictive model in terms of invariance of the

parameters. The item thresholds, factor loadings, and factor covariance matrix are all allowed to

vary across classes. The formulation for the fifth model is:

$$\boldsymbol{y^*}_{ik} = \boldsymbol{\tau_k} + \boldsymbol{\varLambda_k}\boldsymbol{\eta}_{ik} + \boldsymbol{\varepsilon}_{ik}\ ,$$

$$\boldsymbol{\eta}_{ik} = \boldsymbol{\zeta}_{ik}\ ,$$

where,

$$\boldsymbol{\zeta}_{ik} \sim \mathrm{N}(\boldsymbol{0},\ \boldsymbol{\varPsi}_k). \tag{10}$$

FMM and Measurement Invariance (MI)

The FMM variations presented above differ in terms of how much measurement invariance (MI) is present. The amount of MI can have important implications for how an FMM is interpreted, specifically whether or not the same factors are being measured in each latent class. There are many definitions of MI, but this paper will use Meredith's (1993) definition of strong factorial invariance which requires equality of factor loadings, λ, and thresholds, τ, across classes because strong factorial invariance is often considered to be sufficient for the comparison of subpopulations (Little, 1997; Widaman & Reise, 1997). For a further discussion about different definitions of MI and how they apply to FMMs see Lubke and Muthén (2005).

In strong factorial invariance, the factor loadings and thresholds of an item must be held equal across latent classes. If the factor loadings, which are slopes in the regression of the observed variable on the factors, are non-invariant in the latent classes, then a unit increase in the factor score does not result in the same increase of the dependent variable in the different classes. This differential increase can suggest one of two things. The first is that if there are many items with large differences in their factor loadings between classes, then each class has a different underlying factor. The second is if there are relatively few items with differences in their factor loadings between classes, then those items with differential loadings function differently in each class. If the item thresholds, which are negative intercepts in the regression of observed variables on the factor, are not equal across latent classes then one latent class scores consistently higher or lower than the others, independent of scores on the factor. This suggests that observed differences between classes are not entirely due to differences in the factor.

In FMM-1 and -2, the factor loadings and item thresholds are invariant across classes, suggesting that the interpretation of the factor remains the same in each class or, put another way, the factor retains the same meaning at both the low and high values in the population. In

these two variations, the latent classes are used to model the non-normality of the factor in the population. For FMM-1, Figure 4, which plots the factor distribution on the right side of the figure, shows that the factor is categorical with each latent class representing a category. For FMM-2, Figure 5, which also plots the factor distribution, shows that the factor is continuous, but non-normal. The factor in an FMM-2 is a mixture of normal distributions located at different points on the factor distribution, and each normal distribution is a latent class.

In FMM-3 and -4, the item thresholds are non-invariant, while in FMM-5 both the item thresholds and the factor loadings are non-invariant. In these model variations, because there is a violation of strong factorial invariance, the same factor does not apply to the whole population. Instead, there is a different factor, with a different interpretation, in each latent class. The implication of this is that there are several populations, represented by the latent classes, each with their own distributions.

In practice, FMM-1 and -2 often do not fit real data well because the specification of invariant factor loadings and thresholds are likely to be too restrictive for certain items. In these two models, the factor loadings influence both the item mean, which changes over classes for an item j as a function of $\tau_j + \lambda_j * \alpha_k$, and the item correlation between item j and item l within class which is $\lambda_j * \psi * \lambda_l$. Item means are not likely to change over classes as function of $\tau_j + \lambda_j * \alpha_k$, but instead change as a function of τ_{jk}. It is unlikely that α_k is the only parameter that creates changes across classes in item means for these two models. Furthermore, it is because the factor loading has a connection between both the item mean changes over classes and the item correlation within-class that FMM-1 and -2 seldom fit real data well.

Even though FMM-1 and -2 may not fit real data well, partial MI is possible. Items which violate measurement invariance may be pinpointed using one of three possible methods.

To demonstrate these methods, a simulated data set with six continuous items, y1-y6, is used. The data generating model was specified so that one item's thresholds, y5, were non-invariant across classes. The remaining thresholds for the other five items were invariant. This data was specified so that an FMM-2 was the data generating model, but with the factor variance structure equal across classes. One way to possibly identify a non-invariant item is to compare the estimated and observed item means. If an item was non-invariant then the observed and estimated means would be different from one another. In Table 1, which shows the observed and estimated means for each item from an FMM-2 in the top two rows of the table, the observed and estimated means are equivalent for each item. This is because mixture models, like FMMs, do a good job of fitting first and second order moments (means and variances), making the comparison of observed and estimated means an unsuitable method for detecting threshold non-invariance. A second way to identify these items is to examine the within-class residuals for each item's mean from an FMM-1 or -2. An item that is non-invariant may have a larger residual than items that are invariant. In Table 1, all of the items have small residuals which are close to zero, including y5. This suggests that examining the within-class residuals for each item's mean is not the best method for identifying threshold non-invariance.

A third way to detect non-invariance is to conduct a series of analyses in which one item's thresholds are held invariant and the thresholds for the rest of the items freely vary across the classes[1]. Using the model constraint feature in M*plus*, a test of whether the difference between the items whose thresholds are allowed to vary is significantly different from zero can be conducted (See Appendix for example M*plus* code). If the test is significant, those item thresholds should be considered non-invariant and if the test is non-significant, the item thresholds should be invariant. Once this is performed with each item's thresholds being held

---

[1] We thank Tihomir Asparouhov for suggesting this approach.

equal across classes, a tally can be made of the number of times each item's threshold difference test is significant. The more times an item has a significant difference test, the more likely it is that item should have freely varying item thresholds. For the simulated data, y5 had 5 significant difference tests while the remaining items had zero significant difference tests suggesting that item y5 should be non-invariant across classes and the other items should be invariant. Based on these results, a partial MI model with y5 non-invariant and the remaining items invariant should be compared to the original model, FMM-2. This strategy will be demonstrated in the attention-deficit hyperactivity disorder example.

FMM: Factor Measurement Structure

Similar to factor analysis, FMMs can have two possible factor measurement structures: confirmatory and exploratory. As in confirmatory factor analysis, confirmatory FMMs utilize substantive theory to define factors that are measured only by specific symptom items, with the factor having no influence on the remaining items. This results in what is referred to in the factor analysis literature as a simple measurement structure. Confirmatory FMM has the advantage of encouraging researchers to formalize their measurement hypothesis by having latent variables that are grounded in substantive theory, which can often lead to more parsimonious models (Asparouhov & Muthén, 2009). There are, however, disadvantages to using a simple measurement structure including fixing small cross-loadings to zero when a simple structure is used. Ignoring these cross-loadings may force a researcher to specify a more parsimonious model for the data than is necessary, which can lead to poor model fit (Asparouhov & Muthén, 2009).

An alternative to a confirmatory factor structure is an exploratory structure, where the number of factors and the relationships between the factors and the symptom items do not need to be specified a priori. Asparouhov & Muthén (2009) suggest using an exploratory structure in

structural equation models when limited measurement knowledge is available or a more complex measurement structure is needed. An exploratory structure can be applied to an FMM using the Type = Mixture EFA option in M*plus* V5.1 (Muthén & Muthén, 1998-2008). Using this option not only provides rotated factor solutions, but also standard errors for all rotated parameters and overall tests of model fit.

In this paper, all of the FMMs presented used a confirmatory structure that was guided by a priori substantive reasoning as to the number of factors and relationships between the factors and the symptom items. Given that a mixture was found to be a reasonable fit to the data in the examples, an exploratory structure was also fit to the data to examine if the presence of mixtures would change the number of factors or the factor structure since the presence of mixtures may call for a simpler factor measurement structure. In the examples presented in this paper, the presence of mixtures did not change the factor structure or the number of factors. In order to demonstrate the use of the exploratory structure in FMMs, however, it was applied to the second example.

FMM Model Building and Comparison

Because factor mixture models are relatively new to the literature there is a dearth of writing about the steps to build an FMM, but general guidelines for how do them are offered in several papers (Muthén, 2006; Muthén & Asparouhov, 2006; Muthén, Asparouhov, & Rebollo, 2006). In each of these articles, the authors analyze their data using latent class, factor analytic, and factor mixture models and then compare the fit among these three types. None of these articles provide instruction on how to construct the factor mixture models. Specifically, no information is given on how to decide on the number of classes and factors or how to pick among the different variations of the FMM.

A suggested strategy is to first fit LCA models with increasing number of classes and FA models with increasing number of factors. The next step is to fit an FMM with two classes and one factor, and in subsequent models increase the number of classes. After this, the number of factors would be increased to two, and the classes would once again be increased in subsequent models. This pattern of increasing the number of classes and then the number of factors would be repeated. The question arises: at what point does one stop increasing the number of classes and factors? In several papers we have reviewed, the strategy to decide on the number of classes and factors has been to take the best fitting number of classes from an LCA and the best fitting number of factors from an FA and combine them to form an FMM. After this is complete, no further investigation is conducted to see if there is a need for a smaller number of classes and factors. This practice is highly discouraged because having both factors and classes in the model may reduce the need for a greater number of classes and factors. For example, if an LCA solution has many classes that are parallel to one another, which is thought to model severity, then adding one factor will likely reduce the need for the same number of classes because the factor will now model the severity. Instead, it is suggested that the number of classes from the best fitting LCA model and the number of factors from the best fitting FA model be the end point of combinations of classes and factors in model building. The strategy described above should be applied to all of the FMM variations. Once this is complete, the best FMM model for the data will be selected. Each of the examples will apply the strategy outlined above for model building.

Once a best fitting FMM has been selected, it is important to compare this model to the best fitting LCA and FA models to see if a more parsimonious solution can provide a better fit and explanation of the data. This is similar to the "best" candidate approach outlined for the Dimensional-Categorical spectrum in Masyn et al. (in press). Similar to deciding on the number

of classes in a mixture analysis, the LLRT cannot be used for comparing among the different

model types. Model comparison will be based on information criteria, with the lowest value

indicating the best fitting model type, and substantive interpretation.

