

Interpreting and Testing Interactions in Conditional Mixture Models

Sonya K. Sterba
Vanderbilt University

Mixture modeling applications in psychology often include covariates to explain class membership and aid in construct validation of the latent classification variable. These applications tend to use between-class models involving only main effects of predictors. However, a variety of developmental theories posit interactions among risk and protective variables in predicting membership in trajectory classes or behavioral symptom profiles. This article bridges this disconnect between substantive theory and methodological practice by presenting and comparing two approaches for testing interactive effects of predictors on class membership: product term (PT) and multiple group (MG) approaches. For each approach, we discuss alternative interpretation strategies involving predicted probabilities and odds ratios; we also discuss when the approaches provide equivalent inferences. Published longitudinal and cross-sectional mixture model applications that had originally allowed for only additive effects on class membership are re-analyzed to illustrate the testing and interpretation of interactive effects on class membership using both PT and MG approaches.

Mixture models have become increasingly common in psychology research. Developmental psychology and psychopathology researchers in particular employ longitudinal mixtures (e.g., latent class growth models and growth mixture models) to parsimoniously summarize covariation among repeated measures using discrete latent classes with different trajectories of change over time (e.g., Chassin, Curran, Wirth, Presson, & Sherman, 2009 [substance abuse]; Fontaine, Carbonneau, Vitaro, Barker, & Tremblay, 2009 [antisocial behaviors]; Muthén, 2001 [reading achievement]). Psychologists also commonly use cross-sectional mixture models (e.g., latent class or profile analysis) to account for associations among multiple symptoms or behaviors at a single timepoint (e.g., Aldridge & Roesch, 2008 [coping strategies]; Collins & Lanza, 2010 [adolescent sexual risk behavior]; Liu et al., 2010 [neonatal neurobehaviors]). Whereas unconditional mixture models simply allow the description of discrete patterns of change over time

and description of discrete patterns of symptom endorsement, *conditional* mixture models also allow the prediction of membership in these latent classes (e.g., Chung, Flaherty, & Schafer, 2006; Dayton & Macready, 1988; Lubke & Muthén, 2005; Muthén, 2002; Nagin, 2005).

In principle, conditional mixtures serve an important role in allowing researchers to substantively explain class membership (see Bauer & Curran, 2004; Collins & Lanza, 2010; Muthén, 2003). However, when predictors of class membership are included in empirical applications, often only the sign and significance of effects are interpreted rather than their magnitudes. This practice may reflect broader, persistent¹ issues involving “lack of interpretation” (Long, 2012, p. 28) of results in applications with *observed* nominal outcomes. Ultimately, this practice can limit the potential of conditional mixture results to inform theory in a detailed manner.

Address correspondence to Sonya K. Sterba, Vanderbilt University, Quantitative Methods Program, Department of Psychology & Human Development, Peabody #552, 230 Appleton Place, Nashville, TN 37203. E-mail: sonya.sterba@vanderbilt.edu

¹Long (2012) describes the persistence generally of “limit[ing] interpretation to a table of coefficients with a brief discussion of the signs and significance” or “simply presenting odds ratios without information that allows determination of the magnitude of the effect in terms of changes in probabilities” (p. 28).

In principle, conditional mixture models also allow researchers to flexibly accommodate complex interactions, which can be useful when developmentalists posit interactions among risk and protective variables in predicting membership in trajectory classes or symptom profiles (e.g., Duncan & Keller, 2011; Lenroot & Giedd, 2011; Moffitt, Caspi, & Rutter, 2005; Nugent, Tyrka, Carpenter, & Price, 2011; Sterba & Bauer, 2010). However, empirical applications of conditional mixtures typically include only main effects on class membership, not interaction effects. Furthermore, in empirical applications, multiple-group mixtures are rarely capitalized on for flexibly testing interaction effects on class membership in particular.

Even without including product term or multiple group specifications, mixture models can recover certain kinds of interactions among within-class predictors of y -variable(s) and between-class predictors of class membership (assuming a sufficient number of classes and assuming coefficients of within-class predictors are allowed to vary across classes, e.g., Bauer & Shanahan, 2007; Mathiowetz, 2010; Sterba, 2014; Sterba & Bauer, 2014; Van Horn et al., 2009). This is because class-varying (linear or nonlinear) effects of within-class predictors can be locally weighted by functions of between-class predictors. However, to recover other kinds of interactions, mixture models require explicit product term or multiple-group specifications. For instance, such specifications are needed to recover interactions among predictors of class membership, which could be highly relevant to developmental theory (e.g., Bergman, 2001; Cairns, Bergman, & Kagan, 1998; Magnusson, 1985). In a longitudinal mixture, for instance, the effect of treatment on trajectory class membership could depend on another predictor of class, say, gender. Similarly, the effect of adolescents' perceived stress level on adolescents' internalizing symptom class membership could depend on another predictor of class, say, parental drug use. Such interactions would not be accommodated by conditional mixture specifications commonly applied in practice (e.g., Castelao & Kroner-Herwig, 2013; Cook, Roggman, & D'zatko, 2012; Dyer, Pleck, & McBride, 2012; Greenbaum, Del Boca, Darkes, Wang, & Goldman, 2005; Hepworth, Law, Lawlor, & McKinney, 2010; Lincoln & Takeuchi, 2010; Nagin, 2005; Van Horn et al., 2009). Furthermore, multiple group mixtures (which have mainly focused on multiple-group latent class analysis, e.g., Clogg & Goodman, 1985; Collins & Lanza, 2010; Geiser, Lehmann, & Eid, 2006; McCutcheon, 2002) have seen little motivation generally for testing such interactions among observed predictors of class membership.

In this light, this article focuses on the interpretation and testing of interaction effects on class membership in mixture models. This article has three main goals. First, we describe strategies—involving odds ratios and

predicted probabilities—for interpreting the magnitude of predictors' effects on class membership when predictors affect class additively or interactively. We describe advantages and disadvantages of each interpretation strategy. Second, we compare and contrast product term (PT) and multiple-group (MG) mixture approaches for specifying interactions among predictors of class membership. Third, we show how and when the same hypothesis test results for interaction effects, conditional main effects, and simple effects (defined later) on class membership can be obtained using both PT and MG approaches. To demonstrate these interpretation strategies and to illustrate equivalencies of results across PT and MG approaches, we employ two empirical examples involving adolescent antisocial behavior—one involving a conditional longitudinal mixture and the other involving a conditional cross-sectional mixture. An online appendix available at www.vanderbilt.edu/peabody/sterba/ provides *Mplus* (Muthén & Muthén, 1998–2014) software syntax for empirical example analyses using both PT and MG approaches.

Before continuing, some comments about the generality of this presentation are necessary. Specifically, although the empirical examples involve two kinds of mixture models that differ in the specification of the within-class model, all mixture models (see McLachlan & Peel, 2000; Sterba, 2013 for general reviews) allow the incorporation of covariates predicting class membership. Hence, the points made in the main body of this article generalize to other mixtures as well. To emphasize this generality, we initially focus only on the between-class model specification and only later mention illustrative within-class models for the two empirical examples. Additionally, the points made in this article generalize to settings where covariates are incorporated into the mixture specification in a one-step model-based approach (as done here), or in the third step of a three-step approach (as recently introduced in Asparouhov & Muthén, 2013; Vermunt, 2010).²

ADDITIVE EFFECTS ON CLASS MEMBERSHIP

Consider a mixture model with a categorical latent classification variable that has K classes (where classes are indexed $k = 1 \dots K$). We do not know person i 's class membership, c_i . It is unobserved. However, we can estimate the (marginal) probability that a randomly sampled person i is a member of class k , denoted

²The three-step approach has not yet been presented for multiple group specifications, though this is a feasible extension. An alternative two-step (classify-analyze) approach performs worse than either the one-step or three-step approaches (e.g., Clark & Muthén, 2009; Clogg, 1995) and is not considered here.

$p(c_i = k)$. Moreover, if we know background covariate information about person i , such as gender or treatment status, this information can be incorporated into the model to refine our estimate of the probability that person i is a member of class k —making it a conditional probability. Background covariate information can be incorporated by adopting a multinomial logistic regression specification for the conditional probability of class membership (e.g., Bandeen-roche, Miglioretti, Zeger, & Rathouz, 1997; Wedel, 2002). Researchers are likely familiar with the use of multinomial logistic regression when the dependent variable is an observed, nominal variable (e.g., occupation); see Long (1997) for an excellent review. However, a multinomial logistic regression specification can still be used here when our dependent variable, class membership, is an unobserved nominal variable.³

At times, we may theorize that background covariates, here denoted x_{1i} and x_{2i} , have additive effects on class membership, leading to the following multinomial logistic regression specification:

$$p(c_i = k | x_{1i}, x_{2i}) = \frac{\exp(\beta_0^{(k)} + \beta_1^{(k)} x_{1i} + \beta_2^{(k)} x_{2i})}{\sum_{k=1}^K \exp(\beta_0^{(k)} + \beta_1^{(k)} x_{1i} + \beta_2^{(k)} x_{2i})}. \quad (1)$$

Here, \sum indicates summation, β 's are multinomial logistic regression coefficients, and a k superscript indicates that the parameter can differ across latent class. In this specification, one class is designated as the reference class (here, the last, K th, class). In this reference class, for identification, all coefficients are set to 0: $\beta_0^{(K)} = \beta_1^{(K)} = \beta_2^{(K)} = 0$. All comparisons will be made with respect to this reference class. The choice of the reference (last) class is substantively motivated. Often, a normative/low risk class is chosen as the reference class.⁴

The multinomial coefficients, β 's, are not on an intuitive metric—they are in a log odds (or logit) metric. Therefore, rather than interpreting them directly, we consider two interpretation options based on using these coefficients to (1) calculate odds and odds ratios (OR) and (2) calculate predicted probabilities.

³If there were only two classes, $K=2$, the multinomial logistic specification would simplify and would provide equivalent results to a binary logistic specification.

