

Citation:

Lanza, S. T., Bray, B. C., & Collins, L. M. (2013). An introduction to latent class and latent transition analysis. In J. A. Schinka, W. F. Velicer, & I. B. Weiner(Eds.), *Handbook of psychology* (2nd ed., Vol. 2, pp. 691-716). Hoboken, NJ: Wiley.

**An Introduction to Latent Class and Latent Transition Analysis**

Stephanie T. Lanza, Ph.D.<sup>12</sup>

Bethany C. Bray, Ph.D.<sup>3</sup>

Linda M. Collins, Ph.D.<sup>14</sup>

<sup>1</sup>The Methodology Center, The Pennsylvania State University

<sup>2</sup>College of Health and Human Development, The Pennsylvania State University

<sup>3</sup>Department of Psychology, Virginia Polytechnic Institute and State University

<sup>4</sup>Department of Human Development and Family Studies, The Pennsylvania State University

Stephanie T. Lanza, Ph.D.

Scientific Director, The Methodology Center

Research Associate Professor, Health and Human Development

The Pennsylvania State University

204 E. Calder Way, Suite 400

State College, PA 16801

phone: 814-865-7095

fax: 814-863-0000

### Acknowledgments

This study was supported by Award Number P50-DA010075 from the National Institute on Drug Abuse. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Drug Abuse or the National Institutes of Health. The authors thank Brian P. Flaherty for important contributions to an earlier published version of this chapter. The authors also thank Amanda Applegate for significant contributions during the editorial process. This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (<http://www.cpc.unc.edu/addhealth>). No direct support was received from grant P01-HD31921 for this analysis.

## Table of Contents

INTRODUCTION .....	4
LATENT CLASS ANALYSIS.....	5
The Concept of a Latent Class .....	5
The LCA Mathematical Model.....	11
Multiple-Groups LCA and LCA with Covariates.....	14
Estimation .....	16
Standard Errors .....	17
Missing Data .....	18
The Use of Parameter Restrictions .....	20
Model Selection and Goodness-of-Fit .....	22
Absolute model fit.....	22
Relative model fit.....	23
A LATENT CLASS EXAMPLE: ADOLESCENT DEPRESSION .....	27
The Five-Class Solution.....	29
Multiple-Groups LCA: Gender Differences in Depression Classes.....	30
LCA with Covariates: Predicting Depression Latent Class Membership.....	32
LONGITUDINAL EXTENSIONS: REPEATED-MEASURES LCA AND LTA .....	33
Multiple-Groups LTA and LTA with Covariates .....	41
SOME RECENT EXTENSIONS TO LCA.....	43
Associative LTA .....	44
Group-Based Trajectory Analysis.....	45
Other Variations of Mixture Models.....	46
Multilevel LCA.....	46
Additional Types of Indicators .....	48
AREAS FOR FUTURE RESEARCH RELATED TO LCA.....	49
CONCLUSIONS.....	51
REFERENCES .....	53

## AN INTRODUCTION TO LATENT CLASS AND LATENT TRANSITION ANALYSIS

**INTRODUCTION**

Often quantities of interest in psychology cannot be observed directly. These unobservable quantities are known as latent variables. In addition to being unobservable, latent variables tend to be complex, often multidimensional, constructs. Unlike height, which can be measured with a single assessment, depression or temperament cannot be adequately measured with only one observed variable, such as a single questionnaire item. This complexity can be handled using multiple observed variables as indicators of the latent variable; this approach provides a more complete picture of the construct and allows estimation of measurement error. Examples of latent variables in the psychological literature include temperament (Stern, Arcus, Kagan, Rubin, & Snidman, 1995), cognitive ability (Humphreys & Janson, 2000), health behaviors (Maldonado-Molina & Lanza, 2010), and motivation (Coffman, Patrick, Palen, Rhoades, & Ventura, 2007). Knowing that they are imperfect measures, using data from available observed variables provides the best measures of latent variables. When several observed variables are used to assess an underlying latent variable, we have a basis for removing measurement error, leading to better measurement of the latent variable.

The fundamental premise of any latent variable model is that the covariation among observed variables is explained by the latent variable. There are four latent variable frameworks that model the relationship between observed variables and a latent variable. Figure 1 depicts this relationship for the four frameworks. Which framework is appropriate depends on whether the observed variables and latent variable are considered to be continuous or categorical. In *factor analysis* or *covariance structure analysis*, observed variables, usually continuous, map onto

continuous latent variables assumed to be normally distributed (Jöreskog & Sörbom, 1996).

*Latent trait analysis* (Spiel, 1994) refers to discrete observed variables mapping onto a continuous latent variable. For example, a set of observed variables measuring aptitude, coded as correct or incorrect, might be seen as indicators of the underlying latent trait, in this case ability.

In *latent profile analysis*, continuous observed variables map onto a discrete latent variable.

*Latent class analysis (LCA)* models the relationship between discrete observed variables and a discrete latent variable.

---

Figure 1 about here

---

The present chapter will focus on this last framework, LCA. The first section introduces the concept of a latent class and then presents the mathematical model. This is followed by a discussion of parameter restrictions, model fit, and the measurement quality of categorical variables. The second section demonstrates LCA through an examination of the prevalence of depression types in adolescents. The third section presents longitudinal extensions of LCA and contains an empirical example on adolescent depression types, where we extend the previous analysis to examine change in depression types over time. In the final sections we mention several other extensions to the latent class model and areas that merit additional research in the future.

## LATENT CLASS ANALYSIS

### The Concept of a Latent Class

Latent classes can be thought of as a system to classify groups of individuals according to some construct that is not directly measurable. Suppose, for example, a researcher interested in

the construct temperament hypothesizes that this construct is made up of qualitatively different categories. The researcher can measure several indicators of temperament, and then use LCA to identify two or more temperamental types into which people might be classified. In examples presented throughout this chapter, latent class models have been applied in the study of psychological and behavioral phenomena. Note that for some constructs (such as temperament) an individual's latent class membership is generally expected to remain the same over time, whereas for other constructs (such as substance use) it is possible for individuals to move between latent classes over time.

Theory suggests that there are two main temperamental types of children, namely inhibited and uninhibited, characterized by avoidance or approach to unfamiliar situations (Kagan, 1989). Stern, et al. (1995) used latent class analysis to test this theory empirically, comparing a model with two temperamental types of children to models with three and four types. Infants in two cohorts of sample sizes 93 and 76 were measured on three categorical variables: motor activity, fret/cry, and fear. Two-, three-, and four-class models were fit for each cohort. For both cohorts, a two-class solution appeared to represent the data adequately, although the sample size may not have provided enough power to detect additional classes.

Latent class models have also been used to explore the onset of substance use behaviors during adolescence. Following Kandel's (1975) introduction of the concept of stages in substance use, Collins and colleagues (Collins, Graham, Rousculp, & Hansen, 1997; Hyatt & Collins, 2000) explored this construct as a categorical latent variable. Using data from the Adolescent Alcohol Prevention Trial (Hansen & Graham, 1991), Collins et al. (1997) identified a stage-

sequence of substance use made up of the following eight latent classes: No Use; Alcohol Use; Alcohol Use With Drunkenness; Tobacco Use; Alcohol and Tobacco Use; Alcohol Use With Drunkenness and Advanced Use; Alcohol, Tobacco and Advanced Use; and Alcohol Use With Drunkenness, Tobacco Use, and Advanced Use. This model specifies that adolescents can first move from the No Use latent class to either the Alcohol Use or the Tobacco Use latent class, and then progress to latent classes characterized by more advanced substance use. Notice that not all possible combinations of substances are represented in this latent class model. For example, a latent class characterized by alcohol and advanced use without drunkenness and tobacco use does not exist. The eight latent classes specified by this model were sufficient to represent the data.

In addition to modeling actual substance use behavior, LCA has been used to identify four subgroups of high school seniors according to their patterns of motivations to drink alcohol (Coffman et al., 2007). The first latent class was characterized mainly by a desire to experiment with alcohol (Experimenters, 36%); the second latent class used alcohol mainly to get high and have a good time (Thrill-Seekers, 32%); the third subgroup reported that their primary motivation to use alcohol was to relax (Relaxers; 15%); and the fourth latent class was characterized by a wide range of motivations to drink, including anger/frustration and getting away from problems (Multi-Reasoners, 18%). Membership in the Multi-Reasoners latent class was associated with early initiation of alcohol use, past-year drunkenness, and drinking before 4:00 PM.