Examples

In the following section, two examples are explored to illustrate the model building and

comparison strategies described above and to aid in understanding the substantive interpretation

of the factor mixture model. Each example will begin with a brief description of the data,

followed by model building and comparison of latent class, factor analytic, and factor mixture

models.

Example One: Conduct Disorder

The first example examines the underlying structure of conduct disorder (CD) in

interview data from a population-based sample of Finnish twins modestly enriched for familial

alcoholism risk. The study, called FinnTwin12 (FT12; Rose et al. 2004), is longitudinal, but data

for this illustration are from the first follow-up at age 14. In total, there were 1786 children in the

study with about an equal number of boys and girls. The analysis focused on interview reports of

thirteen dichotomous symptom items measuring CD that are from the Child Semi-Structured

Assessment for the Genetics of Alcoholism, Adolescent version (C-SSAGA-A). The C-SSAGA-

A is a polydiagnostic instrument that was developed by the Collaborative Study on the Genetics

of Alcoholism (COGA; Kuperman et al., 2001). Because data from co-twins cannot be

considered independent observations, the relatedness of the twins was handled by correcting the

standard errors using the Type = Complex option in M*plus*. The Type = Complex option use a

sandwich estimator to obtain standard errors which take into account the non-independence of

the twins.

In order to obtain a diagnosis of CD, three or more symptoms must be present. In this sample, 56% of children reported no symptom presence while 12% of the sample reported having three or more symptoms present and, therefore, met criteria for a CD diagnosis. Symptom presence ranged from 0.1% (Has ever been suspended from school) to 25% of the sample (Is physically cruel to animals). For a more extensive description of the interviewed sample and analysis of CD in these 14 year-old Finnish twins see Rose et al. (2004).

According to Table 2, which displays model fit and comparison statistics for this example, the LCA model with the lowest BIC was the two-class model, followed by the three-class model. The first time the LMR p-value was non-significant, or greater than 0.05, was for the four-class model, which suggested that the three-class model was the best fitting model.

Since, the BIC and LMR p-value have indicated that the two- and three-class models were potential candidates for the "best" LCA model, the BLRT p-value and substantive interpretation of these two solutions were examined to help decide between them. The zero p-value of the BLRT for the three-class model indicated that a three-class model should be chosen over a two-class model. Figure 7 shows the item profile plot for the two-class solution. Class One, 78% of the sample, has low symptom presence as indicated by low item endorsement probabilities. Because of the large size of the class and low symptom presence, this can be considered the "unaffected" or non-symptomatic class. Class Two, 22% of the sample, is defined by the items relating to suspension, lying, and truancy, with the other items have relatively low probabilities of symptom endorsement, 0.30 or lower.

Figure 8 shows the item profile plot for the three-class solution. As in two-class solution, there is one class that has low item endorsement across all items. This class is slightly smaller than the non-symptomatic class from the two-class solution, 69%, because some of the

individuals were absorbed into the two remaining classes. Supporting this idea are the lower

endorsement probabilities of the suspension and truancy symptom items in the unaffected class.

The class that was defined by the suspension, lying and truancy items in the two-class solution

has now been split into two classes in the three-class solution. The first of these classes is similar

to the one in the two-class solution with high probabilities of endorsing the suspension and lying

items, but the probability of endorsing truancy item has drastically decreased. The probability of

endorsing other symptoms items is relatively low in this class. Potentially, some individuals in

this class may meet criteria for a CD diagnosis. The second of these classes, 5% of the sample,

still has high probabilities of suspension and lying, but also has high probabilities of running

away, property destruction, forcing someone to participate in sexual activities, and starting

fights. This class can be thought of as those individuals that are likely to meet criteria for CD.

Given that there is a substantive interpretation and reasoning for each of the three classes and the

BLRT p-value points toward the three-class model, it was selected as the "best" LCA model for

this data.

For the factor analytic models, an exploratory factor analysis was first conducted, then

based on the results, a series of confirmatory factor analytic models were explored. The results of

the exploratory factor analysis suggested two possible factor solutions. The first is a one-factor

solution in which the factor is defined by all thirteen CD items. The second is a two-factor

solution with one factor being defined by the symptom items relating to expulsion from school,

cruelty to animals, and use of a weapon when fighting and the second factor being defined by the

remaining items. Potential solutions with more than two factors had the additional factors being

solely defined by one symptom item which lead to the solution being deemed not appropriate.

When comparing the fit results of the one-and two-factor confirmatory models in Table 2, the

log-likelihood is identical, but the one-factor model has one less parameter than the two-factor

model, making it more parsimonious. This means that the BIC is lower for the one-factor model

than the two-factor model. The factor solution for the two-factor model had that several items on

the second factor did not load highly suggesting that there was no need for an additional factor.

Since the one-factor model is more parsimonious, has a lower BIC value, and there are low

loadings on the second factor in the two-factor solution, the one-factor solution was chosen as

the "best" factor analytic model for the data.

The bottom part of Table 2 presents results for the factor mixture models. Models with

one factor and one to three classes were fit to the data, based on the results of FA and LCA. In

Table 2, the FMM with the lowest BIC value is the two-class, one-factor FMM-2, with a value of

9284. But, for this model the LMR p-value is greater than 0.05, which indicates that a one-class,

one-factor model should be chosen over this model. The next lowest BIC value in the table is

9374, which is 90 points higher than the lowest BIC value. This BIC value occurs for two

models, with additional models having BIC values that are not much higher. For these FMMs,

the LMR p-value is less than 0.05, which indicated that the two-class FMM should be chosen

over the one-class model. Because there was no agreement of which model was the "best" FMM

based on the BIC and the LMR p-value, the BLRT p-value was examined in the four models

with the lowest BICs. In these four FMMs, the BLRT p-value was lower than 0.05 indicating

that a two-class FMM should be chosen over a one-class solution. Based on these results, the

two-class, one-factor FMM-2 is chosen as the "best" candidate FMM. It has the lowest BIC and

the substantive interpretation of the model made the most sense.

In Figure 9, which shows the item profile plot of the two-class, one-factor FMM-2 at the

mean of the factor, there are two classes. The first and largest class, 57% of the sample, is the

lowest line on the profile plot, with almost all items having low probabilities of symptom

endorsement. Individuals in this class are likely to never have any symptoms of CD. While this

class is similar in its interpretation to the lowest class in the three-class LCA, the percentage of

individuals in the low class differs between the two solutions suggesting that the latent classes

from a LCA and a FMM are not the same. Also, the probabilities of item endorsement are lower,

almost zero, in the FMM than the LCA and the class proportion in the FMM is similar to the

proportion of individuals in the sample that do not endorse any items, which suggests that the

low class in the FMM solution may be a true asymptomatic class. The second class, 43% of the

sample, has slightly elevated probabilities of endorsing the symptom items related to suspension,

lying, and truancy, but these probabilities are still small. This class is similar to the second class

in the three-class LCA solution, but with smaller item endorsement probabilities and a larger

class proportion. Notice that there are bands around the second class showing what the item

endorsement probabilities are if an individual is one standard deviation away from the factor

mean. The first class does not have bands around the item profile because the factor variance in

this class was non-significant. The bands around the factor mean of the second class give an idea

of the variability within each class. For example, for the symptom item relating to suspension

there is a great range of endorsement (0.20 to 0.70) once an individual is one standard deviation

away from the factor mean. But, for the symptom item relating to expulsion, there is little

variation in this item's probability of endorsement as evidence by the bands being very close to

the factor mean for that item.

Comparing the factor loadings of the FMM and FA solutions, shown in Table 3, suggests

that while the factors are similar in interpretation, they are not equivalent. Table 3, presents the

unstandardized and standardized factor loadings. The factor loadings were standardized to take

into account differences in the factor variance across models in order to to able to compare the loadings. The FMM-2 had class invariant factor loadings, hence the unstandardized factor loadings are the same in each class. But, the factor variances are non-invariant in a FMM-2, so the standardized factor loadings are also different in each class. Comparing the loadings from each class in the FMM-2 to the loadings from the FA solution, the loadings in class two are lower than the loadings in the FA solution while the class one loadings are higher. This suggests that the factor, which is considered the severity, has more influence in class one, the asymptomatic class. The higher loadings in class one also indicate that the with-in class correlation among the items is higher in class one. Note also that the items relating to suspension from school, fire starting, and truancy have the highest loadings in both solutions suggesting that these items are important in both solutions.

Now that the "best" candidate model has been selected for each of the three model types, LCA, FA, and FMA, a comparison can be made among these "best" candidate models to see which one best describes the data. One way to decide between the models is to examine the BIC for the lowest value. For this data, the two-class, one-factor FMM-2 has the lowest BIC. Another way to decide between the models is to examine what each model implies about the underlying structure of CD in this population. The three-class LCA model implies that the underlying structure of CD is categorical with three categories, which are the three-classes in the solution.

The one-factor FA solution implies that the underlying distribution of CD is continuous and normally distributed. The two-class, one-factor FMM-2 also implies that there is one underlying continuous factor of CD because of the class invariant item thresholds and factor loadings. But, unlike the one-factor FA, the distribution of the factor in FMM-2 is non-normal. The distribution of the factor in the FMM-2, shown in Figure 10, is a bi-modal distribution with

a small peak at about 0 and a taller peak at about 6.5. It is these peaks and the distributions

around them that form the mixtures in the FMM-2 solution. The first latent class is the small

peak at the bottom of the plot with the wide distribution around it. The second class is the normal

distribution further up in the plot. By having two latent classes, the FMM can model the non-

normality of the factor more precisely than a FA can. Based on the results of the BIC and

comparing what each model implies about the underlying structure of CD, the "best" model for

this data is the two-class, one-factor FMM-2 because it models the non-normality of CD better

than the other models.