⁴Suppose that a $K=2$ mixture model is fit where the first class evidences low delinquency whereas the second (reference) class evidences high delinquency. Suppose the researcher had instead desired the low delinquency class to be the reference class. One way to achieve this is to refit the model, placing estimates from the low delinquency class as starting values for the parameters of the last class. Such reordering of the classes does not alter model fit (e.g., McLachlan & Peel, 2000).

Odds and OR Interpretation

Exponentiating the multinomial intercept, $\exp(\beta_0^{(k)})$, converts it into an odds metric (see the Online Appendix for details of this conversion). An *odds* is a ratio of two probabilities (here, the probability of membership in class k versus class K , where all covariates = 0):

$$\exp(\beta_0^{(k)}) = \frac{p(c_i = k | x_{1i} = x_{2i} = 0)}{p(c_i = K | x_{1i} = x_{2i} = 0)}. \quad (2)$$

Specifically, $\exp(\beta_0^{(k)})$ is a baseline⁵ odds of membership in class k relative to membership in the reference class K , where all covariates are equal to 0.

Exponentiating a multinomial slope, e.g., $\exp(\beta_1^{(k)})$, converts it into an *odds ratio* (see the Online Appendix for details of this conversion). This OR is a ratio of the odds before and after a+1 unit difference in x_{1i} , holding x_{2i} constant:

$$\exp(\beta_1^{(k)}) = \frac{p(c_i = k | x_{1i} + 1, x_{2i})}{p(c_i = K | x_{1i} + 1, x_{2i})} \bigg/ \frac{p(c_i = k | x_{1i}, x_{2i})}{p(c_i = K | x_{1i}, x_{2i})}. \quad (3)$$

Specifically, $\exp(\beta_1^{(k)})$ is interpretable as the multiplicative factor by which the odds of membership in class k versus the reference class K changes for a+1 unit difference in x_{1i} , holding x_{2i} constant. An OR = 1 implies that the covariate does not distinguish between membership in these two classes; that is, there is a factor change of 1 in the odds of being in class k (vs. K) for every +1 unit difference in x_{1i} . For instance, if x_{1i} were age, an OR = 1 implies that the odds of being in class k (vs. K) stay the same regardless of every +1 year difference in age. If x_{1i} were gender, an OR = 1 implies that the odds of being in class k (vs. K) are the same for boys versus girls. In contrast, OR > 1 means that a+1 unit difference in x_{1i} increases the odds of being in class k (vs. K), but OR < 1 means that a+1 unit difference in x_{1i} reduces the odds of being in class k (vs. K). Thus, if x_{1i} were age, an OR = 1.55 implies that for a+1 year difference in age, the odds of being in class k (vs. K) increase by a factor of 1.55 (i.e., 55% larger odds). And, if x_{1i} were gender, OR = .30 implies that being male decreases the odds of being in class k (vs. K) by a factor of .30 (i.e., 70% smaller odds).

Predicted Probability Interpretation

Interpreting effects in terms of OR is attractive because the same OR applies regardless of the values of the predictor variables. However, the change in the *probability* of class k versus K implied by a particular factor change in

⁵Subsequently, we use the term “baseline odds” as shorthand to refer to the odds where all covariates are equal to 0.

the odds will depend on the values of the predictors. An alternative to interpreting the effects of predictors on class membership in terms of baseline odds and multiplicative changes in these odds is to use the estimated multinomial coefficients (β 's) together with Equation (1) to compute predicted probabilities, $p(c_i = k | x_{1i}, x_{2i})$, themselves. Predicted probabilities are interpretationally useful because they are not specific to the choice of a reference class and they incorporate multinomial intercept and slope information from all predictors, thus facilitating an integrative understanding of the relationship between the entire set of predictors and class membership.

Predicted probabilities can be plotted or tabled for a range of substantively-interesting chosen values of the covariates; this conveys how the probability of class membership varies as a function of predictors. It is also possible to compute discrete change in predicted probabilities as certain covariate(s) are changed by a specified amount whereas others are held constant (perhaps held at their means). (Researchers should be aware that discrete changes in predicted probabilities holding covariates at their means can differ somewhat from discrete changes in predicted probabilities calculated case-wise and then averaged across cases; Long, 1997.)

PRODUCT TERM (PT) APPROACH FOR INCORPORATING INTERACTIVE EFFECTS ON CLASS MEMBERSHIP

The previous subsection concerned additive effects of predictors on class membership in a conditional mixture model. Alternatively, researchers may theorize that x_{1i} and x_{2i} have interactive effects, such that the effect of x_{1i} , say, age, on class membership varies across levels of x_{2i} , say, gender. This can be accommodated in a conditional mixture by including the product term $x_{1i}x_{2i}$ as another predictor of class membership in the multinomial logistic regression:

$$p(c_i = k | x_{1i}, x_{2i}) = \frac{\exp\left(\beta_0^{(k)} + \beta_1^{(k)}x_{1i} + \beta_2^{(k)}x_{2i} + \beta_3^{(k)}x_{1i}x_{2i}\right)}{\sum_{k=1}^K \exp\left(\beta_0^{(k)} + \beta_1^{(k)}x_{1i} + \beta_2^{(k)}x_{2i} + \beta_3^{(k)}x_{1i}x_{2i}\right)}. \quad (4)$$

Again, all coefficients associated with the reference (K th) class are set to 0 for identification: $\beta_0^{(K)} = \beta_1^{(K)} = \beta_2^{(K)} = \beta_3^{(K)} = 0$.

Odds and OR Interpretation

In this interactive model, the OR for x_{1i} , (ratio of the odds of class k vs. K for a+1 unit difference in x_{1i}) will differ depending on x_{2i} , and vice versa. Particularly, considering

x_{1i} (age) the focal predictor and x_{2i} (gender) the moderator, $\exp\left(\beta_3^{(k)}\right)$ is the multiplicative factor by which the OR for x_{1i} (i.e., $\exp\left(\beta_1^{(k)}\right)$) changes for boys ($x_{2i}=1$) versus girls ($x_{2i}=0$). Equivalently, considering x_{2i} (gender) the focal predictor and x_{1i} (age) the moderator, $\exp\left(\beta_3^{(k)}\right)$ is the multiplicative factor by which the OR for x_{2i} (i.e., $\exp\left(\beta_2^{(k)}\right)$) changes given a+1 unit difference in age (x_{1i}). The lower-order effects of each predictor involved in the interaction (here x_{1i} and x_{2i}) are now conditional main effects. For instance, the OR for x_{1i} , $\exp\left(\beta_1^{(k)}\right)$ is now interpretable as the factor by which the odds of membership in class k vs. K changes for a+1 unit difference in x_{1i} , holding the other predictor involved in the interaction (x_{2i}) at 0 and holding any other predictors constant (here, there are no other such predictors). Interpretation of conditional main effects for continuous predictors is aided by centering these predictors at meaningful values. Continuous predictors in both empirical examples will be centered, as described later. Centering also reduces nonessential multicollinearity among predictors (see Aiken & West, 1991).

Predicted Probability Interpretation

Additionally, predicted probabilities can be computed at chosen values of the covariates using Equation (4), and plotted or tabled to show how the probability of class membership can be expected to change as a function of different combinations of covariate values. Note that the plotting utility for predicted probabilities of class membership in some mixture modeling software (e.g., *Mplus* Version 7.11) assumes that predictors of class enter additively, not interactively. This utility allows the plotting of predicted probabilities of class membership across the range of one predictor, holding all other predictors at fixed chosen values. However, in an interactive model, the value of the product $x_{1i}x_{2i}$ is already determined once the values of each constituent variable (x_{1i} and x_{2i}) are chosen. To use this utility to depict interactive relationships of predictors with class, judicious centering would be needed to ensure that the product term equals zero at the desired moderator value.

MULTIPLE GROUP (MG) APPROACH FOR INCORPORATING INTERACTIVE EFFECTS ON CLASS MEMBERSHIP

If theory suggests interactive relationships among predictors of class membership, the approach discussed thus far involves using a conditional mixture model that incorporates product terms as in Equation (4) (i.e., the

PT approach). When at least one of the predictors involved in the interaction is categorical (say, if x_{2i} were gender or treatment), an alternative approach can be used that avoids the need to construct and incorporate product terms as in Equation (4). This approach employs a multiple-group (MG) mixture model, where the observed categorical predictor x_{2i} is treated as the grouping variable. Note that, in this article, we use “group” when referring to levels of a manifest categorical predictor and latent “class” when referring to levels of a latent classification variable. We first introduce a constrained version of a MG mixture for the purpose of conveniently testing such interactions, and then discuss how MG mixtures also afford the flexibility to test assumptions that are untestable using standard conditional mixtures.

Instead of designating a categorical x_{2i} as a covariate, we now designate it as a manifest grouping variable, with $G+1$ levels ($g=0 \dots G$). Here $G=1$. For instance, in a later empirical example the binary grouping variable is gender. (Note that if $G>1$, then the PT approach presented earlier would require representing this categorical covariate by a set of G coding variables and representing the interaction by product terms of x_{1i} with *each* coding variable.)