LCA has been applied to various educational studies, including Aitkin, Anderson, and Hinde's (1981) examination of the number of different classes of teaching style in the UK. Teaching style was characterized by the presence or absence of 38 different teaching behaviors in 468 4<sup>th</sup>-grade teachers. A latent class approach to modeling this type of data is appealing because it summarizes data from many different questionnaire items in a parsimonious way. Although only two teaching styles were originally predicted, formal and informal, evidence was found for a three-class model. The first class was a more formal style, where the teachers used a firm timeline and restricted the students' behaviors. The data suggested that 48 percent of teachers had a formal style. The second class represented an informal teaching style, where teachers tended to have less strict classroom organization, integrated subjects, and encouraged individual work. This class encompassed 32 percent of teachers. The remaining 20 percent of teachers fell in the third latent class, a mixed teaching style, where certain behavior restrictions were enforced as in the formal group, but grading and homework were similar to the informal group.

LCA has been used to apply the transtheoretical model of behavior change to various types of health behaviors, including smoking cessation (Martin, Velicer, & Fava, 1996), condom use (Evers, Harlow, Redding, & LaForge, 1998), and exercise (Gebhardt, Dusseldorp, & Maes, 1999). Velicer and colleagues used LCA and its longitudinal extension latent transition analysis (LTA) to test competing models of the stages of change in smoking behavior (Velicer, Martin, & Collins, 1996; Martin, et al., 1996). Figure 2 shows the model that best represented the data. This model posits that individuals can move both forwards and backwards through the stages of

change, and that forward movement between any two adjacent times does not extend beyond a maximum of two stages.

---

Figure 2 about here.

---

Lanza, Savage, and Birch (2010) used LCA to classify women according to the various healthy and unhealthy weight-control strategies they employed over time. Fourteen strategies were included in the latent class model, including observed variables such as using appetite suppressants, reducing calories, reducing fats, reducing carbohydrates, and increasing exercise. The four latent classes identified were: No Weight Loss Strategies (i.e., non-dieters; 10.0%); Dietary Guidelines (26.5%), characterized by use of healthy weight control strategies such as increased fruits and vegetables intake, increased exercise, and decreased fat intake; Guidelines+Macronutrient (39.4%), including individuals who have used traditionally healthy approaches as well as a low-carbohydrate diet; and a group of women who reported trying all healthy and unhealthy strategies (Guidelines+Macronutrient+Restrictive, 24.2%). The inclusion of covariates permitted an examination of how body mass index (BMI), weight concerns, desire to be thinner, disinhibited eating, and dietary restraint were related to weight-control strategy latent class. Dietary restraint was found to moderate the effect of disinhibition on membership in weight-loss strategy group, providing insight into how one might tailor an intervention program that aims to prevent or reduce unhealthy dieting behaviors.

Several studies have explored subtypes of depression using LCA (e.g. Parker, Wilhelm, Mitchell, Roy, & Hadzi-Pavlovic, 1999; Sullivan, Kessler, & Kendler, 1998). For example, Sullivan et al. (1998) used 14 DSM-III-R depressive symptoms in an epidemiological data set to

identify six latent classes: Severe Typical, Mild Typical, Severe Atypical, Mild Atypical, Intermediate, and Minimal Symptoms. Although the number and type of latent classes identified depended on both the sample and the observed variables included in the model, LCA provided a means of empirically testing competing theories about depressive subtypes for a given data set.

A more recent application of LCA involved the identification of subgroups of individuals based not on their own traits or behaviors, but rather on their ecological/contextual profiles. This framework provides a way, for example, to organize individuals according to the particular set of risk factors to which they are exposed. Latent classes of risk (i.e., risk profiles) that are associated with poorest outcomes might then be useful targets for intervention programs. As an example, Lanza, Rhoades, Nix, Greenberg, and the CPPRG (2010) identified the following four profiles of 13 risk factors across child, family, school, and neighborhood domains in a diverse sample of children in kindergarten: Two-Parent Low-Risk, Single-Parent/History Of Problems, Single-Parent Multilevel Risk, and Two-Parent Multilevel Risk. Membership in each latent class varied substantially across race and urbanicity, and was highly related to 5<sup>th</sup>-grade externalizing problems, failing grades, and low academic achievement. The use of LCA to identify risk profiles was then extended to examine differential treatment effects across latent classes (Lanza & Rhoades, in press). This approach, which is analogous to including a latent moderator in the evaluation of a program's effect, can be used to suggest subgroups of individuals that hold promise to respond most strongly to the program and to describe other subgroups that may benefit from a modified program.

### **The LCA Mathematical Model**

Latent class theory is a measurement theory for a categorical latent variable that divides a population into mutually exclusive and exhaustive latent classes (Lazarsfeld & Henry, 1968; Goodman, 1974). The latent variable is measured by multiple categorical indicators. During the 1970's, two important papers were published which together provided researchers with the theoretical and computational tools for estimating latent class models. First, Goodman (1974) described a maximum-likelihood estimation procedure for latent class models. Second, a broadly applicable presentation of the use of the expectation-maximization (EM) algorithm, an iterative technique that yields maximum-likelihood estimates from incomplete data, was introduced (Dempster, Laird, & Rubin, 1977). In LCA the latent (unobserved) variables can be considered to be missing data. Several software packages are available for conducting LCA and its extensions, including PROC LCA & PROC LTA (Lanza, Dziak, Huang, Xu, & Collins, 2011), Latent Gold (Vermunt & Magidson, 2005), Mplus (Muthén & Muthén, 1998-2010), and LEM (Vermunt, 1997a,b).

Latent class models are particularly useful when the theoretical construct of interest is made up of qualitatively different groups, but the group membership of individuals is unknown and therefore must be inferred from the data. Although it might be tempting to try to classify individuals based on their manifest data, LCA has several important advantages over simple crosstabulation methods. First, a latent variable approach to identifying qualitatively different classes of individuals involves using multiple observed variables as indicators of the latent variable. This provides a basis for estimating measurement error, yielding a clearer picture of the underlying latent variable. Second, LCA can be a confirmatory procedure. For a set of discrete

observed variables, the user must specify the number of latent classes. LCA then estimates the parameters and provides a fit statistic. A confirmatory procedure provides a means of testing a priori models and comparing the fit of different models. Third, when measurement error is present, many individuals' responses do not point unambiguously to membership in one particular group. A latent variable approach can help the researcher interpret large contingency tables, providing a sense of both the underlying group structure and the amount of measurement error associated with particular observed variables.

In latent class models, the data are used to estimate the number of classes in the population, the relative size of each class, and the probability of a particular response to each observed variable given class membership. Suppose that there are  $j = 1, \dots, J$  observed variables, and that variable  $j$  has response categories  $r_j = 1, \dots, R_j$ ; suppose also that the latent variable has  $c = 1, \dots, C$  latent classes. Let  $\mathbf{y}$  represent a particular response pattern (i.e., a vector of possible responses to the observed variables), and let  $\mathbf{Y}$  represent the array of all possible  $\mathbf{y}$ s. Each response pattern  $\mathbf{y}$  corresponds to a cell of the contingency table formed by crosstabulating all of the observed variables, and the length of the array  $\mathbf{Y}$  is equal to the number of cells in this table. As an example, if there are eight dichotomous observed variables corresponding to responses *yes* or *no* to depression symptoms, an individual reporting no symptoms would have the response pattern  $\{1,1,1,1,1,1,1,1\}$ , and someone experiencing the first four symptoms but not the last four would have the response pattern  $\{2,2,2,2,1,1,1,1\}$ .

The estimated proportion of a particular response pattern  $P(\mathbf{Y} = \mathbf{y})$  can be expressed as a function of two types of parameters. First, the *latent class prevalences*, which will be referred to

as  $\gamma$  parameters, represent the proportion of the population that falls into each latent class. Because the latent classes are mutually exclusive and exhaustive (i.e., each individual is placed into one and only one latent class), the  $\gamma$  parameters sum to 1. Second, the *item-response probabilities*, which we will refer to as  $\rho$  parameters, represent the probability of a particular response to a manifest variable, conditioned on latent class membership. These  $\rho$  parameters express the relationship between the observed variables and the latent variable. The item-response probabilities bear a close conceptual resemblance to factor loadings, in that they provide a basis for interpretation of the meaning of the latent classes. However, it is important to remember that they represent probabilities rather than regression coefficients. A  $\rho$  parameter near 0 or 1 represents a strong relationship between the observed variable and the latent construct. This would mean that, given latent class, we can predict with near certainty how an individual would respond to that observed variable. On the other hand, for dichotomous observed variables, a  $\rho$  parameter near .5 means that the observed variable does not provide any information above random chance in placing the individual in the latent class.