Example Two: Attention-Deficit Hyperactivity Disorder (ADHD)

The second data set originated from the Los Angeles ADHD study (Smalley et al., 2000).

Unlike the previous example, which was a population sample, the LA ADHD data contain only

children who have been diagnosed as being affected with ADHD. In total, there were 994

individuals in the sample, with about seventy percent of the sample being male. Ages in the

sample ranged from 4 to 18 with a mean age of 10.47 ($SD = 3.18$). Many of the children in the

study had an affected sibling who also participated; the relatedness of the siblings was handled

by correcting the standard errors using the Type = Complex option in M*plus*, as described in the

previous example.

In this sample, ADHD was measured using the Schedule for Affective Disorders and

Schizophrenia for School-Age Children – Present and Lifetime version (K-SADS-PL; Kaufman

et al., 1997). This diagnostic instrument has 18 symptom items measuring ADHD, with nine

symptom items measuring Inattention and nine symptom items measuring

Hyperactivity\Impulsivity. Symptom presence ranged from 54.3% (Difficulty playing quietly) to

96.2% (Difficulty sustaining attention on tasks\play). There were three possible diagnoses of

ADHD: predominantly Inattentive, predominantly Hyperactive, or Combination (both Inattentive
and Hyperactive). A diagnosis of predominantly Inattentive, was given to individuals having six
or more symptoms from the Inattentive scale. A diagnosis of predominantly Hyperactive, was
given to individuals having six or more symptoms from the Hyperactive scale. An individual was
considered to have Combination ADHD if he/she met criteria for both the predominantly
Inattentive and predominantly Hyperactive. In this sample, 31% of individuals were
predominantly Inattentive, 5% were predominantly Hyperactive, and 64% were Combination.

The latent class analysis of this example was conducted by fitting one- through five-class
models to the data. The results for this example can be seen in Table 4. The lowest value of the
BIC occurred for the four-class model. Because the LMR p-values for the all of the models were
less than 0.05, that LMR was inconclusive as to which is the best LCA model.

Since, the BIC indicated that the four-class model was the best model and the LMR was
inconclusive, the BLRT p-values were examined for the three- through five-class solutions. In
Table 4, the first time the BLRT p-value was non-significant, or greater than 0.05, was for the
five-class model which suggested that a four-class model should be chosen. Since both the BIC
and BLRT have indicated that the four-class solution is the best LCA model for the data, it was
chosen as the best LCA model. The four-class model profile plot is presented in Figure 11. Class
one had high probabilities of item endorsement for the inattentive items and low probabilities of
endorsement for the hyperactivity items, suggesting that this class can be considered a
predominantly inattentive class. The second and smallest class, 15% of the sample, had a high
probability of endorsement for the Hyperactivity and Inattention items with the exception of
items that were related to working memory. The second class can be considered a hyperactive
class with good working memory. The third class, which comprised 44% of the sample and had a

high probability of endorsement for both the Inattention and Hyperactivity items, can be

considered a combination class. The fourth class also had high probabilities of item endorsement

for all the Inattention items, but had moderate to high endorsement probabilities for those items

that measure Impulsiveness.

For the factor analytic models, an exploratory factor analysis was first conducted, then

based on the results, a series of confirmatory factor analytic models were explored. The results of

the exploratory factor analysis suggested two possible solutions, which are shown in Table 5.

The first was a two-factor solution with one factor defined by all nine symptom items on the

Inattention scale and the second factor defined by all nine items of the Hyperactivity scale. The

second potential solution had four factors, with the first being defined by five items on the

Inattention scale related to dreaminess or not paying attention. The second was defined by three

symptom items on the Inattention scale relating to working memory (ex: "Often loses things",

"Forgetful in daily activities"). The third factor was defined by five symptom items on the

Hyperactivity scale and the fourth factor was defined by the remaining four symptom items on

the Hyperactivity scale that were related to Impulsivity (ex: "Blurts out answers"). The model

statics for the two- and four-factor confirmatory solutions are found in Table 4. When comparing

the two models, the four-factor model had the lowest BIC by 172 points. A LLRT also suggested

that the four-factor model fit the data better than a two-factor model, $\chi^2(5) = 206$, $p < .0001$. The

results of the BIC and the LLRT both suggested that a four-factor model fit the data better than a

two-factor model, so the four-factor model was chosen as the "best" candidate model for the FA

model.

For the FMM, the results of the LCA and FA suggested that FMMs with a maximum of

four-classes and four-factors should be fit to the data. Because a large number of factors are

computationally intensive to fit with the already complicated FMMs, the fitting of models

stopped at models with two factors. For this example, FMMs with both a confirmatory and

exploratory measurement structure of the factors were explored. The confirmatory FMMs are at

the top of Table 4 and the exploratory FMMs are at the bottom of the table.

In Table 4, when comparing the FMMs with confirmatory measurement structures to

those with exploratory ones, the models with a confirmatory structure outperformed those with

an exploratory one. This occurred for two possible reasons. The first is that in M*plus*, models

with an exploratory structure are assumed to have non-invariant factor loadings and item

thresholds, and a non-invariant factor covariance structure, like the FMM-5. In this example, the

best fitting model was a FMM-3 which has invariant factor loadings. By forcing an FMM-5 onto

data which has FMM-3 as the best fitting model, there were a large number of unnecessary

parameters which were non-significant, including a large number of non-significant cross

loadings on the factors. The second reason why the exploratory FMMs did not perform as well in

this example is that there was a clean, simple structure in these particular items.

The results of the FMMs in Table 4 suggested that a two-class, one-factor, FMM-4 fit the

data the best. This model had the lowest BIC value by 100 points and the LMR p-value and

BLRT p-value were both close to zero, which indicated that a two-class model should be chosen

over a one-class model. While this model fit the data "best" according to statistical indicators, a

closer inspection of the model revealed that it was not appropriate. Even though the factors in the

two classes do not have the same interpretation due to class-varying threshold values, the factor

loadings, shown in Table 5, for the inattention items are small and non-significant when

compared with the hyperactivity item loadings. This suggested that the model was only

measuring the hyperactivity construct and not inattention, which is a vital part of ADHD.

Further inspection of all the one-factor FMMs revealed this limitation. Because of this, the next

"best" fitting model, the two-class, two-factor FMM-3, was explored. In this model, the first

factor was defined by the inattention items only and second factor was defined by the

hyperactivity items only. All of the item loadings were significant, see Table 5. In Figure 12,

which shows the item profile plot for this solution, the first class, 33% of the sample, had a high

probability of endorsing the inattention items, and high probability of endorsing the hyperactivity

items related to impulsiveness. The second class, 67% of the sample, is similar to a Combination

class because of the high probability of endorsing both hyperactivity and inattention items.

One way to help ease the interpretation of FMMs is to use a pattern response table where

the symptom response pattern, symptom endorsement total, factor scores, and class probabilities

are displayed by latent class. In Table 6, which displays the pattern response table for the ten

most frequent response patterns in each class for the two-class, two-factor FMM-3, both classes

have a high symptom endorsement totals for the inattention items as indicated by the column

labeled Inat. total, but class one generally has lower hyperactivity symptom endorsement than

class two as indicated by the Hyperactivity symptom endorsement column. This supports the

interpretation that the second class is similar to a Combination ADHD class because individuals

in the second class have high endorsement of the both the inattentive and hyperactivity items.

In the two-class, two-factor FMM-3, the thresholds are non-invariant, but there is a

possibility for partial invariance. This example demonstrates the strategy, which was outlined in

a previous section, of how to explore if there is a need partial invariance. A series of analyses

were performed, one for each of the 18 items, in which one item's thresholds were constrained to

be equal across classes and the rest free. A tally was then made of the number of times each

item's difference test was significant. The items with the highest number of times their threshold

difference test was significant are: Motor (9), Seated (8), Fidgets (8), Runs (8), Attention (7), and Instruct (7). Based on the results of the tally, a new model was run where the items mentioned above had freely estimated thresholds in each class and the remaining items had thresholds that were held invariant. Table 4 shows the model fits results for this model in the last line of the table. The BIC for this partially invariant model is only one point lower than the non-invariant two-class, two-factor FMM-3, but the partially invariant model has 12 fewer parameters making it a more parsimonious model.

Comparing the BIC values of the "best" fitting LCA, FA, and FMM located in Table 4, the four-factor FA has the lowest BIC value by 100 points. This example highlights that though the FMM is a powerful and flexible model, it does not always provide the best fit for the data. One possible reason the FMM did not fit as well as the four-factor FA model is that this is a sample of only affected children so there was not much variation in symptom endorsement because the individuals had to endorse a high number of symptoms to be included in this sample. This is contrary to what was seen in the previous example where the FMM was the "best" model for the data. Lubke et al. (2007) applied LCA, FA, and FMM to a population sample in order to investigate the structure of ADHD and found that a FMM fit the data the best.

Discussion

This paper builds upon the work of previous research on factor mixture models in order to explicate how to conduct a factor mixture analysis in practice. This paper discusses several different variations of the FMM and what each suggests about the structure underlying psychological disorders. Steps for building a factor mixture model, including how to decide on the number of classes and factors, and how to compare the factor mixture to other appropriate models, are also discussed.

The two examples examining the underlying structure of conduct disorder and attention-deficit hyperactivity disorder shed light on some issues and challenges that arise when conducting a factor mixture analysis. The first example, which investigated the underlying structure of conduct disorder in a sample of Finnish twins, showed how difficult it can be to decide between a factor mixture model and other appropriate models, such as the factor analytic and latent class models. There is often no clear "winner" based on statistical tests and indicators. The decision between different types of models must be made based on what each model implies about the underlying structure and whether that can be justified substantively. This example also highlights how FMM can be used to model a non-normal, continuous factor.