A MG mixture assumes the same number of latent classes, K , across groups (i.e., configural invariance); this is implicitly assumed by the PT mixture as well. In a K -class MG mixture, *in theory* all model parameters

could be allowed to differ across manifest groups. A g superscript on a parameter indicates that it can differ across manifest groups. However, to specify a MG mixture equivalent to the PT approach discussed earlier, we assume across-group equality of all within-class model parameters (see Empirical Examples for some illustrative within-class models). Only between-class parameters are allowed to differ across group, as follows:

$$p(c_i = k | x_{1i}, g_i) = \frac{\exp(\beta_0^{(k|g)} + \beta_1^{(k|g)} x_{1i})}{\sum_{k=1}^K \exp(\beta_0^{(k|g)} + \beta_1^{(k|g)} x_{1i})}. \quad (5)$$

As an alternative to Equation (4), in Equation (5) only x_{1i} is treated as a covariate and all multinomial coefficients are now conditional on group (i.e., x_{2i}). Table 1 shows how, for a categorical x_{2i} , tests of conditional main effects and interaction effects are accomplished equivalently in the PT approach of Equation (4) versus the MG approach of Equation (5). To summarize, testing that the coefficient for $x_{2i}=0$ or that the coefficient for $x_{1i}x_{2i}=0$ in the PT approach corresponds, in the MG approach, to testing that the multinomial intercepts or multinomial slopes of x_{1i} , respectively, are equal across group. Tests that a predictor (or product term) does not distinguish a specific class k from the reference class K are 1 degree of freedom (df) tests. Tests that a predictor (or product term) does not distinguish any

TABLE 1
Testing Hypotheses Regarding Prediction of Class Membership Involving Interactive Effects in the Product Term (PT) Versus Multiple Group (MG) Mixture Frameworks

	<i>PT mixture approach with a binary x_{2i} In Equation (4):</i>	<i>MG mixture approach with a binary x_{2i} serving as a grouping variable. In Equation (5):</i>
Does the conditional main effect of $x_{2i} \dots$...help distinguish membership in class k vs. K ?	Test* $\beta_2^{(k)} = 0$ ($df = 1$)	Test [†] $\beta_0^{(k g)} = \beta_0^{(k)}$ ($df = 1$)
...help distinguish membership in any class vs. the reference class K ?	Test [†] $\beta_2 = \mathbf{0}$ ($df = K-1$) where $\beta_2 = \{\beta_2^{(1)}, \dots, \beta_2^{(K-1)}\}$	Test [†] $\beta_0^g = \beta_0$ ($df = K-1$) where $\beta_0^g = \{\beta_0^{(1 g)}, \dots, \beta_0^{(K-1 g)}\}$ $\beta_0 = \{\beta_0^{(1)}, \dots, \beta_0^{(K-1)}\}$
Does the interaction term $x_{1i}x_{2i} \dots$...help distinguish membership in class k vs. K ?	Test* $\beta_3^{(k)} = 0$ ($df = 1$)	Test [†] $\beta_1^{(k g)} = \beta_1^{(k)}$ ($df = 1$)
...help distinguish membership in any class vs. the reference class K ?	Test [†] $\beta_3 = \mathbf{0}$ ($df = K-1$) where $\beta_3 = \{\beta_3^{(1)}, \dots, \beta_3^{(K-1)}\}$	Test [†] $\beta_1^g = \beta_1$ ($df = K-1$) where $\beta_1^g = \{\beta_1^{(1 g)}, \dots, \beta_1^{(K-1 g)}\}$ $\beta_1 = \{\beta_1^{(1)}, \dots, \beta_1^{(K-1)}\}$
Does the conditional main effect of $x_{1i} \dots$...help distinguish membership in class k vs. K ?	Test* $\beta_1^{(k)} = 0$ ($df = 1$)	Test [†] $\beta_1^{(k g=0)} = 0$ ($df = 1$)
...help distinguish membership in any class vs. the reference class K ?	Test [†] $\beta_1 = \mathbf{0}$ ($df = K-1$) where $\beta_1 = \{\beta_1^{(1)}, \dots, \beta_1^{(K-1)}\}$	Test [†] $\beta_1^{g=0} = \mathbf{0}$ ($df = K-1$) where $\beta_1^{g=0} = \{\beta_1^{(1 g=0)}, \dots, \beta_1^{(K-1 g=0)}\}$

Note. [†] = This can be a Wald test, for instance where $W = [\mathbf{L}\hat{\boldsymbol{\theta}}]'[\mathbf{L}\widehat{\text{Var}}(\hat{\boldsymbol{\theta}})\mathbf{L}]^{-1}[\mathbf{L}\hat{\boldsymbol{\theta}}]$. \mathbf{L} is a contrast matrix, $\hat{\boldsymbol{\theta}}$ is a vector of estimates of *all* multinomial coefficients involved to the left of the equal sign in the table expressions. For instance, in row 2 column 2 of the table, $\hat{\boldsymbol{\theta}} = \beta_2$, whereas in row 2 column 3, $\hat{\boldsymbol{\theta}} = [\beta_0^{g=0}, \beta_0^{g=1}]'$. $\widehat{\text{Var}}(\hat{\boldsymbol{\theta}})$ is the asymptotic covariance matrix for estimates in $\hat{\boldsymbol{\theta}}$ (obtained with TECH3 in Mplus). W will be χ^2 (df) distributed under the null hypothesis; this test is implemented in Mplus with the *model test* command. * = This can be a z -test (where $z = \hat{\beta}^{(k)} / SE_{\hat{\beta}^{(k)}}$), which is \sqrt{W} where $df=1$). In this table, invariance across manifest groups is assumed for within-class parameters in the MG mixture.

class from the reference class K are $K-1$ *df* tests (e.g., multivariate Wald tests, as described in Table 1, or likelihood ratio tests). Odds ratios and predicted probability interpretation strategies can be used in the MG mixture approach similarly to the PT mixture approach.

Table 1 mentions only x_{1i} and x_{2i} . If, as in empirical example 1, additional covariates are present, the results of tests in columns 1 and 2 of Table 1 would still correspond; each test would simply control for the additional covariates. Also, Table 1 concerns prediction of class membership involving only interaction effects. A constrained special case of the MG mixture model, where $\beta_1^{(k|g)} = \beta_1^{(k)}$, can also be used to test additive effects of predictors on class membership and can yield equivalent results to Equation (1), for categorical x_{2i} .

COMPARING THE PT VS. MG MIXTURE APPROACHES

Three main considerations should be kept in mind when deciding between the PT vs. MG mixture approaches, as follows. First, as stated earlier, the MG approach requires that the moderator, x_{2i} , be categorical. If it is not, the PT mixture approach would typically be considered preferable to dichotomization (see MacCallum, Zhang, Preacher, & Rucker, 2002; Shentu & Xie, 2010).

Second, beyond testing whether the interaction of x_{1i} and x_{2i} is significant, a researcher may be interested in testing whether a focal predictor, say x_{1i} , can significantly distinguish membership in class k vs. K at particular chosen values of a moderator, x_{2i} . In the PT approach, rearranging Equation (4) indicates that the *simple slope* of x_{1i} at a chosen conditional value of x_{2i} , here denoted \tilde{x}_{2i} , is: $\beta_1^{(k)} + \beta_3^{(k)}\tilde{x}_{2i}$. This simple slope can be tested using familiar procedures from multiple linear regression for “probing” interaction effects (e.g., Aiken & West, 1991). Specifically, a *z*-test of this simple slope can be computed using the following (delta-method) standard error:

$$SE_{\beta_1^{(k)} + \beta_3^{(k)}\tilde{x}_{2i}} = \sqrt{VAR_{\beta_1^{(k)}} + 2\tilde{x}_{2i}COV_{\beta_1^{(k)}\beta_3^{(k)}} + \tilde{x}_{2i}^2VAR_{\beta_3^{(k)}}}. \quad (6)$$

This formula includes asymptotic variances and covariances of multinomial coefficients (which can be obtained, for instance, using TECH3 in *Mplus*) or alternatively the entire test can be performed using a model constraint (as shown in the Online Appendix example syntax). In contrast, in the MG approach, simple slopes of x_{1i} for each value of the moderator grouping variable, x_{2i} , are already directly-estimated model parameters. This is convenient particularly when the grouping

variable is expected to moderate effects of several predictors on class membership (as in Example 1).

Third, another strength of the MG approach is that it easily allows investigation of invariance of within-class model parameters across levels of the manifest grouping variable. Invariance would be tested by comparing models in which the within-class parameters are constrained equal across group vs. unconstrained across group. Depending on the nature of the within-class model, conventions may exist regarding the order in which invariance of within-class parameters is tested (see, e.g., Millsap, 2011). In contrast, the standard PT mixture approach a priori assumes invariance of the within-class model parameters across levels of this grouping variable. If this assumption is not met, estimated parameters can be biased. Although an expansion⁶ of the PT mixture approach could indeed be used to test across-group invariance of within-class intercepts and slopes, it is still the case that testing invariance of within-class variance parameters such as residual variances (for continuous y 's) or random effect variances or factor variances (if present) would be more cumbersome than in the MG mixture approach (relatedly, see Muthén & Asparouhov, 2002; Oberski, 2014; Snijders & Bosker, 1999, Sect. 8.1). For categorical y 's, within-class residual variances typically are not estimable parameters, and their invariance across levels of the x_{2i} grouping variable is implicitly assumed in both PT and MG mixture approaches used here (see Long, 2009).

This article's perspective on the general relationship between PT and MG mixture approaches is somewhat different from the perspective of Collins and Lanza (2010), which was given in the context of latent class analysis (LCA) specifically. They argue that the covariate (PT) and MG LCAs “reflect different points of view about the grouping variable” such that in the MG LCA “the grouping variable is automatically allowed to moderate the effect of each covariate on latent class membership” (p. 168). However, if not all of these across-group differences in covariate effects on class membership are substantively motivated, this strategy can sacrifice parsimony (e.g., Cohen, 1968). As the number of across-group parameter constraints is relaxed, a larger sample size should be necessary to maintain the same amount of precision. Instead, we conceptualized PT and MG mixtures as two modeling approaches that can achieve the exact same inferential goals given categorical moderator(s), with differences in convenience for specifying interactions (because MG mixtures do not require use of product terms to represent interactions), and in flexibility (because MG

⁶This extension would involve regressing outcomes within-class (here, y_{ij}) on x_{2i} as well as allowing x_{2i} to interact with each within-class predictor.

mixtures more easily allow optional investigation of invariance assumptions for the within-class submodel).

Also, Collins and Lanza (2010) describe hypothesis tests afforded by the PT versus MG mixture as being somewhat different. They use the MG mixture approach to test a combination of a conditional main effect plus interaction, but use the PT mixture approach with a categorical moderator to test each of these two effects separately (pp. 170–171). Here, in Table 1 we instead articulate how the exact same omnibus or individual-parameter hypothesis tests can be obtained from both modeling frameworks, and we illustrate this in the following empirical examples.