Let us establish an indicator function  $I(y_j = r_j)$  that equals 1 when the response to variable  $j = r_j$ , and equals 0 otherwise. The probability of observing a particular response pattern, or cell in the contingency table cross-classifying the observed variables, can be written as

$$P(\mathbf{Y} = \mathbf{y}) = \sum_{c=1}^C \gamma_c \prod_{j=1}^J \prod_{r_j=1}^{R_j} \rho_{j,r_j|c}^{I(y_j=r_j)}, \quad (1)$$

where  $\gamma_c$  is the probability of membership in latent class  $c$  and  $\rho_{j,r_j|c}^{I(y_j=r_j)}$  is the probability of response  $r_j$  to observed variable  $j$ , conditional on membership in latent class  $c$ . The  $\gamma$  parameters represent a vector of latent class membership probabilities that sum to 1. The  $\rho$  parameters represent a matrix of item-response probabilities conditional on latent class membership.

The latent class model is defined by making two critical assumptions. First, all individuals in a latent class are assumed to have the same item-response probabilities for the observed variables. For example, all individuals in the latent class associated with an inhibited temperament type (Inhibited class), are assumed to have the same probability of displaying high motor activity. Second, there is an assumption of conditional independence given latent class. This implies that within each latent class, the  $J$  indicators are independent of one another. For example, individuals' temperament type explains any relationship among their reports of motor activity, fret/cry and fear. This second assumption allows the probability of a particular response pattern to be expressed as shown in Equation 1, without conditioning on anything in addition to latent class.

### **Multiple-Groups LCA and LCA with Covariates**

Grouping variables can be included in LCA in much the same way that they can be included in structural equation models. In LCA, grouping variables serve two primary purposes. The first is to allow for a statistical test of measurement invariance; that is, a test can be conducted to determine whether the item-response probabilities that define the latent classes differ across two or more groups. Whenever it is reasonable to do so, item-response probabilities should be constrained to be equal across groups so that group comparisons are made using a

common definition of the underlying latent classes. The second purpose of multiple-groups LCA is to compare the latent class prevalences across groups. For example, in the study by Lanza, Rhoades, et al. (2010) described above, the prevalence of the risk profiles varied substantially across race/urbanicity groups. Most notably, the prevalence of the Two-Parent Low Risk latent class was 55% for rural White children, 38% for urban White children, and only 11% for urban African American children. Because the item-response probabilities were constrained to be equal across groups, these comparisons were based on subgroups characterized by the same intersection of risks.

In addition to grouping variables, covariates (also called exogenous or concomitant variables in the literature) can be incorporated into the latent class model in order to predict latent class membership (Collins & Lanza, 2010; Dayton & Macready, 1988; van der Heijden, Dessens, & Bockenholt, 1996) or to predict the item-response probabilities (Pfefferman, Skinner and Humphreys, 1998). Most commonly, covariates are used to predict latent class membership and are added to the latent class model via multinomial logistic regression; covariates can be discrete, continuous or higher-order terms (e.g., powers or interactions). For example, by including alcohol use as a covariate it is possible to investigate questions such as, "For a one-unit increase in alcohol use, how do the probabilities of membership in the depression latent classes change?" Including interaction terms makes it possible to address questions such as whether the effect of alcohol use on membership in latent classes of depression differs for males and females (i.e., whether gender moderates the association between alcohol use and latent class membership).

An extended latent class model including a grouping variable  $V$ , with  $q = 1, \dots, Q$  groups, and a single covariate  $X$  can be described as follows (note that the model can be extended to include two or more covariates). The latent class model can be expressed as

$$P(\mathbf{Y} = \mathbf{y} | \mathbf{V} = \mathbf{q}, \mathbf{X} = \mathbf{x}) = \sum_{c=1}^C \gamma_{c|q}(\mathbf{x}) \prod_{j=1}^J \prod_{r_j=1}^{R_j} \rho_{j,r_j|c,q}^{I(y_j=r_j)},$$

where  $\rho_{j,r_j|c,q}^{I(y_j=r_j)}$  is the probability of response  $r_j$  to observed variable  $j$ , conditional on membership in latent class  $c$  and group  $q$ , and  $\gamma_{c|q}(\mathbf{x})$  is a standard baseline-category multinomial logistic regression model (e.g., Agresti, 2002):

$$\gamma_{c|q}(\mathbf{x}) = P(\mathbf{C} = \mathbf{c} | \mathbf{X} = \mathbf{x}, \mathbf{V} = \mathbf{q}) = \frac{e^{\beta_{0c|q} + \beta_{1c|q}\mathbf{x}}}{\mathbf{1} + \sum_{c'=1}^{C-1} e^{\beta_{0c'|q} + \beta_{1c'|q}\mathbf{x}}}$$

for  $c'=1, \dots, C-1$ , and where  $C$  is the designated reference class. Note that in this model the  $\rho$  parameters are not dependent on the covariate  $X$ .

## Estimation

The EM algorithm is usually used to estimate the parameters of latent class models (Dempster, et al., 1977; Goodman, 1974). This algorithm alternates between the expectation step (E-step), and the maximization step (M-step). At each step of the EM algorithm the current set of estimates is compared with the set from the previous step. When the difference between the estimates becomes smaller than a specified criterion, the program has converged on a maximum of the likelihood function. Depending on the likelihood function of a given model, there may be a distinct global maximum, or there may be one or more local maxima.

The issue of local maxima is critical in latent class models. If there are several local maxima, the set of starting values will determine which local maximum is reached. It is important to explore multiple solutions to ensure that the maximum-likelihood estimates represent the best solution. This can be done by estimating the parameters based on several different sets of random start values. Although the best outcome is one in which a single mode is identified, it is common to find different solutions corresponding to different local maxima. In this case there are several ways to proceed. One approach is to examine the distribution of solutions and assume that the solution reached most often is the best one and thus can be selected as the final model. A second approach is to select the solution with the best fit, which corresponds to the highest likelihood. A third approach is to simplify the model being fit, which will reduce the number of parameters being estimated. Often this is enough to ensure that just one solution is reached. Any combination of these approaches can be used together in deciding upon a final solution. Fortunately, for many latent class analyses only one solution will be identified.

### **Standard Errors**

Standard errors are not a by-product of the EM algorithm, although several methods for obtaining estimates of the standard errors have been proposed. One approach used in most LCA software programs is a likelihood-based one which involves taking the inverse of the information matrix (see Bandeen-Roche, Miglioretti, Zeger, & Rathouz, 1997 for technical details). Several Bayesian approaches also have been proposed, including the addition of a mild smoothing prior so that the information matrix can be inverted (Rubin & Schenker, 1987) and a technique called data augmentation (Tanner & Wong, 1987), where the latent class variable is imputed multiple

times so that variance estimates can be obtained (see Lanza, Collins, Schafer, & Flaherty, 2005 for a demonstration related to LCA).

In addition to yielding standard errors for each parameter, data augmentation allows for more complex hypothesis tests involving multiple parameters. For example, in some applications it may be useful to use data augmentation with LTA to define a stability parameter that is the weighted (by the initial latent class prevalences) sum of all transition probabilities that correspond to membership in the same latent class over time. Group differences in stability can then be examined. Hyatt, Collins, and Schafer (1999) contains an example of this, as well as an explanation of how to calculate the difference in proportions, relative risks, and associated standard errors for combinations of LTA parameters. In this study WinLTA (Collins & Wugalter, 1992) was used to examine differences in the onset of substance use for females with early pubertal timing and females who do not experience early puberty. Early-developing females were less likely to be in the No Substance Use latent class and more likely to be in the Alcohol and Cigarettes and Alcohol, Cigarettes, Drunkenness, and Marijuana latent classes in both 7<sup>th</sup> and 8<sup>th</sup> grades. Also, early-developing females were more likely to begin using substances between 7<sup>th</sup> and 8<sup>th</sup> grades. Although a group difference in the overall stability in substance use over time was found in the anticipated direction, this difference did not reach statistical significance.

### **Missing Data**

As has been reviewed extensively in the missing data literature (e.g. Schafer, 1997; Collins, Schafer, & Kam, 2001), there are three major classifications of missing data. If

missingness on a variable  $Y$  depends on the variable itself (e.g., a participant in a drug use prevention study avoids a measurement session because he is using drugs), this is referred to as *missing not at random* (MNAR). If missingness on  $Y$  does not depend on  $Y$  itself, this is referred to as *missing at random* (MAR). One example of MAR is missingness caused by poor readers failing to finish a drug use questionnaire during an intervention. The special case of MAR where the cause of missingness is completely unrelated to  $Y$  is referred to as *missing completely at random* (MCAR). (For a thorough introduction to modern missing-data procedures, see Schafer and Graham, 2002.)