In the second example, which explored the underlying structure of ADHD in a sample of affected children and adolescents, several issues arose. First, the "best" fitting FMM according to statistical tests and indicators may not always make substantive sense, which suggests models should always be checked for their interpretability. Second, despite the flexibility of the FMM, it is not always the best fitting or most appropriate model for the data. Researchers need to consider whether the FMM is appropriate for their data and what the FMM they are using implies about the underlying structure. Third, despite the advances in computational technology in recent years, more complicated factor mixture models with large numbers of classes and factors can still be difficult to process. In this example, the results of the factor analysis suggested that factor mixture models should be fit with up to four factors. But once the FMMS were increased to four factors, the computational demands of the model were too heavy to run the models. This example also demonstrated how to investigate if there is a need for partial measurement invariance in a factor mixture model.

The two examples have shown that despite the work that has been done with the FMM, there are still avenues for further study. One area for further study is how and when to incorporate covariates or other auxiliary variables into a FMM. Lubke and Muthén (2005) describe different ways that a covariate can influence the latent classes and factors, but not at what point in the analysis covariates should be included or what method should be used to incorporate them. Another area for the further study is the performance of statistical tests and indicators in correctly identifying the correct combination of latent classes and factors. Nylund et al. (2007) have begun this work by investigating a single case with dichotomous items, but this paper only focused on the correct identification of the number of classes with no focus on identifying the number of factors. Lubke and Neale (2006, in press) have conducted simulation studies exploring the use of statistical indicators to identify the presence of measurement invariance in FMMs.

Even though there are some areas that need to be explored further, the factor mixture model is still an important analytic tool for conceptualizing the structure underlying psychological disorders. Using the factor mixture model, does not force the conceptualization to be either categorical or continuous, as previous notions have suggested. Instead the factor mixture model allows for the underlying structure to be modeled as both so that one can simultaneously classify individuals into groups and have an estimate of the severity within those groups.

References

Bartholomew, D.J., and Knott, M. (1999). *Latent variable models and factor analysis*. Kendall's

    Library of Statistics 7. Oxford University Press: New York.

de Boeck, P., Wilson, M., & Acton, G.S. (2005). A conceptual and psychometric

    framework for distinguishing categories and dimensions. *Psychological Review*,

    112, 129-158.

Brown, T.A. (2006). *Confirmatory factor analysis for applied research*. New York: The Guilford

    Press.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor

    analysis. *Pyschometrika*, 34.

Kaufman J., Birmaher B., Brent, D., Rao, U., Flynn, C., Moreci, P., et al. (1997). Schedule for

    affective disorders and schizophrenia for school-age children – present and liftetime

    version (KSADS-PL): Initial reliability and validity data. *Journal of the American

    Academy of Child and Adolescent Psychiatry*, 36, 980-8.

Kim, Y.K. & Muthén, B. (2008). Two-part factor mixture modeling: Application to an

    aggressive behavior measurement instrument. Forthcoming in *Structural Equation

    Modeling*.

Kuperman, S., Schlosser, S. S., Kramer, J. R., Bucholz, K. K., Hesselbrock, V., Reich, T., et al.

    (2001). Developmental sequence from disruptive behavior diagnosis to adolescent

    alcohol dependence. *American Journal of Psychiatry*, 158, 2022-2026.

Lazarsfeld, P., & Henry, N. (1968). *Latent Structure Analysis*. New York: Houghton

    Mifflin.

Little, T.D. (1997) Mean and covariance structures (MACS) analyses of cross-cultural data:

Practical and theoretical issues. *Multivariate Behavioral Research*, 32, 53-76.

Lo, Y., Mendell, N.R., Rubin, D.B. (2001). Testing the number of components in a normal

mixture. *Biometrika*, 88, 767-778.

Lubke, G.H. & Muthén, B. (2005). Investigating population heterogeneity with factor

mixture models. *Psychological Methods*, 10, 21-39.

Lubke, G., Muthén, B., Moilanen, I., McGough, J., Loo, S., Swanson, J., Yang, M., Taanila, A.,

Hurtig, T., Jarvelin, M. & Smalley, S. (2007). Subtypes versus severity differences in the

Attention-Deficit/Hyperactivity disorder in the northern Finnish birth cohort. *Journal of

the American Academy of Child and Adolescent Psychiatry*, 46, 1584-1593.

Lubke, G.H., & Neale, M.C, (2006). Distinguishing between latent classes and continuous

factors: Resolution by maximum likelihood? *Multivariate Behavioral Research*, 41, 499-

532.

Lubke, G. H., & Neale, M.C. (In press). Distinguishing between latent classes and continuous

factors with categorical outcomes: Class invariance of parameters of factor mixture

models. *Multivariate Behavioral Research*.

Masyn, K., Henderson, C., & Greenbaum, P. (In press). Exploring the latent structures of

psychological constructs in social development using the Dimensional-Categorical

Spectrum. *Social Development*.

McLachlan, G.J. (1987). On bootstrapping the likelihood ratio test statistic for the number of

components in a normal mixture. *Applied Statistics*, 36, 318-324.

McLachlan, G.J., Do, K-A., Ambroise, C. (2004). *Analyzing microarray gene expression

data*. New York: Wiley.

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

Meredith, W. (1993) Measurement invariance, factor analysis, and factorial invariance.

    Psychometrika, 58, 525-543.

Mislevy, R.J., & Verhulst, N. (1990). Modeling item response when different subjects

    employ different solution strategies. *Psychometrika*, 55, 195-215.

Mislevy, R.J., & Wilson, M. (1996). Marginal maximum likelihood estimation for a

    psychometric model of discontinuous development. *Psychometrika*, 61, 41-71.

Muthén, B. (1989). Dichotomous factor analysis of symptom data. In Eaton, & Bohrnstedt

    (Eds.), Latent Variable Models for Dichotomous Outcomes: Analysis of Data from the

    Epidemiological Catchment Area Program (pp. 19-65), a special issue of *Sociological*

    Methods & Research, 18, 19-65.

Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling.

    Psychological Methods, 8, 369-377.

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

    Publications.

Muthén, B. (2006). Should substance use disorders be considered as categorical or

    dimensional? *Addiction*, 101 (Suppl. 1), 6-16.

Muthén, B. (2008). Latent variable hybrids: Overview of old and new models. In

    Hancock, G. R., & Samuelsen, K. M. (Eds.), *Advances in latent variable mixture*

    models, pp. 1-24. Charlotte, NC: Information Age Publishing, Inc.

Muthén, B. & Asparouhov, T. (2006). Item response mixture modeling: Application to

    tobacco dependence criteria. *Addictive Behaviors,* 31, 1050-1066.

Muthén, B., Asparouhov, T. & Rebollo, I. (2006). Advances in behavioral genetics

    modeling using Mplus: Applications of factor mixture modeling to twin data.

    Twin Research and Human Genetics, 9, 313-324.

Muthén, L.K., & Muthén, B.O. (2008). M*plus* (Version 5.1). Los Angeles, CA.

Nylund, K.L., 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(4), 535-569.

Rose, R.J., Dick, D.M., Viken, R.J., Pulkkinen, L., Nurnberger, J.I., Jr., & Kaprio, J. (2004).

    Genetic and environmental effects on conduct disorder, alcohol dependence symptoms,

    and their covariation at age 14. *Alcoholism: Clinical and Experimental Research*, 28,

    1541-1548.

Schwarz, G. (1978). Estimating the dimensions of a model. *Annals of Statistics*, 6, 461-464.

Smalley, S.L., McGough, J.J., Del'Homme, M., NewDelman, J., Gordon, E., Kim, T., Liu, A., &

    McCracken, J.T. (2000). Familial clustering of symptoms and disruptive behaviors in

    multiplex families with attention-deficit/hyperactivity disorder. *Journal of the American*

    Academy of Child & Adolescent Psychiatry, 39: 9.

Widamon, K.F., & Reise, S.P. (1997). Exploring the measurement invariance of psychological

    instruments: Applications in the substance use domain. In K.J. Bryant, M. Windle, &

    S.G. West (Eds.), *The science of preventions: Methodological advances from alcohol and*

    substance abuse research (pp. 281-324). Washington, DC: American Psychological

    Association.

Wilson, M. (1989). Saltus: A psychometric model of discontinuity in cognitive

    development. *Psychological Bulletin*, 105, 276-289.

Yung, Y.F. (1997). Finite mixtures in confirmatory factor-analysis models.

Psychometrika, 62, 297-330.

Appendix
Example M*plus* Input and Model Statements for Factor Mixture Model Variations

Input 1a: Factor Mixture Model 1 (FMM-1)
- Estimating both factor means in each class.
- Two-classes, two-factors.
- Class invariant thresholds (τ), class invariant factor loadings (λ), factor covariance matrix is equal to zero ($\psi = 0$), class varying factor means (α_k).

```
Data: File is example.dat;
      ! This is how to comment out text.
Variable:
      Names are u1-u6; !List of variables in dataset.
      Usevariables are u1-u6;
            ! Specifies which variables are to be used in the analysis.
      Categorical = u1-u6;
            ! Specifies which dependent variable is categorical.
      Classes = c(2);
            ! Defines a latent class variable called c with two classes.

Analysis:
      Type = Mixture; !Specifies doing a mixture analysis.
      Starts = 100 10;
            ! Specifies total number of random starts and number of
            ! final stage iterations.
      Process = 2 (STARTS);
            ! Specifies that 2 processors are to be used in analysis
            ! and that random starts are to be spread out on the
            ! multiple processors.

Model:
      %Overall% ! Specifies that model below will be applied to both
                ! latent classes.
            f1 by u1-u3; ! Defining a factor called f1 which is measured by
                         ! items u1 to u3.
            f2 by u4-u6;
            f1-f2@0; ! Fixing factor variance at zero.
            f1 with f2 @0; ! Fixing covariance between f1 and f2 at zero;

      %c#1% ! Indicates statements below apply to class 1 only.
            [u1$1-u6$1] (1-6);

      %c#2%
            [u1$1-u6$1] (1-6);   ! Number in parentheses indicates fixing the
                                 ! threshold of this item to be the same in
                                 ! class 1 and class 2.
```

- By default, the last class will have the factor mean fixed to zero, and because there is a factor loading fixed at one, a factor mean for a different class can be estimated as well.