In the next two sections, two empirical examples are used to illustrate interpretation of interactive effects on class membership. Each example is an extended version of a previously published analysis which originally investigated only additive effects of predictors on class membership—specifically, Sterba (2014) and Muthén (2002). Here, in Examples 1 and 2, the previously published analyses are extended to investigate interactive effects of predictors on class membership. Example 1 is a fully worked example, whereas Example 2 is focused on one particular result to avoid redundancy. When covariates predict only class membership and do not enter the within-class model, as in Examples 1 and 2, class enumeration results have been shown to correspond whether an unconditional or conditional mixture is used for enumeration (e.g., Bandeen-roche et al., 1997; Collins & Lanza, 2010; Lubke & Muthén, 2007; Nagin, 2005). Hence, in both examples, we continue to use the same K that was employed in each original analysis, and we focus on results of the between-class model. Syntax for Example 1 and 2 analyses is in the Online Appendix.

EMPIRICAL EXAMPLE 1

A previous study (Sterba, 2014) summarized individual variation in young adult aggressive behavior across the transition out of foster care facilities using $K = 3$ latent trajectory classes of aggressive behavior (high-stable, declining, and low-stable). The aggressive behavior repeated measures were collected yearly for three years starting at age 17–18 (see Courtney & Cusick, 2007 for details).⁷ In Sterba (2014), latent trajectory classes of aggressive behavior were predicted by four time-invariant covariates that were assumed to have additive effects on

⁷Exact ages were not made available for participants in this dataset, at www.icpsr.umich.edu. The analysis sample of $N = 438$ consists of young adults with at least one outcome and all covariate data present. Under missing-at-random assumptions, missing outcomes were accommodated with full information maximum likelihood, implemented with the Expectation-Maximization algorithm.

class membership: presence of an alcohol or substance abuse diagnosis at time 1 (d_i), male gender (m_i), level of social support, (s_i , a standardized scale score from Sherbourne and Stewart (1991)), and presence of a prior arrest record, (a_i). Although, in this prior study, gender, social support, and an arrest record did distinguish membership among some classes, controlling for substance abuse, there is theoretical reason to suspect that this model may be too parsimonious. The effect of the protective factors such as social support and gender (being female) on aggression trajectory class membership may depend on the levels of a risk factor, such as, prior arrest record (e.g., Yoshikawa, 1994; Youngstrom, Weist, & Albus, 2003). That is, having higher social support and being female may be most protective for those with a prior delinquency history. Here, revisiting this previous $K = 3$ analysis, Example 1 investigates the interactive effect of arrest \times social support ($s_i a_i$) and arrest \times male ($m_i a_i$) on class membership.

The *within-class* model for Example 1 uses a latent class growth model (LCGM) specification (e.g., Muthén, 2001; Nagin, 1999) for binary aggression⁸ repeated measures:

$$\pi_{ij}^{(k)} = \frac{1}{1 + \exp\left(\tau_j^{(k)} - \left(\eta_0^{(k)} + \eta_1^{(k)} \text{time}_{ij}\right)\right)}. \quad (7)$$

Here, $\tau_j^{(k)}$ is a threshold for repeated measure j (where $j = 1 \dots J$) in class k that is fixed to 0 for identification. $\pi_{ij}^{(k)}$ is the probability of endorsing repeated measure j for person i (where $i = 1 \dots N$) within class k . It is allowed to depend on intercept, $\eta_0^{(k)}$ and linear, $\eta_1^{(k)}$ aspects of change. These aspects of change differ across latent classes (as indicated by the k superscript on these parameters), to accommodate differently-shaped class-specific trajectories. time_{ij} was centered at initial status.

Using the PT mixture approach, adapting Equation (4), the Example 1 *between-class model* is:

$$p(c_i = k | \mathbf{x}_i) = \frac{\exp\left(\beta_0^{(k)} + \beta_1^{(k)} d_i + \beta_2^{(k)} a_i + \beta_3^{(k)} m_i + \beta_4^{(k)} s_i + \beta_5^{(k)} m_i a_i + \beta_6^{(k)} s_i a_i\right)}{\sum_{k=1}^K \exp\left(\beta_0^{(k)} + \beta_1^{(k)} d_i + \beta_2^{(k)} a_i + \beta_3^{(k)} m_i + \beta_4^{(k)} s_i + \beta_5^{(k)} m_i a_i + \beta_6^{(k)} s_i a_i\right)}, \quad (8)$$

where \mathbf{x}_i is a vector of all covariates in the between-class model. Since two anticipated interactions involved a

⁸Endorsement of an aggressive conduct offense occurred if an adolescent over the past 12 months: fought in a group, shot or stabbed, pulled a knife or gun, badly injured someone, or threatened with a weapon. It would be possible to adopt other representations of this aggression construct.

categorical moderator (prior arrest status, a_i), an equivalent specification under the MG mixture approach uses a_i as a manifest grouping variable, and assumes invariance of the within-class model across arrest status. a_i can take on values $g = 0 \dots G$. Herein, $G = 1$. Using this MG mixture approach, adapting Equation (5), the Example 1 *between-class model* is instead:

$$p(c_i = k | \mathbf{x}_i, a_i = g) = \frac{\exp(\beta_0^{(k|g)} + \beta_1^{(k)} d_i + \beta_3^{(k|g)} m_i + \beta_4^{(k|g)} s_i)}{\sum_{k=1}^K \exp(\beta_0^{(k|g)} + \beta_1^{(k)} d_i + \beta_3^{(k|g)} m_i + \beta_4^{(k|g)} s_i)}. \quad (9)$$

In the fitted LCGM, within class k , repeated measure j follows a Bernoulli probability mass function (PMF) where $p(y_{ij} | c_i = k) = \left(\pi_{ij}^{(k)}\right)^{y_{ij}} \left(1 - \pi_{ij}^{(k)}\right)^{1-y_{ij}}$. Under the assumption that repeated measures are locally independent within-class, their PMFs are multiplied to form the joint PMF for all J repeated measures within class, $p(\mathbf{y}_i | c_i = k)$. In the PT mixture, this within-class joint PMF is weighted by the between-class model in Equation (8) and summed across class to form the full model PMF:

$$p(\mathbf{y}_i | \mathbf{x}_i) = \sum_{k=1}^K p(c_i = k | \mathbf{x}_i) p(\mathbf{y}_i | c_i = k). \quad (10)$$

For more details, see Sterba (2013). In the MG mixture approach, the within-class joint PMF is instead multiplied by Equation (9) and summed across class and group to form the full model PMF:

$$p(\mathbf{y}_i | \mathbf{x}_i) = \sum_{g=0}^G \sum_{k=1}^K p(a_i = g) p(c_i = k | \mathbf{x}_i, a_i = g) p(\mathbf{y}_i | c_i = k). \quad (11)$$

Here, $p(a_i = g)$ indicates the probability that the observed grouping variable takes on the value g . Note that degrees of freedom in these PT and MG mixtures match when $p(a_i = g)$ is fixed to the observed proportion (rather than counted as an estimated parameter).

Results

The within-class model of our $K = 3$ LCGM⁹ summarized individual variation in aggressive behavior trajectories

⁹The previous study found $K = 3$ classes to be best fitting using a more complex within-class model that involved time-varying covariates, which were not used here. Yet the $K = 3$ conditional LCGM used here had marginal class trajectories nearly identical to the previous study and was also found best fitting according to information criteria (for AIC, $K = 2$: 1397.88, $K = 3$: 1344.82, $K = 4$: 1348.16; for BIC, $K = 2$: 1442.78; $K = 3$: 1426.46; $K = 4$: 1466.55).

with a low-stable aggression class (marginally 44% of participants, where $\eta_0^{(1)} = -4.19, \eta_1^{(1)} = .90$), a high-stable aggression class (marginally 18%, where $\eta_0^{(2)} = 2.03, \eta_1^{(2)} = -.29$), and a declining class (marginally 38%, where $\eta_0^{(3)} = 3.03, \eta_1^{(3)} = -3.91$). By time 3, class 3 had decreased to approximately the same level as the low-stable class 1. Prior to turning to the prediction of class membership and interpretation of results using OR and predicted probabilities, a possible preliminary step could involve testing invariance of within-class growth parameters across arrest-status group, i.e.,: $\eta_0^{(k|g)} = \eta_0^{(k)}$ and $\eta_1^{(k|g)} = \eta_1^{(k)}$, under assumptions mentioned earlier (Long, 2009). Doing so, we find that we cannot reject the null hypothesis of invariance using a multivariate Wald test (5.97 (6), $p = .43$). Turning then from the within-class model to the between-class model, we may consider starting with the omnibus tests described in Table 1, each of which evaluates the hypothesis that a predictor cannot discriminate membership in *any* class vs. the reference (declining) class. Results of these omnibus Wald tests for Example 1 are given in Table 2; these results apply to both the PT and MG mixtures, as was described in Table 1. Table 2 results indicate that two conditional main effects (gender, prior arrest), the covariate substance abuse, and the interaction of support \times arrest each significantly distinguish at least one class versus the reference (declining) class.