Most LCA procedures, including the program used in the empirical analyses reported below (PROC LCA; Lanza, et al., 2011), employ a maximum-likelihood routine that adjusts for MAR missingness, but not for MNAR missingness. Several simulation studies have documented the success of parameter recovery under various conditions. For situations where the  $\rho$  parameters are above .8 or below .2 for dichotomous indicators, parameter recovery for data that are MCAR or MAR is not substantially biased regardless of the amount of missing data, the sample size, or the latent class model (Hyatt & Collins, 1998; Kolb & Dayton, 1996). It is important to note that the maximum-likelihood missing data procedure will be fully successful only if all variables relevant to missingness are included in the model. If there are variables relevant to missingness that cannot be included, it may be preferable to use a multiple imputation approach to include these variables (Collins, et al., 2001).

## The Use of Parameter Restrictions

Restricted parameters are either *fixed*, where the value is set to a particular value and not estimated, or *constrained* to be equal to other parameters in an equivalence set, so that only one parameter is estimated for the entire set. In a classic article, Goodman (1974) presented the estimation of restricted latent class models using the EM algorithm. Often latent class models are fit without the use of any parameter restrictions. Such unrestricted models can be quite informative, and they are especially useful for exploring new models. However, parameter restrictions can help perform two important tasks in latent class models: achieving identification and specifying or testing specific features of a model.

Parameter restrictions can be useful in achieving identification, because fixing or constraining parameters reduces the number of parameters to be estimated. Underidentification refers to the situation when there are too many parameters to estimate given the information available in a certain data set. One necessary condition for identification is that the *number of independent parameters to be estimated* be less than the *number of possible response patterns*. In some cases, such as for latent class models with few possible response patterns, parameter restrictions can be used to achieve this condition. However, satisfying this condition does not ensure an identified model. Having a large sample size relative to the number of response patterns helps identification. When the sample size is small relative to the number of response patterns (i.e. the contingency table is *sparse*), parameter restrictions can greatly aid in identification.

Second, parameter restrictions are useful tools for specifying or testing various features of a model. For example, the  $\rho$  parameters define the meaning of the latent classes. If we have measurement invariance across groups then meaningful comparisons of the latent class prevalences can be made. Structural invariance across groups in the measurement of the latent variable is a testable hypothesis. The parameter estimates can be freely estimated in one analysis, and constrained equal across groups in another; then, the fit of the two models can be compared. As an example, a researcher interested in comparing gender differences in the prevalence of two temperament types in a sample of infants may wish to compare a model with all parameters freely estimated to one in which the item-response probabilities by class are constrained equal across the two genders. If measurement invariance by class can be established across groups (i.e., male and female), this is evidence that the same construct is being measured in males and females; therefore, meaningful cross-gender comparisons of the prevalence of the various temperament types can be made.

The hierarchical or Guttman model (Rindskopf, 1983) is another example of a latent class model for which there is a theoretical justification to impose parameter restrictions. In Guttman models, it is assumed that there is an order among the observed variables that form a scale. In models of learning, this order corresponds to the difficulty of passing each task assessed by the observed variables. For example, if three skills are measured and assumed to be hierarchical in level of difficulty, and a pass is denoted  $1$  and a fail denoted  $0$ , we might restrict the parameters so that the latent classes correspond to the patterns 000, 100, 110, and 111. The fit of such a model can then be compared to the fit of an unrestricted model to see if the Guttman scale holds in the data.

### Model Selection and Goodness-of-Fit

**Absolute model fit.** Choosing the number of latent classes is an important issue and can be somewhat subjective; the choice can be driven by both empirical evidence and theoretical reasoning. For example, if theory suggests that there are only two temperament types, the fit of a two-class model can be assessed. A more empirical approach would be to examine models with two, three and four latent classes to see which solution is most interpretable or provides the best fit. The typical approach to model fit in LCA is to compare the response pattern frequencies predicted by the model with the response pattern frequencies observed in the data. The predicted response pattern frequencies are computed based on the parameter estimates produced in the LCA. The two most common measures of fit in a contingency table analysis are the Pearson chi-square statistic,  $X^2$ , and the likelihood-ratio statistic,  $G^2$ . The likelihood-ratio statistic has the advantage that nested models can be compared by a likelihood-ratio test, with the resulting statistic distributed as chi-square, and thus is often preferred to the Pearson chi-square. The likelihood-ratio statistic is calculated by

$$G^2 = 2 \sum_y obs \log \left( \frac{obs}{exp} \right)$$

where  $y$  represents a response pattern (i.e., a cell in the contingency table formed by cross-tabulating all observed variables). This statistic expresses the degree of agreement between these predicted frequencies and the observed frequencies.

The  $G^2$  is asymptotically distributed as chi-square, with degrees of freedom equal to *number of possible response patterns* minus *number of parameters estimated* minus 1. The term

"asymptotically" means that a chi-square distribution is a good approximation when the number of observations in each cell of the contingency table is sufficiently large. However, latent class models can often involve large contingency tables, resulting in a contingency table with many sparsely populated cells. When the ratio  $N/k$  becomes small, where  $N$  is the sample size and  $k$  is the number of cells in the contingency table, the distribution of the  $G^2$  is not well-approximated by the chi-square distribution. Unfortunately, under these circumstances the true distribution of the  $G^2$  is not known, rendering it of limited utility for model selection. This applies to all contingency table models, and is particularly a problem in latent class models with large numbers of indicators. For a  $G^2$  difference test comparing two nested models, however, the distribution usually is better approximated by the chi-squared distribution and thus hypothesis testing is more reliable.

**Relative model fit.** A pair of nested models consists of a simpler model and a more complex model. Some parameters restricted in the simpler model are estimated in the more complex model. Thus, the simpler model can be considered a version of the more complex model. Two nested latent class models can be compared statistically by taking the difference of their  $G^2$  values. This difference is distributed as chi-square with degrees of freedom equal to the difference in the degrees of freedom associated with the two  $G^2$ s. If the difference in  $G^2$  is nonsignificant, it means that the more parsimonious model fits about as well as the more complex model, and thus there is no benefit to estimating the parameters in the more complex model. If the difference in  $G^2$  is significant, it means that the additional parameters estimated in the more complex model are necessary to achieve adequate fit. The  $G^2$  difference test can be quite useful when comparing various patterns of parameter restrictions for latent class models

with a given number of classes. For example, an investigator interested in conducting an omnibus test for gender differences in the prevalence of temperament types in infants can fit a model that constrains the  $\gamma$  parameters (the probability of membership in each temperament type) to be equal across males and females in one model, and fit a second model that freely estimates the  $\gamma$  parameters. The  $G^2$  difference test between the two models will indicate whether or not it is reasonable to impose this equality restriction. If the  $G^2$  difference test reaches statistical significance, then the conclusion would be that the probabilities of membership in the temperament types vary by gender.

Ideally, it would be helpful to use this approach to help determine the appropriate number of latent classes by comparing the fit of two models, such as when an investigator wants to compare a model with two temperament types to a model with three temperament types. Unfortunately, two models with different numbers of latent classes cannot be compared in this way because parameters of the simpler model take on boundary values of the parameter space (Everitt, 1988; Rubin & Stern, 1994). In this case, the distribution of the  $G^2$  test statistic is undefined (McLachlan & Peel, 2000).

Various model selection information criteria have been proposed for comparing models with different numbers of classes, including the Akaike Information Criterion (AIC; Akaike, 1974), Bayesian Information Criterion (BIC; Schwarz, 1978), consistent AIC (CAIC; Bozdogan, 1987), and adjusted BIC (a-BIC; Sclove, 1987). These information criteria are penalized log-likelihood test statistics, where the penalty is two times the number of parameters estimated for the AIC and the log of  $N$  times the number of parameters estimated for the BIC. Results of a

simulation study conducted by Lin and Dayton (1997) suggest that, although the AIC performs better than the BIC or the CAIC, the AIC tends to err on the side of selecting models that are more complex than the true model. Another drawback to this approach is that these methods serve only to compare the relative fit of several models under consideration, but do not help in determining whether a particular model has sufficiently good fit in an absolute sense. A parametric bootstrap of the goodness-of-fit measures (referred to as the bootstrap likelihood-ratio test, or BLRT) has also been proposed for latent class models (Collins, Fidler, Wugalter, & Long, 1993; Langeheine, Pannekoek, & van de Pol, 1996) and has been shown to perform well for latent class models with item-response probabilities at or above .8 (Nylund, Asparouhov, & Muthén, 2007). This method, also referred to as Monte Carlo sampling, involves repeatedly sampling from the model-based parameter estimates to get a distribution of the fit statistic under the assumption that the model is true. This method yields an empirical distribution of the fit statistic, forgoing the use of a theoretical distribution altogether.