Input 1b: Factor Mixture Model 1 (FMM-1)
- Three-classes, one-factor.

- Estimating all factor loadings.
- Class invariant thresholds (τ), class invariant factor loadings (λ), factor covariance matrix is equal to zero ($\psi = 0$), class varying factor means (α_k).

```
Model:
     %Overall%
          f1 by u1-u3*;   !* indicates that all factor loadings are to
                           ! be freely estimated.
          f1@0;

     %c#1%
          [f1*]; !* indicates the factor mean is freely estimated.
          [u1$1-u6$1] (1-6);
     %c#2%
          [f1@1]; !Fixing factor mean at one.
          [u1$1-u6$1] (1-6);
     %c#3%
          [u1$1-u6$1] (1-6);
```

- By default the last class will have the factor mean fixed to zero. Because all of the factor loadings are freely estimated, at least one of the factor means must be fixed for model identification. Here, class 2 has the factor mean set to one.

Input 2: Factor Mixture Model 2 (FMM-2)
- Two-classes, two-factors.
- Class invariant thresholds (τ), class invariant factor loadings (λ), class varying factor covariance matrix (ψ_k), class varying factor means (α_k).

```
Model:
     %Overall%
          f1 by u1-u3;
          f2 by u4-u6;
          f1-f2@0;
          f1 with f2 @0;
     %c#1%
          f1-f2;
          [f1-f2@0];
          [u1$1-u6$1] (1-6);
     %c#2%
          f1-f2;
          [f1-f2*];
          [u1$1-u6$1] (1-6);
```

Input 3: Factor Mixture Model 3 (FMM-3)
- Two-classes, two-factors.
- Class varying thresholds (τ_k), class invariant factor loadings (λ), class invariant factor covariance matrix (ψ), factor means fixed at zero ($\alpha_k = 0$).

```
Model:
     %Overall%
          f1 by u1-u3;
          f2 by u4-u6;
```

```
        f1 - f2;
        f1 with f2;
        [f1-f2@0];
    %c#1%
        [u1$1-u6$1];
    %c#2%
        [u1$1-u6$1];
```

Input 4: Factor Mixture Model 4 (FMM-4)
  - Two-classes, two-factors.
  - Class varying thresholds (τ_k), class invariant factor loadings (λ), class varying factor covariance matrix (ψ_k), factor means fixed at zero ($\alpha_k = 0$).

```
Model:
    %Overall%
        f1 by u1-u3;
        f2 by u4-u6;
        [f1-f2@0];
    %c#1%
        f1-f2;
        f1 with f2;
        [u1$1-u6$1];
    %c#2%
        f1-f2*;
        f1 with f2;
        [u1$1-u6$1];
```

Input 5: Factor Mixture Model 5 (FMM-5)
  - Two-classes, two-factors.
  - Class varying thresholds (τ_k), class varying factor loadings (λ_k), class varying factor covariance matrix (ψ_k), factor means fixed at zero ($\alpha_k = 0$).

```
Model:
    %Overall%
        f1 by u1-u3;
        f2 by u4-u6;
        [f1-f2@0];
    %c#1%
        f1 by u2-u3;
        f2 by u3-u6;
        f1-f2;
        f1 with f2;
        [u1$1-u6$1];
    %c#2%
        f1 by u2-u3;
        f2 by u5-u6;
        f1-f2;
        f1 with f2;
        [u1$1-u6$1];
```

-The overall statement (%Overall%) defines the factors and in the class specific statements (%c#1%, %c#2%) there is what looks like another definition of the factor, but this is not defining the factors. This allows the loadings to be freely estimated in both classes. Note that the in the

class specific statement, the factor definition starts with the second item and not the first. This is because in the overall statement where the factor is first defined Mplus defaults to fixing the loading of the first item to one. Since the loading of the first item is already fixed, it does not need to be set to be freely estimated in the class specific statement so the class specific statements begins with the second item.

Input 6: MI Investigation Example
- Two-classes, one-factor.
- Class invariant threshold anchor item: u1.

```
Model:
    %Overall%
            f1 by u1 - u6;
            f1;
            [f1@0];
    %c#1%
            [u1$1]          (1);
            [u2$1]          (p1_2);
            [u3$1]          (p1_3);
            [u4$1]          (p1_4);
            [u5$1]          (p1_5);
            [u6$1]          (p1_6);
    %c#2%
            [u1$1]          (1);
            [u2$1]          (p2_2);
            [u3$1]          (p2_3);
            [u4$1]          (p2_4);
            [u5$1]          (p2_5);
            [u6$1]          (p2_6);

Model Constraint:

        New(difi1 difi2 difi3 difi4 difi5 difi6);
        difi1= p2_1 - p1_1;
        difi2= p2_2 - p1_2;
        difi3= p2_3 - p1_3;
        difi4= p2_4 - p1_4;
        difi5= p2_5 - p1_5;
        difi6= p2_6 - p1_6;
```

Note: All of the example inputs and model statements above used dichotomous items, but factor mixture modeling can be conducted with other types of items as well with slight adjustments to the code above. Taking Input 4 as an example, the code can be adjusted for polytomous item by mentioning all the thresholds in the items. In the example below, the code has been adjusted for use with trichotomous items.

```
Model:
    %Overall%
            f1 by u1-u3;
            f2 by u4-u6;
            [f1-f2@0];
```

```
%c#1%
      f1-f2;
      f1 with f2;
      [u1$1-u6$1];
      [u1$2-u6$2]; ! Mentioning second threshold of trichotomous item.
%c#2%
      f1-f2*;
      f1 with f2;
      [u1$1-u6$1];
      [u1$2-u6$2];
```

Also using Input 4 as an example, the code can be adjusted for continuous items by mentioning the mean of the items, rather than threshold.

```
Model:
    %Overall%
          f1 by u1-u3;
          f2 by u4-u6;
          [f1-f2@0];
    %c#1%
          f1-f2;
          f1 with f2;
          [u1-u6]; ! Setting mean of items to be different across classes.
    %c#2%
          f1-f2*;
          f1 with f2;
          [u1-u6];
```

Table 1

Simulated MI Data: Observed vs. Estimated Mean and Within-Class Residuals

Item	y1	y2	y3	y4	y5	y6
Observed Mean	0.948	0.784	0.755	0.729	1.721	0.74
Estimated Mean	0.948	0.784	0.755	0.729	1.721	0.74
Class 1 Residual	0.00	-0.01	-0.02	-0.01	0.04	-0.01
Class 2 Residual	0.00	0.01	0.02	0.01	-0.04	0.02

Table 2

Conduct Disorder Example: Model Comparison Results (n =1758)

Model	Log-likelihood	Par.	BIC	LMR p-value	BLRT p-value
Latent Class Analysis					
One-Class	-5018	13	10062		
Two-Class	-4586	27	9374	0	0
Three-Class	-4540	41	9387	0.033	0
Four-Class	-4524	55	9460	0.095	
Five-Class	-4513	69	9543	0.132	
Factor Analysis					
One-Factor	-4561	26	9317		
Two-Factors	-4561	27	9324		
Factor Mixture Analysis					
2-Class, 1-Factor					
FMM-1	-4586	27	9374	0	0
FMM-2	-4537	28	9284	0.235	0.04
FMM-3	-4357	40	9374	0.006	0
FMM-4	-4536	41	9379	0.029	0
FMM-5	-4526	53	9449	1	
3-Class, 1-Factor					
FMM-1	-4540	41	9387	0.034	
FMM-2	*				
FMM-3	-4524	54	9452	0.527	
FMM-4	-4523	55	9481	0	
FMM-5	-4496*	81	9600	1	

Note. * Log-likelihood not replicated.

Table 3

Conduct Disorder Example: Unstandardized and Standardized Factor Loadings

Loadings	Unstandardized			Standardized		
Model	FA	FMM		FA	FMM	
Number of Factors	1 Factor	2-Class, 1-Factor FMM-2		1 Factor	2-Class, 1-Factor FMM-2	
		Class 1	Class 2		Class 1	Class 2
Items						
Suspend	0.90	0.90	0.90	1.81	4.41	1.07
Expelled	0.68	0.47	0.47	1.36	2.31	0.56
Stolen	0.49	0.36	0.36	0.99	1.75	0.42
Runaway	0.61	0.47	0.47	1.22	2.31	0.56
Lie	0.72	0.64	0.64	1.45	3.13	0.76
Fire	0.96	0.81	0.81	1.92	3.95	0.96
Truant	1.00	1.00	1.00	2.01	4.88	1.18
Property	0.59	0.44	0.44	1.18	2.17	0.53
Animal	0.67	0.42	0.42	1.34	2.06	0.50
Sexual	0.53	0.4	0.4	1.07	1.97	0.48
Fights	0.61	0.5	0.5	1.23	2.44	0.59
Weapon	0.36	0.26	0.26	0.72	1.25	0.30
Stolen	0.72	0.57	0.57	1.45	2.79	0.68
Factor Variance	4.03	23.8*	1.40			

Note. * $p > 0.05$.