Odds and OR Interpretation

More specific information regarding differentiating membership in a particular class k from the reference class can be obtained for each predictor in terms of odds and OR's; this information is provided in Table 3 for PT mixture results and Table 4 for MG mixture results. Confidence intervals for ORs are provided in Tables 3 and 4; for instance, the CI for $\exp(\beta_1^{(k)})$ would be $\exp(\beta_1^{(k)} \pm z_{1-\alpha/2} SE_{\beta_1^{(k)}})$ (e.g., Hosmer & Lemeshow, 2000). Effects superscripted by the same letter across Tables 3 and 4 are equivalent in both models; effects that are directly estimated in only one model can also

TABLE 2
Multivariate Wald Tests for Each Predictor Distinguishing Membership in at Least One Class Versus Class K Using the Product Term (PT) or Multiple Group (MG) Mixture in Example 1

Predictor of class membership	Wald test
Substance abuse	6.09 (2), $p = .048$
Prior arrest	14.13 (2), $p = .001$
Gender	13.24 (2), $p = .001$
Social support	0.41 (2), $p = .814$
Gender \times arrest	3.32 (2), $p = .190$
Support \times arrest	6.22 (2), $p = .045$

TABLE 3
Product Term (PT) Mixture Results for Example 1: Odds Ratios for Each Predictor Distinguishing Membership in Class *k* Versus *K*

Multinomial coefficient	Class 1 (low-stable) vs. class 3 (declining)			Class 2 (high-stable) vs. class 3 (declining)		
	Est. (SE)	Exp (Est)	CI	Est. (SE)	Exp(Est)	CI
Intercept ^a	1.11* (.34)	3.03	(1.56, 5.91)	-1.08* (.54)	.34	(.12, .98)
Substance abuse ^b	-0.78* (.34)	.46	(.24, .89)	0.01 (.39)	1.01	(.47, 2.17)
Arrest	-1.30* (.35)	.27	(.14, .54)	-0.63 (.57)	.53	(.17, 1.63)
Male ^c	-0.94* (.36)	.39	(.19, .79)	0.51 (.55)	1.67	(.57, 4.89)
Support ^d	-0.11 (.20)	.90	(.61, 1.33)	-0.17 (.30)	.84	(.47, 1.52)
Male × arrest	0.71 (.54)	2.04	(.71, 5.86)	1.23 (.73)	3.43	(.82, 14.31)
Support × arrest	0.71* (.29)	2.04	(1.15, 3.59)	0.49 (.38)	1.64	(.78, 3.44)

Note. * = $p < .05$. Superscripted letters indicate that an effect is equal across the modeling approaches in Tables 3 versus 4. For the intercept, Exp(Est) is a baseline odds. For each slope, Exp(Est) an odds ratio. CI = confidence interval, $\text{Exp}(\text{Est} \pm 1.96 \times SE)$.

be tested in the other model using procedures described in Table 1. Results for intercepts indicate that, for non-substance abusing, nonarrested females with average social support, the baseline odds of low-stable aggression (class #1) were higher and baseline odds of high-stable aggression (#2) were lower, each compared to the declining aggression class (#3). But the odds of low-stable aggression (vs. declining aggression) decrease by 54% for having a substance abuse diagnosis, and decrease 61% for being male—given that one has not been arrested and controlling for other predictors. Interesting results also included a significant interaction of social support × arrest, for distinguishing the low-stable (#1) versus declining (#3) classes. Simple slopes for social support, considering prior arrest as moderator, were conveniently already directly estimated in the MG mixture (Table 4). For those with an arrest record, a +1 standard deviation difference in social support increased by 82% the odds of low-stable versus declining aggression (controlling for gender and substance abuse). But, for those without an arrest record, the effect of social support is nonsignificant.

Predicted Probability Interpretation

Next, we substitute the values of the estimated multinomial coefficients in Tables 3 (or 4) into Equation (8) (or (9)) and depict predicted probabilities at a meaningful combination of covariate values in Figure 1, here chosen specifically to aid in interpreting the support × arrest interaction. Figure 1 depicts the probability of membership (vertical axis) in classes 1, 2, or 3, across the observed range of social support (horizontal axis) where prior arrest = 0 (Panel A) or prior arrest = 1 (Panel B), holding other covariates at their means. For youth with a prior arrest record, Figure 1 shows that the more support received, the more likely they are to be in a low-stable aggression class (#1). If youth with a prior arrest record receive the least support, they are most likely (probability = .81) to be in class 3 (declining aggression), but if they receive maximal support, they are most likely (probability = .47) to be in class 1. On the other hand, for youth without a prior arrest record, the low-stable class (#1) is most likely across the range of social support. Furthermore, for these youth declining (#3) and high-stable (#2)

TABLE 4
Multiple-Group (MG) Mixture Results for Example 1: Odds Ratios for Each Predictor Distinguishing Membership in Class *k* Versus *K*

Multinomial coefficient	Class 1 (low-stable) vs. class 3 (declining)			Class 2 (high-stable) vs. class 3 (declining)		
	Est. (SE)	Exp(Est)	CI	Est. (SE)	Exp(Est)	CI
Intercept arrest	-0.19 (.45)	.83	(.34, 2.00)	-1.72* (.57)	.18	(.06, .55)
Intercept no arrest ^a	1.11* (.34)	3.03	(1.56, 5.91)	-1.08* (.54)	.34	(.12, .98)
Substance abuse ^b	-0.78* (.34)	.46	(.24, .89)	0.01 (.39)	1.01	(.47, 2.17)
Male arrest	-0.22 (.41)	.80	(.36, 1.79)	1.74* (.51)	5.70	(2.10, 15.48)
Male no arrest ^c	-0.94* (.36)	.39	(.19, .79)	0.51 (.55)	1.67	(.57, 4.89)
Support arrest	0.60* (.23)	1.82	(1.16, 2.86)	0.33 (.23)	1.39	(.89, 2.18)
Support no arrest ^d	-0.11 (.20)	.90	(.61, 1.33)	-0.17 (.30)	.84	(.47, 1.52)

Note. In the multiple-group (MG) mixture, prior arrest status (arrest = 1, no arrest = 0) is the grouping variable. * = $p < .05$. Superscripted letters indicate that an effect is equal across the modeling approaches in Tables 3 versus 4. For each intercept, Exp(Est) is a baseline odds. For each slope, Exp(Est) an odds ratio. CI = confidence interval, $\text{Exp}(\text{Est} \pm 1.96 \times SE)$.

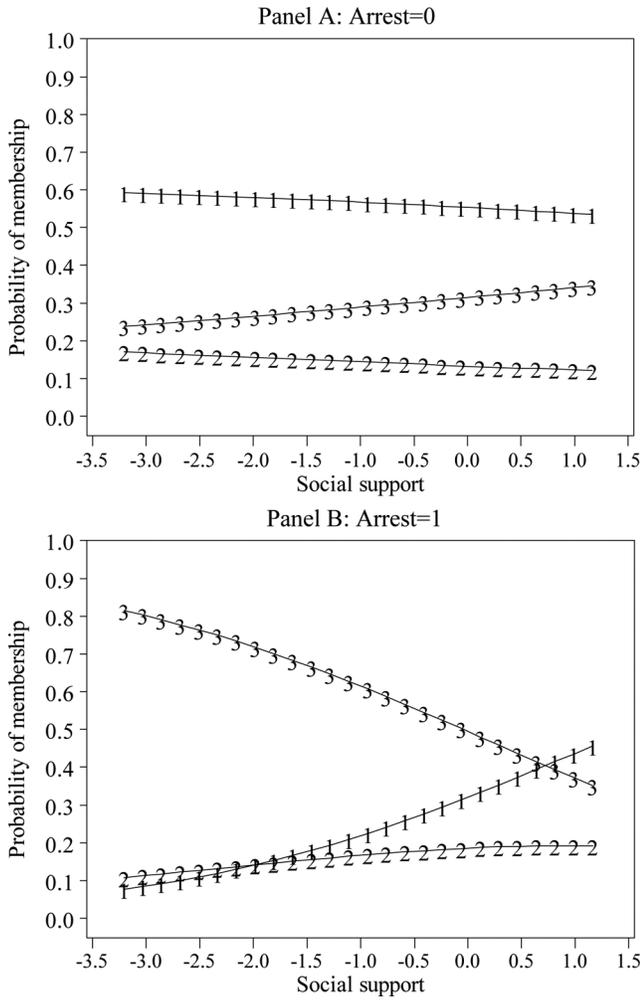


FIGURE 1 Example 1: Predicted probabilities of class membership from PT or MG mixture approaches for alternative values of social support and prior arrest, holding other covariates at their means. *Note.* The number identifying each line corresponds to the latent trajectory class for which the probability of membership is plotted. Class 1 = low-stable aggression; Class 2 = high-stable aggression; Class 3 = declining aggression. The x -axis spans the observed range of standardized social support. PT = product term; MG = multiple group.

aggression are about equally common at low social support, whereas as social support increases, declining aggression becomes relatively more common than high-stable.

EMPIRICAL EXAMPLE 2

A previous study by Muthén (2002) (see also Muthén & Muthén, 2000) extracted $K=4$ classes of antisocial behavior symptoms from a cross-sectional sample of $N=7,326$ 16–23 year olds from the National Longitudinal Survey of Youth (NLSY; The Ohio State University Center for Human Resource Research, 1979). The nine

binary antisocial behavior symptoms were: property damage, fighting, shoplifting, stealing <\$50, making threats, smoking marijuana, other drug involvement, conning someone, and holding stolen goods. In this prior study, the latent classes of antisocial behavior were predicted by three covariates: age in years (a_i , centered at 16 years), gender (m_i , where 1 = male and 0 = female), and race (r_i , where 1 = Black and 0 = white or Hispanic). Herein, we revisit Muthén's (2002) analysis in order to investigate a three-way interaction of age \times gender \times race in predicting class membership, using both PT and MG mixture specifications. This reanalysis is for purposes of methodological illustration.