A Bayesian approach to model monitoring using a posterior predictive check distribution has been proposed as a method for comparing models with different numbers of latent classes (Rubin, 1984; Rubin & Stern, 1994). This is an empirical, simulation-based procedure where the observed value of the test statistic is compared with its posterior predictive check distribution under the null model to determine if the data are consistent with the null model. The  $p$ -value indicates the probability of a result more extreme than the data under the posterior predictive check distribution of the test statistic. This is an appropriate method for comparing two models where there is a set of parameters distinguishing them, instead of a single parameter. For example, the posterior predictive check distribution can be used to test the validity of a five-class

model versus a six-class model. The first step is to draw a set of parameter values from their joint posterior distribution using a Markov chain sampling approach; specifically, a data augmentation algorithm. The second step is to create a data set the same size as the original using random draws from appropriate distributions with the chosen parameter values. The fit of the model to this imputed data set is assessed, and this process is repeated many times to obtain a distribution of the fit statistic. This is a very promising but complex approach, and widely available software for this procedure is not yet available. Hoijtink (1998; Hoijtink & Molenaar, 1997) presents several examples where the fit of constrained latent class models are assessed by applying the posterior predictive check distribution to several goodness-of-fit statistics.

In considering model selection in LCA, it may be helpful to draw a comparison with factor analysis. In factor analysis, multiple continuous observed variables are mapped onto several latent factors. In contrast, the latent class model maps multiple categorical observed variables onto several categories of a latent variable. In factor analysis both exploratory and confirmatory approaches can be used in selecting the number of factors. All factors with eigenvalues greater than one are often selected in an exploratory factor analysis. An exploratory approach to LCA might involve the user fitting a two-class solution to a data set, then a three-class solution, and so on, and comparing the various solutions in a rough way using fit statistics and criteria. In this framework there is no rule of thumb for selecting the smallest number of latent classes that can adequately explain the structure in the data. The closest analog might be to create a table that summarizes the  $G^2$  value, degrees of freedom, and information criteria for each number of classes fit to the data. The most parsimonious model (i.e., the model with the smallest number of classes) that provides adequate fit could be selected as the one with the most

appropriate number of latent classes. This approach is demonstrated in the empirical example on adolescent depression presented below.

### **A LATENT CLASS EXAMPLE: ADOLESCENT DEPRESSION**

We will illustrate the latent class model by examining adolescent depression in data from The National Longitudinal Study of Adolescent Health (Add Health; Harris et al., 2009). One theory suggests an underlying latent variable made up of two mutually exclusive and exhaustive groups of individuals: those who are depressed and those who are not. An alternate theory hypothesizes that there are several different types or levels of depression. These two competing theories can be examined empirically by fitting several different latent class models to the data set.

The dataset for this example comes from the Add Health study, which was mandated by Congress to collect data for the purpose of measuring the effect of social context on the health and well-being of adolescents in the United States. The first wave of the sample included 11,796 students 7<sup>th</sup> through 12<sup>th</sup> grades who were surveyed between April and December, 1995, and the second wave included the same individuals interviewed again between April and August of 1996. The sample used in the LCA includes all 1,043 adolescents (51.6% male) from the public-use dataset who were in 11<sup>th</sup> grade at Wave I.

The prevalence of depression in this sample will be explored by examining eight observed variables, listed in Table 1: four indicators of sadness, two indicators of feeling disliked by others, and two indicators related to feelings of failing at life. These six-level variables were recoded so that 1 represents never or rarely experiencing the symptom in the past week, and 2

represents experiencing the symptom sometimes, a lot, most of the time, or all of the time during the past week. It is important to have only a few levels for each variable because LCA is based on data in the form of a contingency table and extreme sparseness of cells can lead to problems in model estimation. As the number of levels within variables increases, so does the number of parameters in the model. A balance between retaining information in the original variables and collapsing categories must be struck, as identification can be difficult to attain when too many parameters are estimated.

---

Table 1 about here

---

The eight observed variables are manifest indicators of the latent variable Depression. In our example, LCA will allow us to arrive at a model that we believe best represents the relation between these eight variables and Depression. We are interested in the latent classes of depression that emerge, and the prevalence of each class.

For our example, we first fit a one-class solution, which yielded a  $G^2$  of 1837.5 with 247 degrees of freedom. We then fit models with two classes, three classes, four classes, and so on, up to eight classes. Table 2 reports for each model the degrees of freedom,  $G^2$  likelihood-ratio test statistic, four information criteria (AIC, BIC, CAIC, and a-BIC), BLRT, entropy, and the percentage of 1000 sets of random starting values that converged to the maximum-likelihood solution (referred to as Solution % in Table 2). This last statistic indicates how confident the investigator can be that the maximum-likelihood solution, as opposed to a local maximum, was, in fact, identified. The information criteria suggest that models between four (based on the BIC and CAIC) and eight (for the AIC) latent classes should be considered as optimizing the balance

between parsimony and fit. Upon careful inspection of the solutions corresponding to these models, along with the BLRT, entropy, and solution %, we selected the five-class solution ( $G^2$  of 240.6 with 211 degrees of freedom) as our final model.

---

Table 2 about here

---

### **The Five-Class Solution**

Table 3 contains the parameter estimates for the five-class solution. The matrix of item-response probabilities ( $\rho$  parameters) shows the probability of responding *sometimes, a lot, most of the time, or all of the time* to each of the eight observed variables given class membership.

(The probability of responding *never* or *rarely* to each observed variable given class membership can be calculated by subtracting this value from 1.0.) Probabilities greater than .50 are marked in bold to facilitate interpretation, indicating that individuals in that latent class are more likely to report that symptom than not. Interestingly, although no restrictions were imposed on the model parameters, the four indicators of sadness tend to operate similarly, as do the two indicators of being disliked and the two indicators of failure. The pattern of item-response probabilities lends itself to a straightforward interpretation of the five latent classes; thus, appropriate class labels can be assigned. Figure 3 presents the patterns of item-response probabilities graphically.

Individuals in Latent Class 1, Non-Depressed, are expected to have a low probability of reporting any depression symptom. Members of Latent Class 2, Sad, are likely to report experiencing all sadness-related symptoms but not likely to report the other symptoms. Similarly, Latent Class 3 is labeled Disliked, and Latent Class 4 is called Sad + Disliked. Latent Class 5, Depressed, is associated with a high probability of reporting all eight depression symptoms.

---

Table 3 about here

---

---

Figure 3 about here

---

Table 3 also reports the  $\gamma$  parameters, which represent the prevalence of each latent class. Note that these parameters must sum to 1, as they represent the distribution of a discrete random variable. The Non-Depressed latent class has the highest prevalence, with 39% of the sample expected to be members of this class. We expect 23%, 17%, and 15%, respectively, to be in the Sad, Disliked, and Sad + Disliked classes. Only 7% are expected to experience the highest level of depression.

### **Multiple-Groups LCA: Gender Differences in Depression Classes**

Now that a five-class model for depression has been established in this sample of adolescents, it may be useful to examine how the prevalence of each depression latent class varies across different groups. For example, we might be interested in whether the prevalence rates vary for males and females. This can be investigated by estimating the  $\gamma$  and  $\rho$  parameters for each group and comparing the latent class prevalences. However, in order for these comparisons to be meaningful, it may be useful to impose measurement invariance across gender. This involves applying constraints such that each item-response probability is the same for males and females. The unrestricted five-class model where all parameters are freely estimated for both genders, and an alternative model where measurement is constrained to be equal across gender, are nested and can be compared directly using a  $G^2$  difference test. This test

will indicate whether it is reasonable to conclude that there is measurement invariance across groups. Because this test can be highly sensitive when many parameters are involved, however, it also can be useful to examine information criteria in deciding whether the restricted model (with equal measurement for males and females) is sufficient. In our example, the model with all parameters estimated freely has a  $G^2$  of 384.8 with 423 degrees of freedom ( $df$ ; AIC = 560.8, BIC = 996.3), and the model that imposes measurement invariance across gender has a  $G^2$  of 441.2 with 463  $df$  (AIC = 537.2, BIC = 774.8). The  $G^2$  difference test is significant ( $G^2 = 56.4$ ,  $df = 40$ ,  $p = .04$ ), indicating that there are some differences between males and females in the measurement model. However, the information criteria suggest that the more parsimonious model is preferable given the number of  $df$  saved. After considering the measurement models for males and females when measurement varied across groups, we decided to impose measurement invariance by constraining the  $\rho$  parameters to be equal across gender.