Table 4

ADHD Example: Model Comparison Results (n = 994)

Model	Log-likelihood	Par.	BIC	LMR p-value	BLRT p-value
Latent Class Analysis					
One-Class	-8378	18	16879		
Two-Class	-7534	37	15323	0	
Three-Class	-7320	56	15025	0	0
Four-Class	-7214	75	14944	0.016	0
Five-Class	-7155	94	14956	0.026	1
Factor Analysis					
One-Factor	-7448	36	15143		
Two-Factors	-7232	37	14719		
Four-Factors	-7129	42	14547		
Factor Mixture Analysis - Confirmatory					
2-Class, 1-Factor					
FMM-1	-7534	37	15323	0	
FMM-2	-7423	39	15115	0	
FMM-3	-7217	55	14813	0	
FMM-4	-7217	56	14547	0	0
FMM-5	-7161	73	14823	0	
3-Class, 1-Factor					
FMM-1	-7451	39	15170	0	
FMM-2	-7417	42	15123	0.002	
FMM-3	-7135	74	14780	0.083	
FMM-4	-7135	76	14792	0.164	
FMM-5	-7072	110	14901	0.461	
2-Class, 2-Factor					
FMM-1	-7534	38	15330	0	
FMM-2	-7191	45	14691	0.339	
FMM-3	-7131	56	14647	0.015	
FMM-4	-7124	59	14655	0.535	
FMM-5	-7090	75	14695	0.004	
3-Class, 2-Factor					
FMM-1	-7338	40	14952	0	
FMM-2	-7211	49	14659	0.004	
FMM-3	-7080	75	14675	1	
FMM-4	-7067	81	14691	0.015	
FMM-5	-7038	113	14852	0.011	
2-Class, 2-Factor Partial MI	-7171	44	14646		
Factor Mixture Analysis - Exploratory					
2-Class,1-Factor	-7165	73	14833		
2-Class, 2-Factor	-7044	107	14824		
2-Class, 3-Factor	-7001	139	14958		
2-Class, 4-Factor	**				
3-Class, 1-Factor	-7065	110	14886		

3-Class, 2-Factor	-6979	161	15064
3-Class, 3-Factor	**		
3-Class, 4-Factor	**		

Note. ** Heywood case.

Table 5

ADHD Example: Factor Loadings

Model	FA						FMM-4: 2-Class	FMM-3: 2-Class	
Number of Factors	2		4				1	2	
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 1	Factor 2
Items									
Attention	0.98*		1.00				-0.02	0.97*	
Distract	0.65*		0.68*				0.021	0.62*	
Mistake	0.53*		0.36*				-0.13	0.40*	
Listen	0.65*		0.52*				0.08	0.52*	
Instruct	0.86*		0.71*				-0.06	0.75*	
Organization	0.73*			0.34*			-0.12	0.61*	
Avoid	0.41*		0.27*				-0.01	0.32*	
Loses	1.00		1.00				-0.19	1.00	
Forget	0.95*			0.46*			-0.22	0.76*	
Seated		0.63*			0.58*		0.61*		0.61*
Fidget		0.66*			0.62*		0.63*		0.65*
Runs		0.91*			0.84*		0.91*		0.88*
Quiet		0.46*			0.38*		0.45*		0.48*
Blurt		0.53*				0.80*	0.50*		0.73*
Turn		0.59*				0.78*	0.57*		0.67*
Interrupt		0.59*				1.00	0.55*		0.95*
Talk		0.38*				0.51*	0.37*		0.50*
Motor		1.00			1.00		1.00		1.00

Note. * p-value < 0.05.

Table 6

ADHD Example: Pattern Response Table for Two-class, Two-factor FMM-3.

Freq.	Inattention Items										Hyperactivity Items										Item Total	Factor Score		Class Probs.	
	Attn.	Distract	Mistake	Listen	Instruct	Organ.	Avoid	Loses	Forget	Inat. Total	Seated	Fidget	Runs	Quiet	Blurt	Turn	Inter.	Talk	Motor	Hyper. Total		Factor 1	Factor 2	Class 1	Class 2
Class One																									
14	1	1	1	1	1	1	1	1	1	9	0	0	0	0	0	0	0	0	0	0	9	1.22	-5.13	0.58	0.42
9	1	1	1	1	1	1	1	1	1	9	1	1	0	0	1	1	1	1	0	6	15	1.50	0.38	0.85	0.15
6	1	1	1	1	1	1	1	1	1	9	1	1	1	1	1	1	1	1	0	8	17	1.59	2.08	0.62	0.38
5	1	1	1	1	1	1	1	1	1	9	1	1	1	0	1	1	1	1	0	7	16	1.55	1.43	0.61	0.39
5	1	1	1	1	1	1	1	1	1	9	0	1	0	0	0	0	1	1	0	3	12	1.38	-1.91	0.77	0.23
4	1	1	1	1	1	1	1	1	1	9	1	1	0	1	1	1	1	1	0	7	16	1.53	0.94	0.85	0.15
4	1	1	1	1	1	1	1	1	1	9	0	0	0	0	0	0	1	0	0	1	10	1.30	-3.39	0.90	0.10
4	1	1	1	1	1	1	1	1	1	9	0	0	0	0	0	0	0	1	0	1	10	1.27	-4.10	0.77	0.23
4	1	1	1	1	1	1	1	1	1	9	0	0	0	0	1	0	1	0	0	2	11	1.35	-2.41	0.97	0.03
4	1	1	1	1	1	1	1	1	1	9	1	1	0	1	1	0	1	1	0	6	15	1.48	0.16	0.76	0.25
Class Two																									
122	1	1	1	1	1	1	1	1	1	9	1	1	1	1	1	1	1	1	1	9	18	1.65	2.84	0.18	0.82
22	1	1	1	1	1	1	1	1	1	9	1	1	1	0	1	1	1	1	1	8	17	1.57	1.40	0.20	0.80
18	1	1	1	1	1	1	0	1	1	8	1	1	1	1	1	1	1	1	1	9	18	0.65	2.78	0.23	0.78
16	1	1	1	1	1	1	1	1	1	9	1	1	1	1	1	0	1	1	1	8	17	1.55	0.96	0.12	0.88
15	1	1	1	1	1	1	1	1	1	9	1	1	1	0	1	1	1	0	1	7	16	1.52	0.39	0.07	0.93
11	1	1	1	1	1	1	1	1	1	9	1	1	1	1	1	1	1	0	1	8	17	1.57	1.35	0.07	0.93
8	1	1	1	1	1	1	1	1	1	9	1	1	1	1	0	1	1	1	1	8	17	1.54	0.84	0.05	0.95
8	1	1	1	1	1	1	1	0	1	8	1	1	1	1	1	1	1	1	1	9	18	-.92	2.7	0.03	0.97
7	1	1	1	1	1	1	1	1	1	9	1	1	1	0	0	1	1	1	1	7	16	1.5	0.02	0.05	0.95
7	1	1	1	1	1	0	1	1	1	8	1	1	1	1	1	1	1	1	1	9	17	-.11	2.74	0.05	0.95

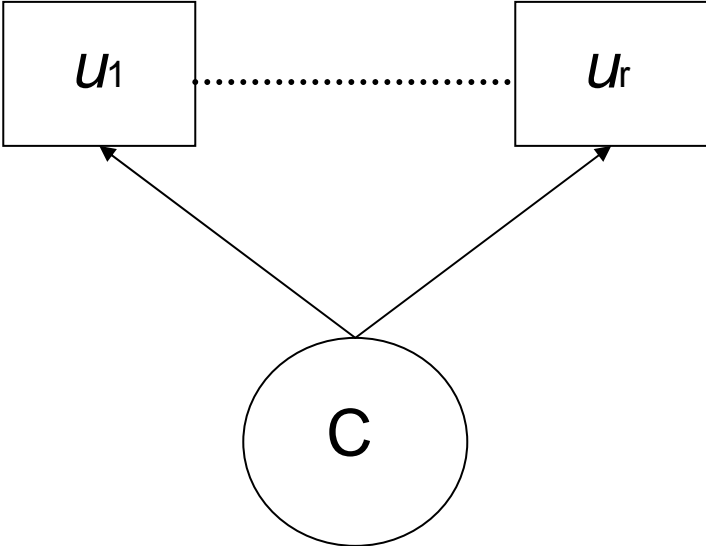*Figure 1*. Latent Class Analysis Model Diagram.

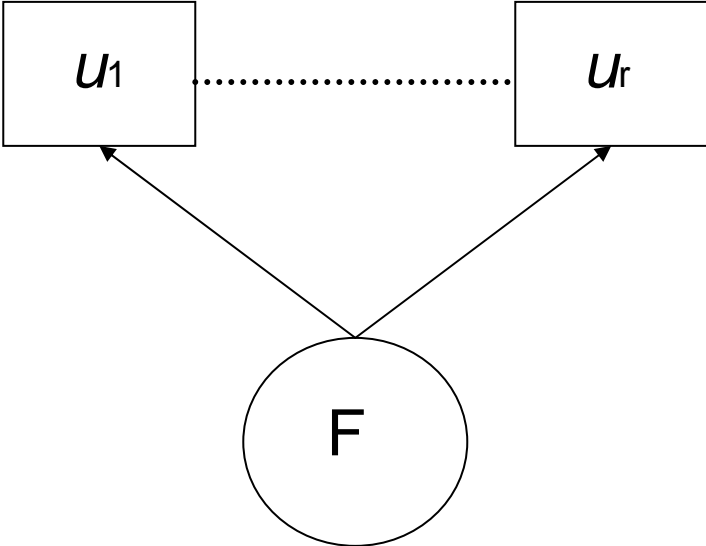*Figure 2.* Factor Analysis Model Diagram.