The *within-class model* for Example 2 uses a latent class analysis (LCA) specification for binary outcomes (e.g., Clogg, 1995; Muthén, 2004). For item j , we have:

$$\pi_j^{(k)} = \frac{1}{1 + \exp(\tau_j^{(k)})}, \quad (12)$$

where $\pi_j^{(k)}$ is the probability of endorsing outcome j , (where $j=1 \dots J$) and $\tau_j^{(k)}$ is an estimated threshold. Herein, $J=9$. Using the PT mixture approach, the Example 2 *between-class model* is then:

$$\begin{aligned} p(c_i = k | \mathbf{x}_i) &= \frac{\exp(\beta_0^{(k)} + \beta_1^{(k)} a_i + \beta_2^{(k)} m_i + \beta_3^{(k)} r_i + \beta_4^{(k)} a_i m_i \\ &\quad + \beta_5^{(k)} m_i r_i + \beta_6^{(k)} a_i r_i + \beta_7^{(k)} a_i m_i r_i)}{\sum_{k=1}^K \exp(\beta_0^{(k)} + \beta_1^{(k)} a_i + \beta_2^{(k)} m_i + \beta_3^{(k)} r_i + \beta_4^{(k)} a_i m_i \\ &\quad + \beta_5^{(k)} m_i r_i + \beta_6^{(k)} a_i r_i + \beta_7^{(k)} a_i m_i r_i)}. \end{aligned} \quad (13)$$

Notice that we now have a $a_i m_i r_i$ product term included in order to accommodate the posited three-way interaction. Two of these variables can be considered moderators and one, say a_i , a focal predictor.

Two alternative MG mixture specifications for the Example 2 *between-class model* could be used that are both equivalent to the above PT mixture specification, when assuming across-group invariance of the within-class model. One MG mixture specification could involve constructing a manifest grouping variable with four levels to represent each combination of the two categorical moderators (m_i and r_i). Allowing effects of a_i to differ across these groups would accommodate the posited three-way interaction. Alternatively, an equivalent MG mixture specification could employ one categorical moderator (say, m_i) as the manifest

TABLE 5
Within-Class Symptom Endorsement Probabilities for Example 2

	Class 1 (13%)	Class 2 (29%)	Class 3 (21%)	Class 4 (37%)
Property damage	.77	.20	.14	.01
Fight	.72	.52	.15	.02
Shoplift	.79	.22	.39	.05
Steal <\$50	.66	.13	.25	.04
Threaten	.80	.56	.37	.06
Use marijuana	.85	.29	.92	.19
Use of other drugs	.49	.01	.52	.00
Con	.60	.29	.21	.06
Hold stolen goods	.58	.07	.09	.00

grouping variable, and incorporate two-way product terms ($a_i r_i$); a test of the equivalence of their effects across group (m_i) tests the posited three-way interaction. The latter specification is implemented here as the Example 2 MG-mixture between-class model. The grouping variable m_i can take on values $g = 0 \dots G$. Here, $G = 1$:

$$p(c_i = k | \mathbf{x}_i, m_i = g) = \frac{\exp(\beta_0^{(k|g)} + \beta_1^{(k|g)} a_i + \beta_3^{(k|g)} r_i + \beta_5^{(k|g)} a_i r_i)}{\sum_{k=1}^K \exp(\beta_0^{(k|g)} + \beta_1^{(k|g)} a_i + \beta_3^{(k|g)} r_i + \beta_5^{(k|g)} a_i r_i)} \quad (14)$$

Note that this specification is actually a hybrid approach because it involves both product terms and a multiple-group architecture, but here for simplicity we continue to call it a MG mixture.

In the LCA, binary item j in class k is again Bernoulli distributed, with $\pi_j^{(k)}$ defined as in Equation (12). The product of the Bernoulli PMFs for items $j = 1 \dots J$ in class k forms a joint within-class PMF, $p(\mathbf{y}_i | c_i = k)$, similarly to the LCGM. This joint within-class PMF and Equation (13) are both substituted into Equation (10)

to construct the PMF for the PT mixture. This joint within-class PMF and Equation (14) are substituted into Equation (11) (now using gender, m_i , rather than arrest, for the grouping variable) to construct the PMF for the MG mixture.

Results

The within-class antisocial item endorsement probabilities for each of the $K = 4$ classes are given in Table 5. Behaviors in class 1 (here, 13% of participants marginally) can be described as multidomain delinquent, class 2 (here, 29% marginally) as fighting, class 3 (here, 21% marginally) as drug use, and class 4 (here, 37% marginally) as low antisocial behavior. We focus here on the prediction of class membership using OR and predicted probability interpretations.

Odds and OR Interpretation

Baseline odds and OR results for distinguishing each class from the reference class (#4) are given in Tables 6 and 7 for the PT mixture and MG mixture, respectively. Effects superscripted by the same letter across Tables 6 and 7 are directly estimated in both models; effects that are directly estimated in only one model can be tested in the other model by extending procedures outlined in Table 1. Our specific focus in Example 2 is investigating the posited three-way interaction effect. Table 6 indicates that this three-way interaction is significant in discriminating membership in class 2 (fighting) versus 4 (low). Table 7 breaks this interaction into two two-way interactions, for male = 1 and male = 0 respectively, each of which can be probed using Equation (6). This probing indicates that a 1 year increase in age decreases the odds of membership in the fight versus low class by 38% for white women ($p < .05$), 21% for black men ($p < .05$), 19% for black women ($p < .05$), and 25% for white men ($p < .05$).

TABLE 6
Product Term (PT) Mixture Results for Example 2: Odds Ratios for Each Predictor Distinguishing Membership in Class k Versus K

Multinomial coefficient	Class 1 vs. 4			Class 2 vs. 4			Class 3 vs. 4		
	Est. (SE)	Exp(Est)	CI	Est. (SE)	Exp(Est)	CI	Est. (SE)	Exp(Est)	CI
Intercept ^a	-1.66* (.20)	.19	(.13, .28)	-0.34* (.15)	.71	(.53, .96)	-0.42* (.11)	.66	(.53, .82)
Age ^b	-0.40* (.08)	.67	(.57, .78)	-0.48* (.07)	.62	(.54, .71)	-0.02 (.03)	.98	(.92, 1.04)
Male	2.34* (.23)	10.38	(6.61, 16.29)	1.45* (.19)	4.26	(2.94, 6.19)	-0.64* (.21)	.53	(.35, .80)
Black ^c	-0.31 (.37)	.73	(.36, 1.51)	0.85* (.20)	2.34	(1.58, 3.46)	-1.53* (.30)	.22	(.12, .39)
Age × male	0.14 (.08)	1.15	(.98, 1.35)	0.19* (.08)	1.21	(1.03, 1.41)	0.21* (.05)	1.23	(1.12, 1.36)
Black × male	-0.25 (.43)	.78	(.34, 1.81)	-0.61* (.30)	.54	(.30, .98)	0.71 (.52)	2.03	(.73, 5.64)
Age × black ^d	0.26* (.12)	1.30	(1.03, 1.64)	0.27* (.08)	1.31	(1.12, 1.53)	0.20* (.07)	1.22	(1.06, 1.40)
Age × black × male	-0.14 (.14)	.87	(.66, 1.14)	-0.22* (.10)	.80	(.66, .98)	-0.11 (.11)	.90	(.72, 1.11)

Note. * = $p < .05$. Superscripted letters indicate that an effect is equal across the modeling approaches in Tables 6 versus 7. For the intercept, Exp(Est) is a baseline odds. For each slope, Exp(Est) an odds ratio. CI = confidence interval, Exp(Est ± 1.96 × SE).

TABLE 7
Multiple-group (MG) Mixture Results for Example 2: Odds Ratios for Each Predictor Distinguishing Membership in Class k Versus K

Multinomial coefficient	Class 1 vs. 4			Class 2 vs. 4			Class 3 vs. 4		
	Est. (SE)	Exp(Est)	CI	Est. (SE)	Exp(Est)	CI	Est. (SE)	Exp(Est)	CI
Intercept male	0.68* (.13)	1.97	(1.53, 2.55)	1.11* (.14)	3.03	(2.31, 3.99)	-1.06* (.20)	.35	(.23, .51)
Intercept female ^a	-1.66* (.20)	.19	(.13, .28)	-0.34* (.15)	.71	(.53, .96)	-0.42* (.11)	.66	(.53, .82)
Age male	-0.27* (.03)	.76	(.72, .81)	-0.29* (.04)	.75	(.69, .81)	0.19* (.04)	1.21	(1.12, 1.31)
Age female ^b	-0.40* (.08)	.67	(.57, .78)	-0.48* (.07)	.62	(.54, .71)	-0.02 (.03)	.98	(.92, 1.04)
Black male	-0.56* (.23)	.57	(.36, .90)	0.24 (.23)	1.27	(.81, 2.00)	-0.82 (.43)	.44	(.19, 1.02)
black female ^c	-0.31 (.37)	.73	(.36, 1.51)	0.85* (.20)	2.34	(1.58, 3.46)	-1.53* (.30)	.22	(.12, .39)
Age × black male	0.12 (.06)	1.13	(1.00, 1.27)	0.06 (.06)	1.06	(.94, 1.19)	0.10 (.09)	1.11	(.93, 1.32)
Age × black female ^d	0.26* (.12)	1.30	(1.03, 1.64)	0.27* (.08)	1.31	(1.12, 1.53)	0.20* (.07)	1.22	(1.06, 1.40)

Note. In the multiple-group (MG) mixture, gender (male = 1, female = 0) is the grouping variable. * = $p < .05$. Superscripted letters indicate that an effect is equal across the modeling approaches in Tables 6 versus 7. For each intercept, Exp(Est) is a baseline odds. For each slope, Exp(Est) an odds ratio. CI = confidence interval, Exp(Est) $\pm 1.96 \times SE$.

Predicted Probability Interpretation

Substituting estimated multinomial coefficients from Table 6 (or Table 7) into Equation (13) (or (14)) allows computation of predicted probabilities at chosen combinations of covariates. Figure 2 plots the probability of class membership (vertical axis) across age (horizontal axis) for each combination of race and gender in this cross-sectional analysis.¹⁰ Figure 2 indicates that the fight class (#2) stops becoming the most likely class for white males around age 20, black females around age 18.5, black males by age 22, and for white females before 16. The probability of membership in the drug class (#3) is predicted to increase across the transition to adulthood. For males, the drug class eventually becomes more common than the low class (#4). For females, the low class (#4) is most common starting at age 18.