Table 4 shows results for a latent class model of depression with gender as a grouping variable. The matrix of item-response probabilities is very similar to that in Table 3. The latent class prevalences for males and females can now be compared. We can conduct an omnibus test for differences in the distribution of males and females' prevalence of depression by constraining all elements of the vector of  $\gamma$  parameters to be equal across gender ( $G^2 = 467.8$ ,  $df = 467$ ), and comparing that to our present model ( $G^2 = 441.2$ ,  $df = 463$ ). The  $G^2$  difference test is highly significant ( $G^2 = 26.6$  with 4  $df$ ), indicating that the prevalences of the five latent classes of depression differ substantially between males and females. Table 4 shows that the proportion of adolescents expected to be in the Depressed latent class is similar for males and females (.06 for each group), but the Non-Depressed latent class is more prevalent among males (.44 for males

versus .30 for females). In addition, gender differences appear in the middle levels of depression. Female adolescents are more likely to be in the Sad latent class (.25 versus .18) and the Sad + Disliked latent class (.22 versus .12), whereas male adolescents are more likely to be in the Disliked latent class (.20 versus .16). Figure 4 shows the gender differences in the prevalences of depression latent classes.

---

Table 4 about here

---



---

Figure 4 about here

---

### **LCA with Covariates: Predicting Depression Latent Class Membership**

Covariates can be used to predict latent class membership in much the same way that predictors of an observed categorical outcome can be used in logistic regression analysis. An indicator for academic achievement (Grades) will be incorporated as a covariate in the model involving five latent classes of depression to determine whether it is a significant predictor of membership in the different latent classes. The Grades covariate represents the average grade on a four-point scale across English or language arts, mathematics, history or social studies, and science, and is based on the most recent grading period as reported by the adolescents ( $mean = 2.78$ ,  $sd = .75$ ). A multinomial logistic regression model was specified to predict latent class membership from Grades; because the outcome is latent, the logistic regression parameters are estimated simultaneously with the latent class model so that class membership uncertainty is taken into account. The Non-Depressed latent class was specified as the reference group so that

any effect of Grades could be expressed in terms of increased odds of membership in each latent class involving depression symptoms relative to the latent class involving no depression symptoms. The specification of a different reference latent class would yield a mathematically equivalent solution, but provide parameter estimates that express the association differently.

Table 5 shows the logistic regression parameter estimates, denoted  $\beta$ , as well as the corresponding odds ratios (exponentiated  $\beta$ s) and inverse odds ratios. Odds ratio confidence intervals not containing the value 1.0 indicate a significant effect of Grades. The results suggest that having higher grades is associated with lower odds of membership in the Disliked and Depressed latent classes relative to the Non-Depressed latent class.

---

Table 5 about here

---

### **LONGITUDINAL EXTENSIONS: REPEATED-MEASURES LCA AND LTA**

Often investigators in psychology are interested in developmental processes; that is, how constructs change over time. Development is often characterized by increases or decreases in a particular continuous variable over time. However, qualitative development can also be examined by modeling how individuals pass through categories or stages over time. Examples of stage sequences are abundant in psychology. These include Piaget's stage model of cognitive development (Piaget, 1973), Freud's stages of psychosocial development (Freud, 1961), Kohlberg's stage sequence of moral development (Kohlberg, 1966), and Erikson's stages of ego identity (Erikson, 1950). Several extensions of LCA to repeated-measures data have been proposed. Below we introduce two different frameworks for applying LCA to repeated-measures

data: repeated-measures LCA and LTA. We then extend our analysis of adolescent depression to two time points and demonstrate LTA in the context of this example.

Repeated-measures LCA involves fitting a standard latent class model, but using as indicators one or more observed variables that were assessed at multiple time points. In this case, the latent classes are characterized by discontinuous patterns of change over three or more times. As an example, Lanza and Collins (2006) presented an eight-class model fit to six repeated assessments of past-year heavy episodic drinking. Participants comprised a national sample assessed at Time 1, high school; Times 2 and 3, college-age; Times 4 and 5, young adulthood; and Time 6, age 30. Each latent class represented a longitudinal pattern of heavy drinking. For example, the 2.6% of participants in the College Age Only latent class were characterized by the absence of heavy drinking in high school, young adulthood and adulthood, but engaged in the behavior during college ages. Although heavy drinking rates among college students did not exceed heavy drinking rates among the non-collegiate group at any age, college-enrolled individuals were significantly more likely to belong to the College Age Only latent class. Non-collegiate individuals, however, even those who were not drinking heavily during high school or college ages, were found to be at increased risk for heavy drinking during adulthood.

LTA is another extension of LCA to repeated-measures data (Collins & Lanza, 2010; Collins & Wugalter, 1992). Continuing our examination of adolescent depression, we describe LTA in the context of examining stability and change in depression during adolescence. In LTA, individuals transition between latent classes over time; in other words, change over time in a discrete latent variable is measured. This framework provides a way to estimate and test models

of stage-sequential development (i.e., change in latent class membership over time) in longitudinal data. This approach allows researchers to estimate the prevalence of latent classes and the incidence of transitions to different latent classes over time for multiple groups and to predict initial latent class membership and transitions over time. It is worth noting that the multiple indicator latent Markov model (Langeheine, 1994; Langeheine & Van de Pol, 1994; Macready & Dayton, 1994) is closely related to LTA. Below some applications of LTA in the behavioral sciences are briefly described. The LTA mathematical model and related issues are then presented by extending the empirical example on adolescent depression.

Previously we discussed the modeling of substance use onset and smoking cessation within the latent class framework. Because individuals' class membership in each of these sequences can change over time, it may be interesting to collect data at more than one wave and examine patterns of change between consecutive times. The transtheoretical model of behavior change presented earlier (see Figure 2) is, in this case, a sequence of four stages: precontemplation, contemplation, action and maintenance. LTA has been used to test competing models of smoking behavior as individuals move from one stage to another (Velicer, et al., 1996; Martin, et al., 1996). As another example, the stage-sequential model of substance use onset has been examined extensively using LTA: Collins et al. (1997) examined the relationship between heavy caffeine use and adolescent substance use. Other covariates of substance use onset have been explored using LTA, including pubertal timing (Hyatt & Collins, 1999), parental permissiveness (Hyatt & Collins, 2000) and exposure to adult substance use (Tracy, Collins, & Graham, 1997).

A similar application of LTA incorporated a DSM-IV diagnosis of alcohol abuse and dependence (AAD) at age 21 as the grouping variable, and modeled group differences in the transitions through alcohol use during elementary, middle, and high school (Guo, Collins, Hill, & Hawkins, 2000). The behavior was characterized with the following four latent classes: Nonuse, Initiated Only, Initiated and Currently Using, and Initiated and Currently Using With Heavy Episodic Drinking. Different drinking patterns that emerged in middle and high school were related to individuals' subsequent AAD diagnosis. Current alcohol use in middle school was related to an AAD diagnosis at age 21, as was heavy episodic drinking in high school. This evidence suggests the need for different intervention programs at various developmental periods throughout adolescence.

More recently, Maldonado-Molina and Lanza (2010) demonstrated LTA as a useful tool for testing research questions motivated by the gateway hypothesis of drug use (Kandel, 2002). LTA provided a way to clearly operationalize and test gateway relations between two different drugs. Data from a national sample of adolescents were used to test gateway relations between cigarettes and marijuana, alcohol and marijuana, and alcohol and cigarettes. Results suggested that alcohol use served as a gateway drug for both recent cigarette use and recent marijuana use, but statistical evidence was not found for cigarettes as a gateway drug for either alcohol or marijuana use.

LTA has also been used to model change from year to year in the safety of sexual behaviors among injection drug users (Posner, Collins, Longshore, & Anglin, 1996). Four subscales were used to estimate transitions in latent class membership across two times:

knowledge, denial of personal risk, self-protective behaviors, and sexual risk behavior. A six-class solution that allowed individuals to move to any latent class over time fit the data well. The six latent classes were labeled High Risk; Knowledge Only; Safe Sex Only; Sex Risk; Low Risk, but Denial; and Low Risk. It was found that individuals moved to a different latent class over time quite often, indicating that sexual behavior prevention efforts should be ongoing and provided to all injection drug users regardless of the safety of their current sexual behaviors.

LTA can be used to aid both the design and evaluation of intervention programs. Competing models can be fit in order to describe developmental processes for which researchers may want to develop intervention programs. In addition, by incorporating a grouping variable, it is possible to determine whether certain subgroups of the population are at higher risk, providing useful information to researchers as they design interventions for prevention or treatment. For example, Collins et al. (1994) explored differences in substance use onset for Anglo, Latino, and Asian-American adolescents and discussed possible implications for intervention design.