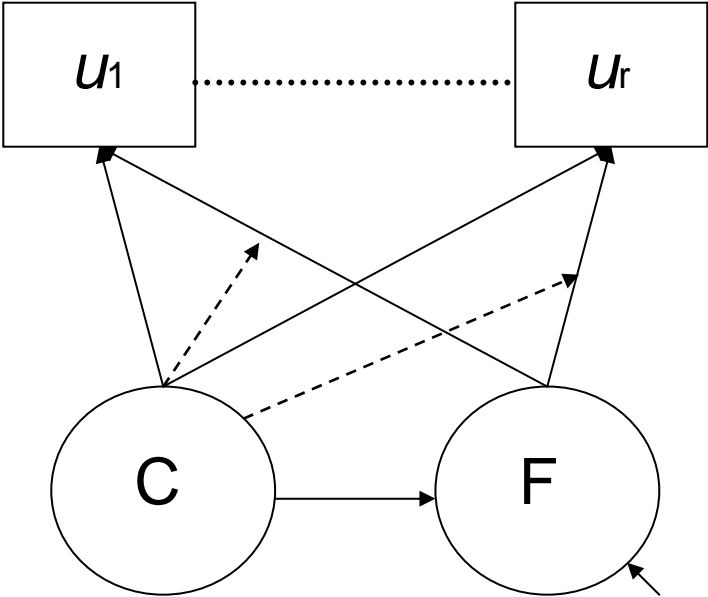*Figure 3.* General Factor Mixture Model Diagram.

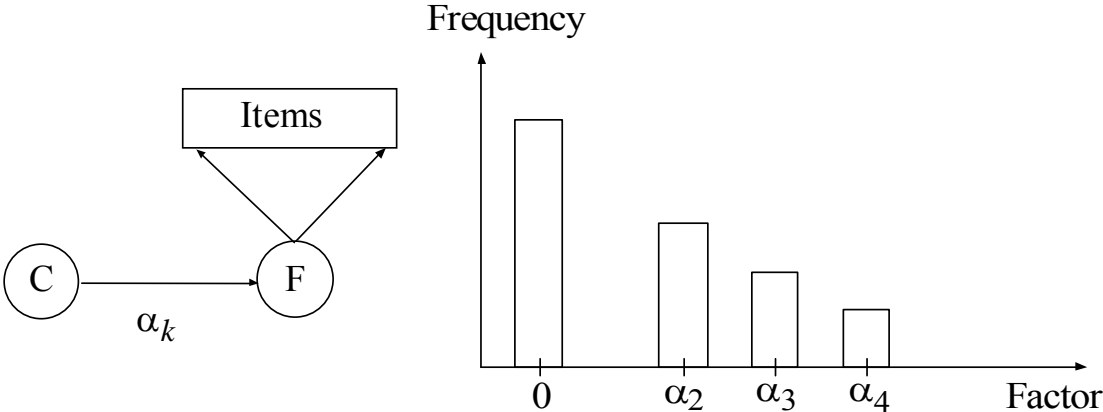*Figure 4.* Factor Mixture Model-1: Diagram and Factor Distribution Plot.

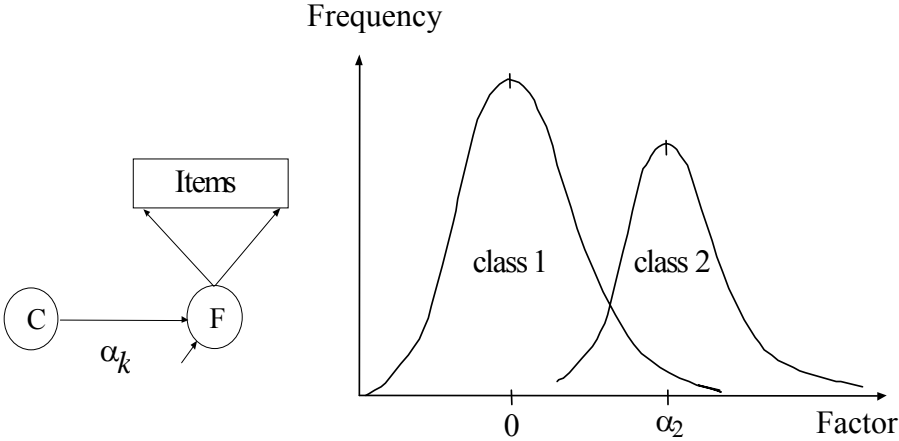*Figure 5.* Factor Mixture Model-2: Diagram and Factor Distribution Plot.
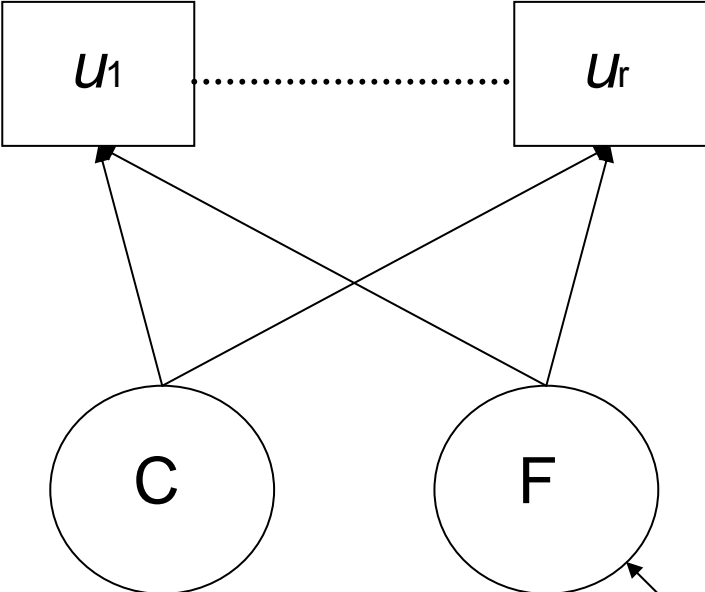
Figure 6. Factor Mixture Model-3: Diagram.

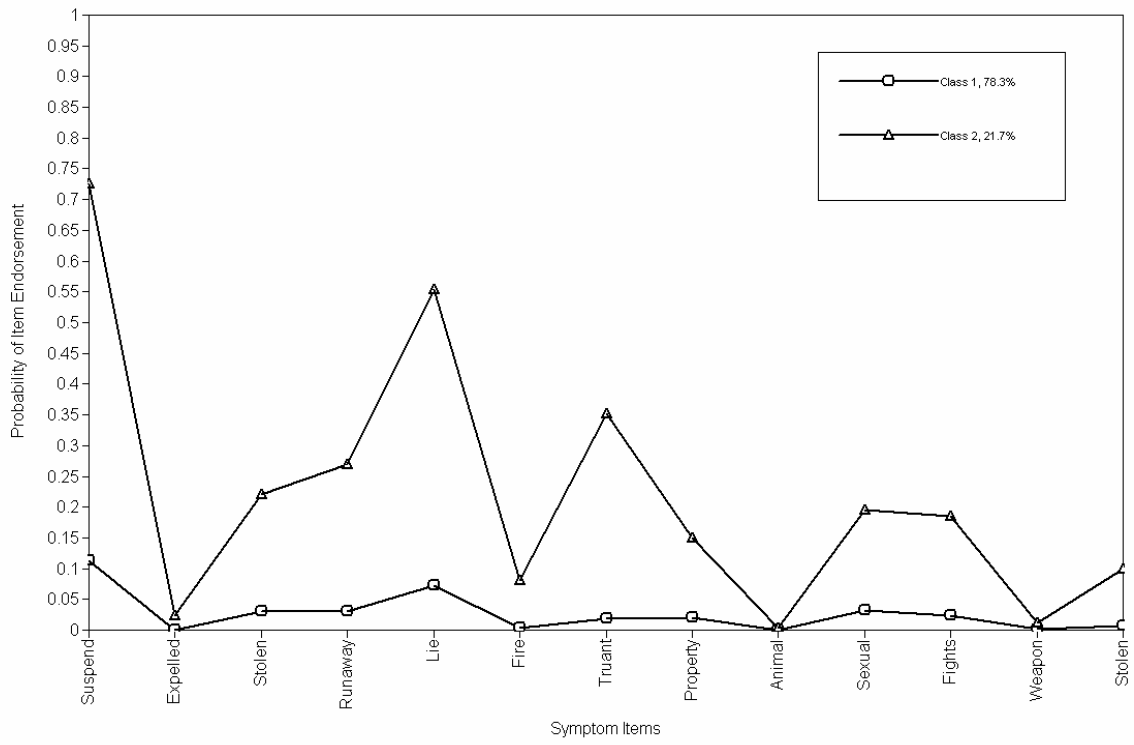*Figure 7.* Conduct Disorder Example: Two-class Latent Class Analysis Profile Plot.

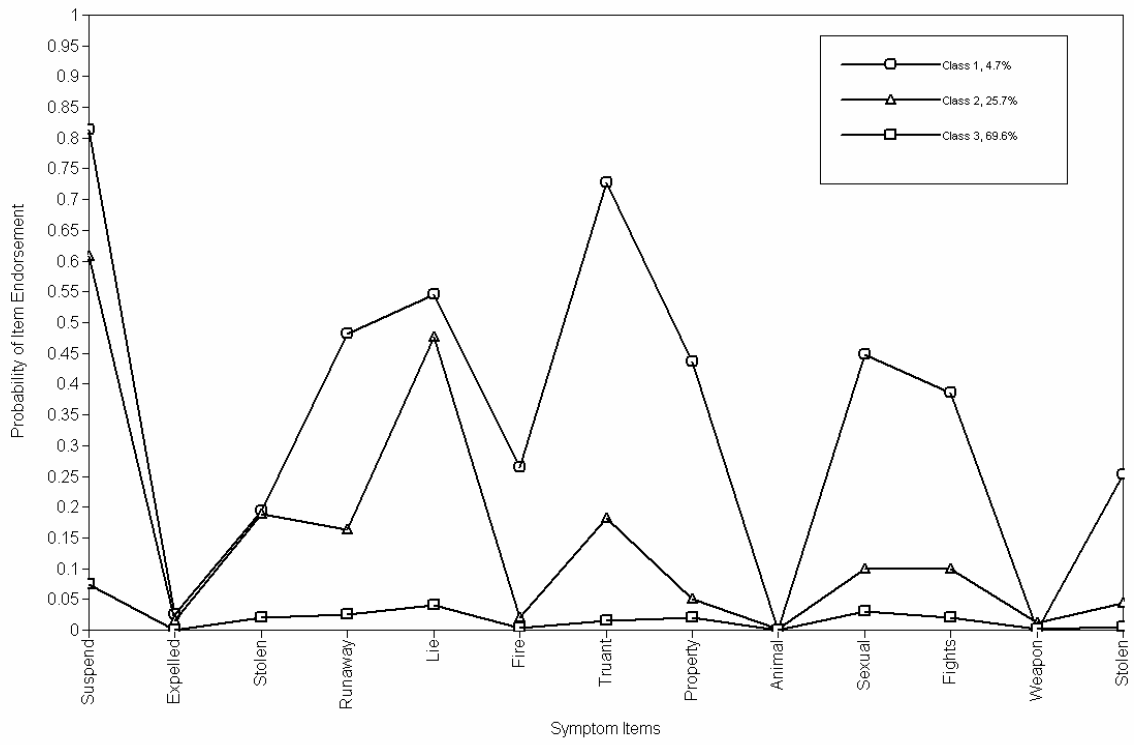*Figure 8.* Conduct Disorder Example: Three-class Latent Class Analysis Profile Plot.