SUMMARY OF EMPIRICAL EXAMPLES

In sum, in Example 1, we used $K = 3$ trajectory classes to account for individual differences in change in aggression among youth transferring out of foster care, and found that risk (prior arrest record) and protective (social support) variables interactively predicted which youth were more likely to follow which trajectory type, in line with Yoshikawa (1994) and Youngstrom et al. (2003). An additive conditional mixture model would not have captured these effects. Furthermore, the MG mixture gave us ready access to simple slopes that aided in interpreting this interaction. In Example 2, we used $K = 4$ classes to account for associations among

antisocial symptoms in a cross-sectional analysis. Although prior studies had allowed only additive effects of age, gender, and race predictors on delinquent behavior class membership (Muthén, 2002; Muthén & Muthén, 2000), evidence of a three-way interaction of these predictors was found, which was explicated by plotting predicted probabilities of class membership. As these three interacting demographic predictors are themselves proxies for other explanatory variables, future research should instead examine interactive effects among potential underlying causal factors.

DISCUSSION

Mixture applications in psychology often include covariates to explain class membership and aid in construct validation of the categorical latent classification variable (e.g., Castelao & Kroner-Herwig, 2013; Cook et al., 2012; Dyer et al., 2012; Greenbaum et al., 2005; Hepworth et al., 2010; Lincoln & Takeuchi, 2010; Nagin, 2005). However, applications largely have used between-class models involving only main effects of predictors to help distinguish class membership. This is inconsistent with many developmental psychology and developmental psychopathology theories that posit interactions among person characteristics and/or contextual variables in predicting membership in classes of growth trajectories or classes of behavioral symptom profiles (e.g., Bergman, 2001; Lenroot & Giedd, 2011; Magnusson, 1985; Moffitt et al., 2005; Nugent et al., 2011; Sterba & Bauer, 2010). As noted by Cohen, Cohen, West, and Aiken (2003) “the testing of interactions is at the very heart of theory testing in the social sciences” (p. 255). To address this disconnect between substantive theory and methodological practice, this article provided a general description of model specifications incorporating interactive effects of predictors on class membership—using both conditional product term (PT) and multiple group (MG) mixture modeling

¹⁰If the reader is interested in comparing the appearance of predicted probability plots assuming additive effects of predictors on class (in Muthén, 2002) vs. allowing interactive effects of predictors on class (here, in Figure 2), note that in the corresponding additive-effect plot in Muthén (2002), class 2 and 3 labels were switched.

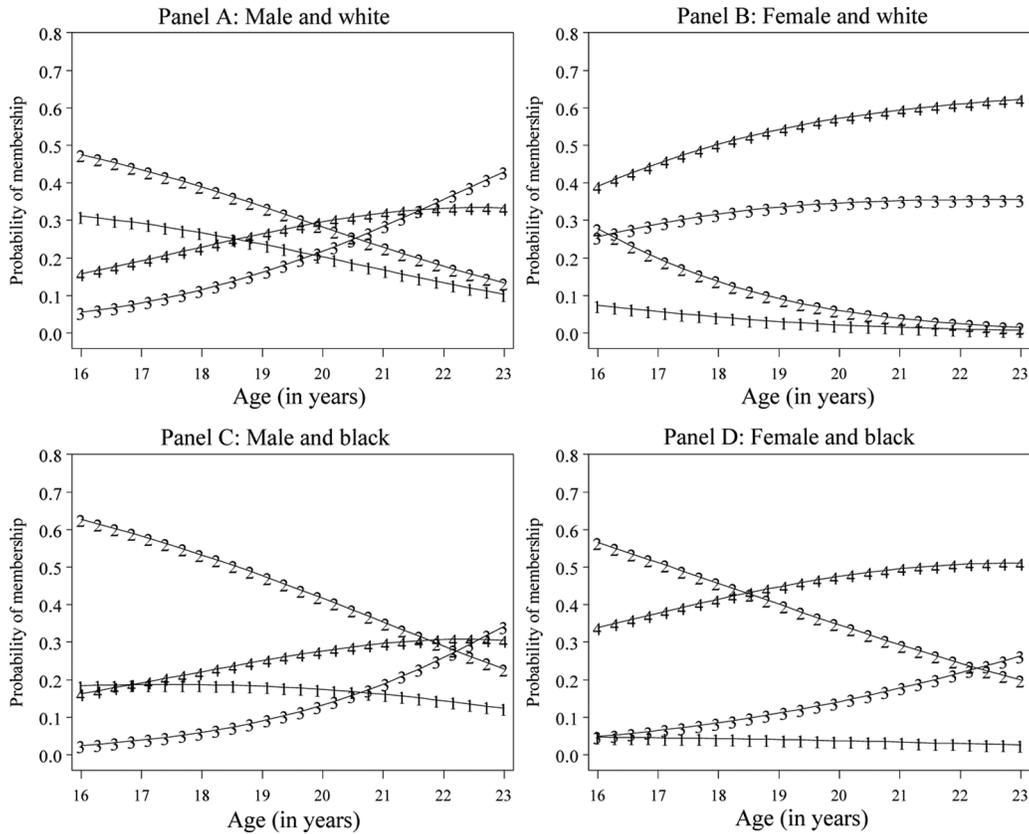


FIGURE 2 Example 2: Predicted probabilities of class membership from PT or MG mixture approaches for alternative values of age, gender and race. *Note.* The number identifying each line corresponds to the class for which the probability of membership is plotted. Class 1 = multidomain delinquent; Class 2 = fighting; Class 3 = drug use; Class 4 = low. In the corresponding additive-effect plot in Muthén (2002), class 2 and 3 labels were reversed. The race = 0 category is herein termed white to be consistent with the prior study, although it includes Hispanic. The x -axis spans the observed range of age; age has a mean = 19 and $SD = 2.18$.

frameworks and using both odds ratio and predictive probability interpretation strategies. Relative advantages and disadvantages of OR and predicted probability interpretation strategies were described. Furthermore, this article showed how and when equivalent inferences could be obtained from PT vs. MG mixture modeling frameworks. Importantly, points made in this article apply to mixture models in general—regardless of the within-class outcome distributions and regardless of within-class model. Here, empirical examples concerned two illustrative mixtures—one longitudinal and the other cross-sectional—and software syntax for each example analysis was provided in our Online Appendix.

Extensions

We conclude with a discussion of three extension topics. First, although here we have discussed incorporating only two-way and three-way linear interactions among predictors of class membership, nonlinear interactions could also be of interest (e.g., x_{1i}^2, x_{2i}) and could be similarly incorporated. In particular, when considering

higher-order and nonlinear interaction effects, researchers are encouraged to center continuous predictors to reduce nonessential collinearity among predictors within-class. Multicollinearity could increase the risk of estimation problems for mixtures particularly with small classes. Second, researchers interested in recovering interaction effects involving predictors of class membership could compare the MG mixture approach advocated herein versus the variation recommended by Collins and Lanza (2010) and discussed earlier. The latter strategy involves potentially unnecessarily relaxing all across-group constraints on covariate effects, even for covariates not involved in testing the hypothesized interaction(s). This strategy should reduce power (and risk estimation problems due to the increased model complexity) when the constrained model is true. If the unconstrained model is true (i.e., unhypothesized interactions do exist), this strategy may have lower risk of bias. Third, here we discussed plotting of predicted probabilities of class membership using information in the between-class model. It is also possible to combine information from the between-class model and within-class models (in

Equations (10) and (11)) to plot predicted probabilities for y -outcomes (rather than for class) at chosen covariate values (see Bauer & Shanahan, 2007; Nagin & Tremblay, 2005; Sterba & Bauer, 2014). By combining information from within-class and between-class models, plots of such predicted outcomes can depict interactions of between-class predictors with within-class predictors that are accommodated by the mixture.

CONCLUSIONS

Our focus here was on interactions among between-class predictors. It is important for developmental researchers to be aware of options for how to incorporate and interpret such interactions, when they are anticipated based on theory. Otherwise, researchers risk overly parsimonious explanations of class membership that may not fully reflect substantive theory.