LTA also allows researchers to assess the effectiveness of interventions by incorporating treatment as a grouping variable. Because the outcome process is broken down into a stage sequence, differential effects of the intervention can be identified. The Adolescent Alcohol Prevention Trial (Hansen & Graham, 1991) was a school-based substance use intervention program. LTA was used to evaluate the effectiveness of a normative education curriculum. An overall program effect was detected, such that individuals who received the education were less likely to advance in substance use between 7<sup>th</sup> and 8<sup>th</sup> grades. More specifically, adolescents who received the education were more likely to stay in the No Use latent class in 8<sup>th</sup> grade if they

were in that stage in 7<sup>th</sup> grade, and less likely to begin advanced use during that time (Graham, Collins, Wugalter, Chung, & Hansen, 1991).

We will present LTA by modeling change in depression from 11<sup>th</sup> grade to 12<sup>th</sup> grade. Building on the previously described latent class example, we will use the same eight observed variables from the Add Health depression index, in this case measured at two times. The sample used in this longitudinal analysis includes all adolescents who were in 11<sup>th</sup> grade at Wave I and 12<sup>th</sup> grade at Wave II. We will explore gender differences in the stability of depression over time. First an overall model will be fit to the two waves of data for all adolescents, followed by a model incorporating the grouping variable sex.

The mathematical model will be presented in terms of  $T$  times of measurement, with depression as the dynamic (changing) latent variable. Note that the model can be extended to include a grouping variable and covariates; these extensions are discussed below (we refer readers to Collins & Lanza, 2010 and Lanza & Collins, 2008 for mathematical details of this extended model). There are  $S_1 = 1, \dots, S$  latent classes at Time 1 and  $S_2 = 1, \dots, S$  latent classes at Time 2. Let  $\mathbf{y}$  represent a response pattern, a vector of possible responses to the eight observed variables indicating latent class membership at Time 1 and the same eight variables at Time 2 (i.e., a cell of the contingency table made by crosstabulating all observed variables in the model). In this example, the response pattern is a vector made up of the eight depression indicators at Time 1 and the same eight indicators measured at Time 2. Let  $\mathbf{Y}$  represent the complete array of response patterns. The LTA model can be expressed as follows. Using an indicator function

$I(y_{j,t} = r_{j,t})$  that equals 1 when response  $r_{j,t}$  is given to variable  $j$  at Time  $t$ , and equals 0 otherwise, the probability of observing a particular response pattern,  $P(\mathbf{Y} = \mathbf{y})$ , is expressed as

$$P(\mathbf{Y} = \mathbf{y}) = \sum_{s_1=1}^S \cdots \sum_{s_T=1}^S \delta_{s_1} \tau_{s_2|s_1} \cdots \tau_{s_T|s_{T-1}} \prod_{t=1}^T \prod_{j=1}^J \prod_{r_{j,t}=1}^{R_j} \rho_{t,j,r_{j,t}|s_t}^{I(y_{j,t}=r_{j,t})},$$

where  $\delta_{s_1}$  is the probability of membership in latent class  $s_1$  at Time 1 (e.g. the probability of being in the Sad latent class in 11<sup>th</sup> grade);  $\tau_{s_2|s_1}$  is the probability of membership in latent class  $s_2$  at Time 2 conditional on membership in latent class  $s_1$  at Time 1 (e.g. the probability of membership in the Sad latent class in 12<sup>th</sup> grade given membership in the Not-Depressed latent class in 11<sup>th</sup> grade); and  $\rho_{t,j,r_{j,t}|s_t}$  is the probability of response  $r_{j,t}$  to observed variable  $j$  at Time  $t$ , conditional on membership in latent class  $s_t$  at Time  $t$  (e.g., the probability of reporting that one “Could not shake the blues” in 11<sup>th</sup> grade given membership in the Sad latent class at that time). In general, there can be any number of observed variables and response categories, and the number of response categories can vary across variables. As in LCA, LTA employs the EM algorithm. Issues of model selection and fit are identical to those in the latent class framework.

The use of parameter restrictions serves the same purpose as in a latent class framework. However, there are several reasons why these restrictions often play a larger role in LTA models. First, because complex models are often estimated in LTA, the number of parameters involved can be quite large. Imposing restrictions on the parameters can greatly aid in model identification. Second, as in all longitudinal models, it is important to consider the issue of measurement invariance over time. This can be explored by imposing constraints across the sets

of  $\rho$  parameters, or item-response probabilities, estimated at each time. Constraining these parameters to be equal across times ensures that the meaning of the latent classes remains consistent across times. If the measurement is structurally identical at both times then changes in the class membership over time can be attributed solely to development rather than development mixed with changes in the relations between the indicators and the latent variable. Third, parameter restrictions often play an important theoretical role in the transition probabilities; restrictions can be imposed to test the nature of development over time. For example, models allowing all the  $\tau$  parameters to be freely estimated (implying that individuals can move between latent classes freely over time) can be compared to models which do not allow backsliding to stages earlier in a stage sequence by fixing certain  $\tau$  parameters to 0.

Table 6 presents results from a latent transition model with five latent classes of depression across two times. Measurement invariance across time was tested by comparing a model with all  $\rho$  parameters estimated freely to one in which  $\rho$  parameters were constrained to be equal across time. The  $G^2$  difference test for these nested models is  $G^2_{diff}=33.1$  with 40  $df$ ,  $p > .05$ . This test provides evidence that measurement invariance over time holds. Therefore, all subsequent findings are based on the more parsimonious model (i.e., one in which measurement was constrained to be equal over time).

---

Table 6 about here

---

The depression latent class prevalences at each grade show that the overall proportion of students at each level is quite stable between 11<sup>th</sup> and 12<sup>th</sup> grades, with a slightly higher proportion in the Non-Depressed latent class in 12<sup>th</sup> grade. The matrix of transition probabilities provides information about intra-individual change in depression between 11<sup>th</sup> and 12<sup>th</sup> grades. Elements on the diagonal are bolded; these numbers represent the probability of membership in a particular depression latent class in 12<sup>th</sup> grade conditional on membership in the same depression latent class in 11<sup>th</sup> grade. Interestingly, it is estimated that 75% of adolescents in the Non-Depressed latent class at 11<sup>th</sup> grade were in that same latent class a year later. The probability of being in the same depression latent class over time is somewhat lower for the Depressed latent class (65%), and considerably lower for the remaining three groups.

### **Multiple-Groups LTA and LTA with Covariates**

Grouping variables and covariates can be incorporated in LTA in much the same way as in LCA (see Collins & Lanza, 2010 for a detailed presentation of LTA, including multiple-groups LTA and LTA with covariates). Multiple-groups LTA is useful when a test of measurement invariance across groups is desired; often of more scientific interest is an examination of group differences in latent class prevalences and in transition probabilities. In our empirical demonstration we next incorporate gender as a grouping variable ( $N=538$  males,  $N=506$  females) in order to examine gender differences in the prevalence of depression classes and in the incidence of transitions over time in depression. The  $\rho$  parameters were constrained to be equal across time and gender so that the latent classes of depression are defined in the same way in 11<sup>th</sup> and 12<sup>th</sup> grades for males and females. Table 7 shows the latent class prevalences at each time and the transition probabilities for males and females. As seen by the overall latent

class membership probabilities, the prevalence of the Disliked latent class increases over time for males, from 23% to 27%, whereas the prevalence decreases for females, from 17% to 12%. The relative increase in the prevalence of the Non-Depressed latent class was larger for females (26% to 35%) than for males (37% to 39%).

---

Table 7 about here

---

Overall, the pattern of transition probabilities appears to be generally similar for males and females. For example, 72% of 11<sup>th</sup>-grade males in the Non-Depressed latent class and 42% of those in the Sad latent class are in this same latent class in 12<sup>th</sup> grade. Similar rates of membership in the same latent class over time are observed for females (79% for Non-Depressed and 45% for Sad). However, several interesting but subtle gender differences in the transition probabilities emerge. Notice that, among adolescents in the Non-Depressed, Sad, Disliked, or Sad+Disliked latent class in 11<sup>th</sup> grade, males appear to be more likely than females to transition to (or remain in) the Disliked latent class in 12<sup>th</sup> grade. In contrast, among adolescents in the Disliked latent class, females appear to be more likely than males to transition to the Depressed latent class (20% for females versus 5% for males). These parameter estimates also indicate that females are less likely than males to remain in the Depressed latent class over time (54% for females versus 79% for males).

Covariates can be included in LTA models in order to predict initial latent class prevalences or to predict transitions over time between latent classes (e.g., Humphreys & Janson, 2000; Lanza & Collins, 2008; Pfefferman, et al., 1998; Reboussin, Reboussin, Liang, & Anthony, 1998). In the current example, including covariates would allow us to address

questions such as (1) “How does involvement in high school sports relate to depression latent class membership in 11<sup>th</sup> grade?” (2) “Does this relation vary across gender?” and (3) “Given 11<sup>th</sup>-grade depression class, how is regular alcohol use related to 12<sup>th</sup>-grade depression class?”