Figure 9. Conduct Disorder Example: Two-class, One-Factor FMM-2 Profile Plot.
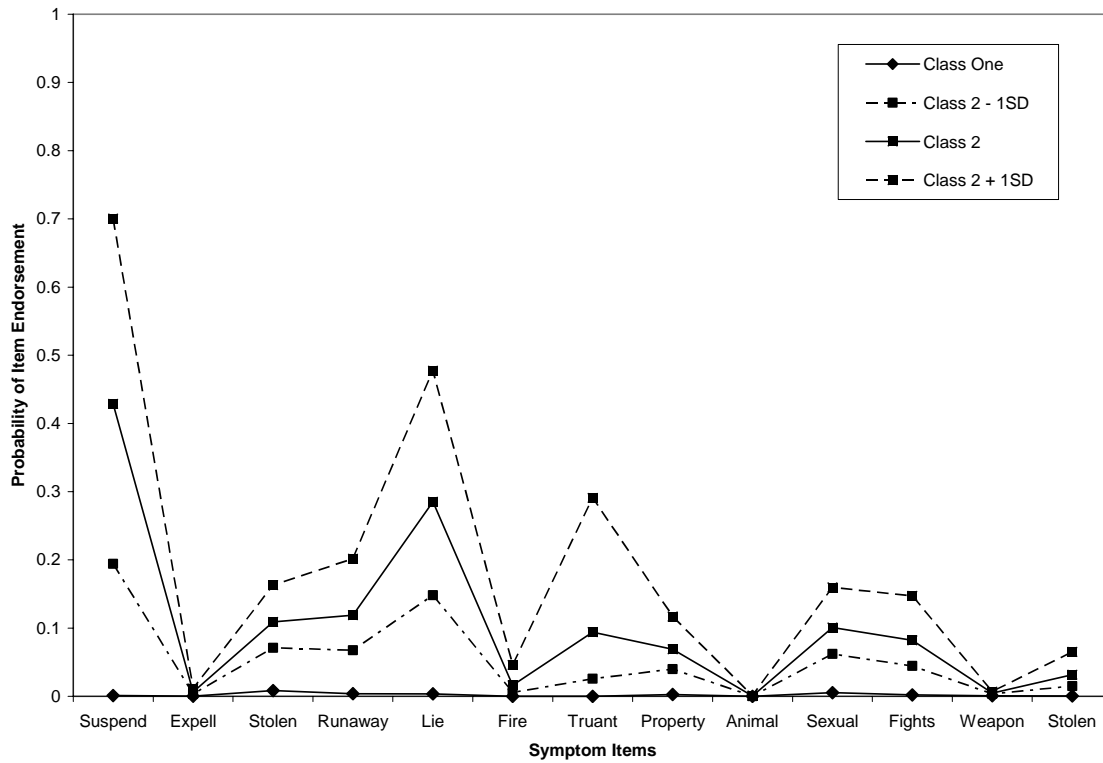
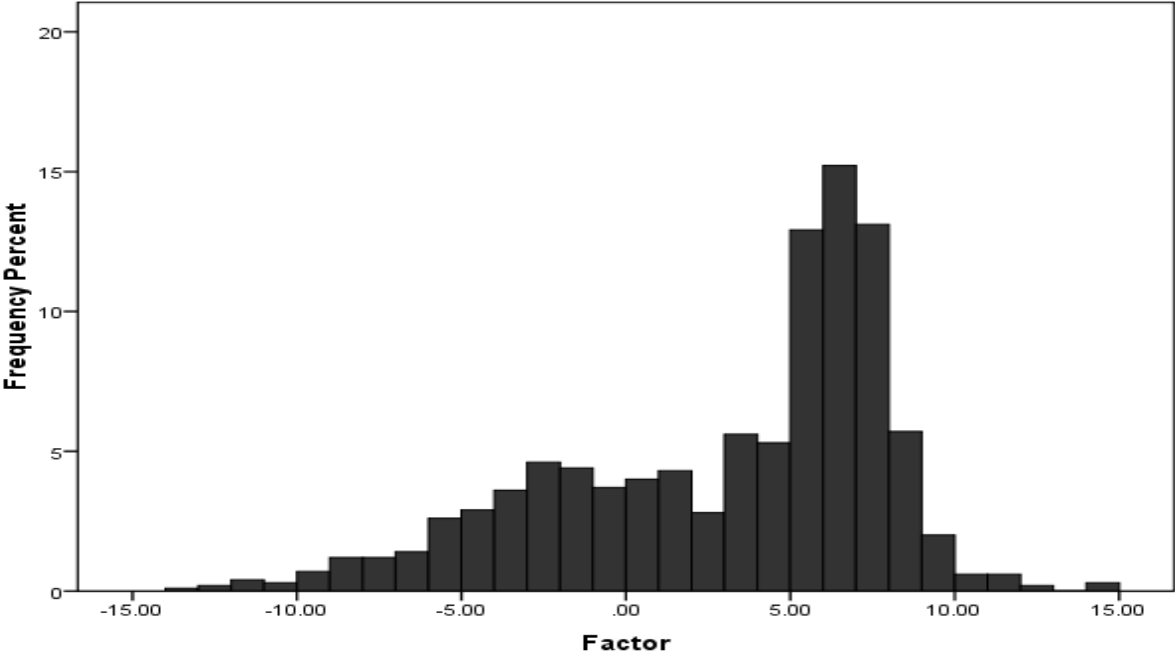*Figure 10.* CD Example: Two-class, One-factor FMM-2 Factor Distribution Plot.

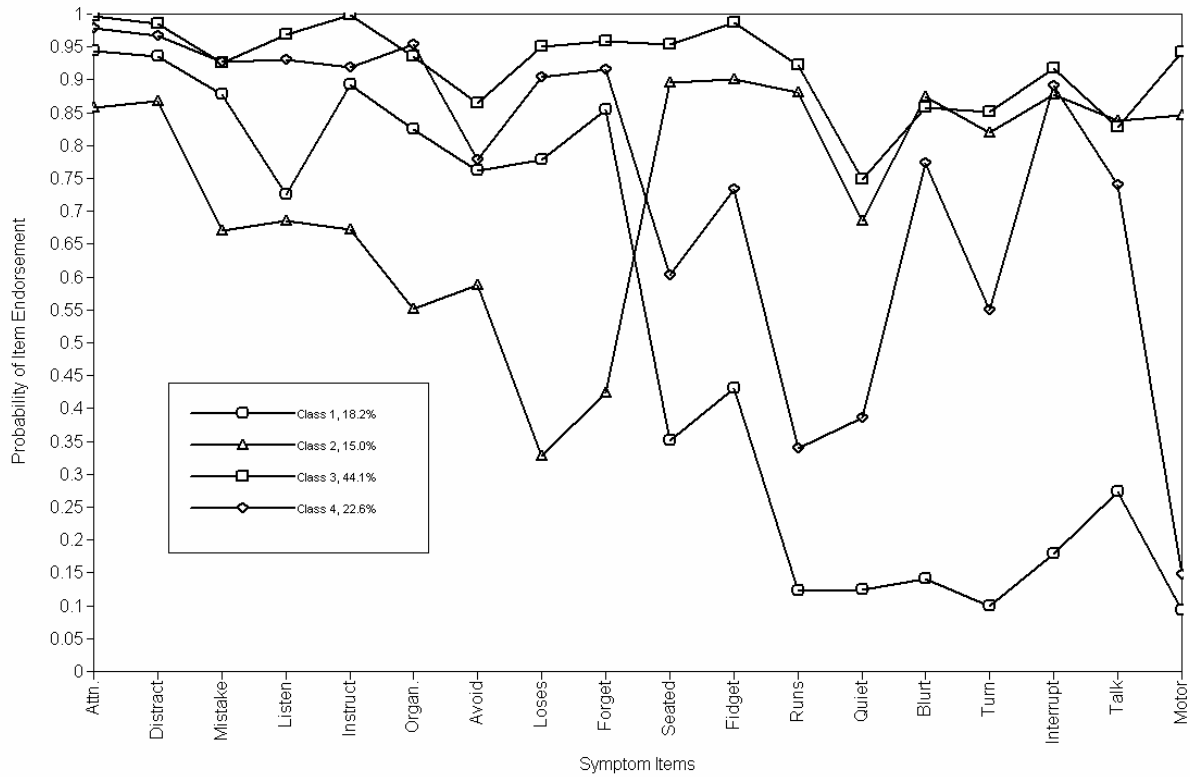*Figure 11.* ADHD Example: Four-class Latent Class Analysis Profile Plot.

Figure 12. ADHD Example: Two-class, Two-factor FMM-3 Profile Plot.