REFERENCES

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Aldridge, A., & Roesch, S. C. (2008). Developing coping typologies of minority adolescents: A latent profile analysis. *Journal of Adolescence and Youth, 31*, 499–517. doi:10.1016/j.adolescence.2007.08.005
- Asparouhov, T., & Muthén, B. (2013). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal, 21*(3), 329–341. doi:10.1080/10705511.2014.915181
- Bandeau-roche, K., Miglioretti, D. L., Zeger, S. L., & Rathouz, P. J. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association, 92*, 1375–1386. doi:10.1080/01621459.1997.10473658
- Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods, 9*, 3–29. doi:10.1037/1082-989X.9.1.3
- Bauer, D. J., & Shanahan, M. J. (2007). Modeling complex interactions: Person-centered and variable centered approaches. In T. D. Little, J. A. Bovaird, & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 255–284). Mahwah, NJ: Lawrence Erlbaum.
- Bergman, L. R. (2001). A person approach in research on adolescence: Some methodological challenges. *Journal of Research Adolescence, 16*, 28–53. doi:10.1177/0743558401161004
- Cairns, R. B., Bergman, L. R., & Kagan, J. (1998). *Methods and models for studying the individual*. London, UK: Sage.
- Castelao, C. F., & Kroner-Herwig, B. (2013). Different trajectories of depressive symptoms in children and adolescents: Predictors and differences in girls and boys. *Journal of Youth and Adolescence, 42*, 1169–1182. doi:10.1007/s10964-012-9858-4
- Chassin, L., Curran, P. J., Wirth, R. J., Presson, C. C., & Sherman, S. J. (2009). *Developmental trajectories of cigarette smoking from adolescence to adulthood: An overview and relations to indicators of tobacco dependence* (pp. 189–244). Tobacco Control Monograph No. 20. Bethesda, MD: U.S. Dept. of Health and Human Services.
- Chung, H., Flaherty, B. P., & Schafer, J. L. (2006). Latent class logistic regression: Application to marijuana use and attitudes among high school seniors. *Journal of the Royal Statistics Society: Series A (Statistics in Society), 169*, 723–743. doi:10.1111/j.1467-985X.2006.00419.x
- Clark, S., & Muthén, B. O. (2009). *Relating latent class analysis results to variables not included in the analysis*. Unpublished manuscript. Retrieved from www.statmodel.com/download/relatinglca.pdf
- Clogg, C. C. (1995). Latent class models. In G. Arminger, C. Clogg, & M. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 311–359). New York, NJ: Plenum.
- Clogg, C. C., & Goodman, L. A. (1985). Simultaneous latent structure analysis in several groups. In N. B. Tuma (Ed.), *Sociological methodology*. San Francisco, CA: Jossey-Bass.
- Cohen, J. (1968). Multiple regression as a general data-analytic system. *Psychological Bulletin, 70*, 426–443. doi:10.1037/h0026714
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Hillsdale, NJ: Erlbaum.
- Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis with applications in the social, behavioral, and health sciences*. Hoboken, NJ: Wiley.
- Cook, G. A., Roggman, L. A., & D'zatko, K. (2012). A person-oriented approach to understanding dimensions of parenting in low-income mothers. *Early Childhood Research Quarterly, 27*, 582–595. doi:10.1016/j.ecresq.2012.06.001
- Courtney, M., & Cusick, G. (2007). *Crime during the transition to adulthood: How youth fare as they leave out-of-home care in Illinois, Iowa and Wisconsin*. [Computer file]. <http://doi.org/10.3886/ICPSR27062.v1>
- Dayton, C. M., & Macready, G. B. (1988). Concomitant-variable latent-class models. *Journal of the American Statistical Association, 83*, 173–178. doi:10.1080/01621459.1988.10478584
- Duncan, L. E., & Keller, M. C. (2011). A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *The American Journal of Psychiatry, 168*, 1041–1049. doi:10.1176/appi.ajp.2011.11020191
- Dyer, W. J., Pleck, J., & McBride, B. (2012). Using mixture regression to identify varying effects: A demonstration with paternal incarceration. *Journal of Marriage and Family, 74*, 1129–1148. doi:10.1111/j.1741-3737.2012.01012.x
- Fontaine, N., Carbonneau, R., Vitaro, F., Barker, E. D., & Tremblay, R. E. (2009). Research review: A critical review of studies on the developmental trajectories of antisocial behavior in females. *Journal of Child Psychology and Psychiatry, 50*, 363–385. doi:10.1111/j.1469-7610.2008.01949.x
- Geiser, C., Lehmann, W., & Eid, M. (2006). Separating “rotators” from “nonrotators” in the mental rotations test: A multigroup latent class analysis. *Multivariate Behavioral Research, 41*, 261–293. doi:10.1207/s15327906mbr4103_2
- Greenbaum, P. E., Del Boca, F. K., Darkes, J., Wang, C.-P., & Goldman, M. S. (2005). Variation in the drinking trajectories of freshman college students. *Journal of Consulting and Clinical Psychology, 73*, 229–238. doi:10.1037/0022-006X.73.2.229
- Hepworth, S. J., Law, G. R., Lawlor, D. A., & McKinney, P. A. (2010). Early life patterns of common infection: A latent class analysis. *European Journal of Epidemiology, 25*, 875–883. doi:10.1007/s10654-010-9518-5
- Hosmer, D., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York, NY: Wiley.
- Lenroot, R. K., & Giedd, J. N. (2011). Annual research review: Developmental considerations of gene by environment interactions. *Journal of Child Psychology and Psychiatry, 52*, 429–441. doi:10.1111/j.1469-7610.2011.02381.x
- Lincoln, K. D., & Takeuchi, D. T. (2010). Variation in the trajectories of depressive symptoms: Results from the Americans' changing lives study. *Biodemography and Social Biology, 56*, 24–41. doi:10.1080/19485561003709180

- Liu, J., Bann, C., Lester, B., Tronick, E., Das, A., Lagasse, L., ... Bada, H. (2010). Neonatal neurobehavior predicts medical and behavioral outcomes. *Pediatrics*, *125*, e90–e98. doi:10.1542/peds.2009-0204
- Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage.
- Long, J. S. (2009). *Group comparisons in logit and probit using predicted probabilities*. Unpublished manuscript. Retrieved from http://www.indiana.edu/~jslsoc/files_research/groupdif
- Long, J. S. (2012). Regression models for nominal and ordinal outcomes. In H. Best & C. Wolf (Eds.), *The SAGE handbook of regression analysis and causal inference*. New York, NY: Sage.
- Lubke, G., & Muthén, B. O. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*, 26–47. doi:10.1207/s15328007sem1401_2
- Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, *10*, 21–39. doi:10.1037/1082-989X.10.1.21
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomizing quantitative variables. *Psychological Methods*, *7*, 19–40. doi:10.1037/1082-989X.7.1.19
- McCutcheon, A. (2002). Basic concepts and procedures in single- and multiple-group latent class analysis. In J. A. Hagenaars & A. L. McCutcheon (Eds.), *Applied latent class analysis* (pp. 56–88). New York, NY: Cambridge.
- Magnusson, D. (1985). Implications of an interactional paradigm for research on human development. *International Journal of Behavioral Development*, *8*, 115–137. doi:10.1177/016502548500800201
- Mathiowetz, R. (2010). *Evaluating latent variable interactions with structural equation mixture models*. (Unpublished thesis). University of North Carolina at Chapel Hill.
- McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York, NY: Wiley.
- Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York, NY: Routledge.
- Moffitt, T. E., Caspi, A., & Rutter, M. (2005). Strategy for investigating interactions between measured genes and measured environments. *Archives of General Psychiatry*, *62*, 473–481. doi:10.1001/archpsyc.62.5.473
- Muthén, B. (2001). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 1–33). Mahwah, NJ: Erlbaum.
- Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran (2003). *Psychological Methods*, *8*, 369–377. doi:10.1037/1082-989X.8.3.369
- Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, *24*, 882–891. doi:10.1111/j.1530-0277.2000.tb02070.x
- Muthén, B. O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, *29*, 81–117. doi:10.2333/bhmk.29.81
- Muthén, B. O. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Newbury Park, CA: Sage.
- Muthén, B. O., & Asparouhov, T. (2002). *Modeling of heteroscedastic measurement errors*. Mplus Web Note #3. Retrieved from <https://www.statmodel.com/download/webnotes/mc3.pdf>
- Muthén, L. K., & Muthén, B. O. (1998–2014). *Mplus users guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Nagin, D. S. (1999). Analyzing developmental trajectories: Semi-parametric, group-based approach. *Psychological Methods*, *4*, 39–157. doi:10.1037/1082-989X.4.2.139
- Nagin, D. S. (2005). *Group-based modeling of development*. Cambridge, MA: Harvard Press.
- Nagin, D. S., & Tremblay, R. E. (2005). Developmental trajectory groups: Fact or a useful statistical fiction? *Criminology*, *43*, 873–904. doi:10.1111/j.1745-9125.2005.00026.x
- Nugent, N. R., Tyrka, A. R., Carpenter, L. L., & Price, L. H. (2011). Gene–environment interactions: Early life stress and risk for depressive and anxiety disorders. *Psychopharmacology*, *214*, 175–196. doi:10.1007/s00213-010-2151-x
- Oberski, D. L. (2014). *A flexible method to explain differences in structural equation model parameters over subgroups*. Retrieved from <http://daob.nl/wp-publications/oberskiwpssem-ipc/>
- The Ohio State University Center for Human Resource Research. (1979). *National longitudinal survey of youth*. Retrieved from <http://doi.org/10.3886/ICPSR04683.v1>
- Shentu, Y., & Xie, M. (2010). A note on dichotomization of continuous response variable in the presence of contamination and model misspecification. *Statistics in Medicine*, *29*, 2200–2214. doi:10.1002/sim.3966
- Sherbourne, C. D., & Stewart, A. L. (1991). The MOS social support survey. *Social Science & Medicine*, *32*(6), 705–714. doi:10.1016/0277-9536(91)90150-b
- Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London, UK: Sage.
- Sterba, S. K. (2013). Understanding linkages among mixture models. *Multivariate Behavioral Research*, *48*, 775–815. doi:10.1080/00273171.2013.827564
- Sterba, S. K. (2014). Modeling strategies in developmental psychopathology research: Prediction of individual change. In M. Lewis & K. D. Rudolph (Eds.), *Handbook of developmental psychopathology* (3rd ed., pp. 109–124). New York, NY: Springer.
- Sterba, S. K., & Bauer, D. J. (2010). Matching method with theory in person-oriented developmental psychopathology research. *Development and Psychopathology*, *22*, 239–254. doi:10.1017/S0954579410000015
- Sterba, S. K., & Bauer, D. J. (2014). Predictions of individual change recovered with latent class or random coefficient growth models. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*, 342–360. doi:10.1080/10705511.2014.915189
- Van Horn, M. L., Jaki, T., Masyn, K., Ramey, S. L., Smith, J. A., & Antaramian, S. (2009). Assessing differential effects: Applying regression mixture models to identify variations in the influence of family resources on academic achievement. *Developmental Psychology*, *45*, 1298–1313. doi:10.1037/a0016427
- Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis*, *18*, 450–469. doi:10.1093/pan/mpq025
- Wedel, M. (2002). Concomitant variables in finite mixture models. *Statistical Neerlandica*, *56*, 362–375. doi:10.1111/1467-9574.t01-1-00072
- Yoshikawa, H. (1994). Prevention as cumulative protection: Effects of early family support and education on chronic delinquency. *Psychological Bulletin*, *115*, 28–54. doi:10.1037//0033-2909.115.1.28
- Youngstrom, E., Weist, M. D., & Albus, K. E. (2003). Exploring violence exposure, stress, protective factors and behavioral problems among inner-city youth. *American Journal of Community Psychology*, *32*, 115–129. doi:10.1023/A:1025607226122