As another example of LTA with covariates, Lanza and Collins (2008) fit a latent transition model of adolescent sexual risk behavior across three annual time points. The five latent classes that emerged were Nondaters, Monogamous, Daters, Multi-Partner Safe, and Multi-Partner Exposed. By including a covariate measuring recent heavy episodic drinking (drunkenness), they were able to address questions including (1) “How does drunkenness relate to membership in the five latent classes?” and (2) “For each initial class of sexual risk behavior, how is drunkenness related to the probability of transitioning to a latent class characterized by high-risk sexual behavior?” The authors found that adolescents who reported drunkenness at Time 1 were 8.4 times more likely than adolescents not reporting drunkenness to belong to the high-risk class Multi-Partner Exposed relative to the Non-Daters class. Further, Time 1 drunkenness was found to be a significant risk factor for making a transition to high-risk sexual behavior, particularly among adolescents who were initially in the Non-Daters and Daters latent classes. The inclusion of covariates in LTA can provide a nuanced view of the behavior change.

### **SOME RECENT EXTENSIONS TO LCA**

Here we describe several extensions of LCA and LTA, as well as other closely-related finite mixture models that hold much promise for addressing questions in psychology research.

### **Associative LTA**

Recently, the LTA framework has been extended to model two developmental processes simultaneously over time (Bray, Lanza, & Collins, 2010; Flaherty, 2008a,b). Associative LTA (ALTA) can be used to address a number of interesting research questions, including how one discrete process (e.g., transitions through latent classes of substance use) predicts another (e.g., transitions through latent classes of depression). For example, this approach would be useful to address questions about how transitions over time in depression are linked to transitions in drinking patterns. Suppose we wish to examine initial level and changes in depression (say, as modeled in our example discussed previously) as they relate to initial level and changes in drinking patterns. Some questions may pertain to a single time of measurement; for example, how are people in various drinking pattern latent classes (e.g., No Use; Infrequent Light Drinking; Infrequent Heavy Drinking; Frequent Heavy Drinking) at one time of measurement distributed among depression latent classes? Are Frequent Heavy Drinkers more likely to be in one depression latent class than another? Also, are Frequent Heavy Drinkers more likely to be in a particular depression latent class than Infrequent Heavy Drinkers? Other questions may pertain to change over time. With ALTA, interesting and complex patterns of contingent change can be examined. For example, are people who change from No Use to Infrequent Heavy Drinking, compared to those who remain in the No Use latent class, more likely to transition from no Depression to a particular depression latent class? Are people who remain in the Infrequent Light Drinking latent class more or less likely to remain in the No Depression latent class, compared with those who advance from the Infrequent Light Drinking latent class to a more advanced level of drinking behavior?

ALTA can be used to examine cross-sectional and longitudinal associations between processes, which may stem from lagged or concurrent effects of one process on the other. It can also be used to test specific hypotheses about these effects. Applications of the ALTA approach have examined the links between alcohol use and sexual behavior (Bray et al., 2010), tobacco use and alcohol use (Flaherty, 2008a), psychological state and substance use (Flaherty, 2008b), and negative affect and alcohol use (Witkiewitz & Villarroel, 2009).

### **Group-Based Trajectory Analysis**

Over the past decade, an extension of LCA called group-based trajectory analysis (Nagin, 2005; Nagin & Tremblay, 1999) or general growth mixture modeling (Muthén & Shedden, 1999) has been applied in many studies. Using this framework, heterogeneity in intra-individual development is examined by modeling subgroups that are identified by individual growth parameters, such as intercept and slope. In other words, two or more latent classes are estimated, each of which is characterized by a different mean growth curve trajectory. For example, Nagin and Tremblay used this approach to identify four latent classes of individuals based on their trajectory in physical aggression from age 6 to 15. The latent classes included a subgroup characterized by low aggressive behavior over time (14.4%), a moderate declining subgroup characterized by desistance in adolescence (53.7%), a high declining subgroup characterized by elevated aggression in childhood but a modest level in adolescence (27.6%), and a subgroup of individuals displaying chronic aggression across the whole range of ages (4.3%). Predictors of latent class membership can be included as covariates in this framework in order to study characteristics that are related to the different developmental patterns.

### **Other Variations of Mixture Models**

Mixture models such as LCA provide an opportunity to explore and explain heterogeneity in populations by identifying a set of subgroups comprised of individuals with shared characteristics. ALTA and group-based trajectory analysis are two different approaches to understanding heterogeneity in developmental processes or behavioral patterns over time. There are numerous other mixture models that hold promise for research in psychology; an edited book by Hancock and Samuelson (2007) includes introductions to and applications of several of these. One such model is the mixture item response theory (IRT) model (e.g., Mislevy, Levy, Kroopnick, & Rutstein, 2008), which relaxes the assumption that the performance tasks behave in the same way for all examinees. This approach allows the examinee and task parameters to be conditioned on latent class, and may be useful in detecting population subgroups based on, for example, the cognitive strategies they employed in performing tasks. Another mixture model that may be useful in psychological research is the factor mixture model (Lubke & Spies, 2007). An example by Kuo, Aggen, Prescott, Kendler, and Neale (2008) demonstrated how a model involving both a continuous factor and a latent class variable provided better fit to alcohol dependence measures than either a factor model or a latent class model alone. The three-class factor model was used to identify groups labeled Non-Problem Drinking, Moderate Dependence, and Severe Dependence. This type of mixture model may provide a basis for an alternative classification system to the DSM-IV diagnostic system.

### **Multilevel LCA**

Random effects (or mixed) models are popular in social science research. They provide a powerful way to accommodate clustered responses, which violate the standard data analysis

assumption that observations are independently sampled. Much of the work in this domain has been in the area of regression modeling (e.g., Raudenbush & Bryk, 2002; Hedeker & Gibbons, 1994; Hedeker & Mermelstein, 1998). The inclusion of random effects in both latent class (Qu, Tan, & Kutner, 1996) and latent transition (Humphreys, 1998) models has been discussed, and multilevel LCA is receiving growing attention. For instance, multilevel LCA recently has been used to examine cigarette smoking in rural communities (Henry & Muthén, 2010), individual- and community-level predictors of heavy alcohol use (Rindskopf, 2006), intervention effects in group-randomized trials (Van Horn, et al., 2008), and gender differences in mathematics achievement (Muthén & Asparouhov, 2009).

By adding random effects to the latent class model, it is possible to expand the assumption of conditional independence to include the situation where item-response probabilities are conditionally independent given latent class membership and unmeasured, subject-specific random effects. These random effects account for unmeasured factors that induce relations among item responses beyond those due to latent class membership. Two common situations that often employ random effects are repeated measures on the same person and a design with people nested within organizations like schools.

Consider the case of students nested within classrooms. The incorporation of random effects makes it possible to account for the fact that students in the same classroom tend to report more similarly than students in different classrooms. Rather than measuring and modeling all the different factors that could be causing this observed similarity, random effects models account for the similarity automatically. This would allow, for example, the probability of belonging to a

particular depression latent class to vary across classrooms and the item-response probabilities for the depression latent classes to be held constant across classrooms. This conceptualization of multilevel LCA is known as the parametric approach (Vermunt, 2003, 2008; Asparouhov & Muthén, 2008). A competing nonparametric approach has also been discussed (Vermunt, 2003, 2008; Asparouhov & Muthén, 2008). It could be used to identify, for example, latent classes of classrooms that have similar high and low probabilities of membership in the depression latent classes in order to account for the random effects of classrooms. These two approaches focus on latent class-specific random effects, but it is also possible to incorporate item-specific random effects (Asparouhov & Muthén, 2008; Henry & Muthén, 2010) that can be used to examine, for example, how classrooms influence the individual observed variables that measure the depression latent classes. Note that random effects do not normally constitute factors that are of substantive interest. If there were additional substantively interesting variables that accounted for within-class similarity, then these other factors should be measured and modeled explicitly.

### **Additional Types of Indicators**

The latent class model can be generalized to accommodate ordinal and continuous indicators. For example, Kim & Böckenholt (2000) presented the stage-sequential ordinal (SSO) model. The SSO model estimates stage-sequential development measured by one or more ordinal indicators. Measurement error in the ordinal data is estimated according to a graded-response model (Samejima, 1969), which assumes that there is an underlying continuous characteristic that is discretized to form an ordinal scale. Interval- and ratio-level indicators have been added by modeling them as normally distributed, conditional on latent class membership (Moustaki, 1996). In this model, categorical indicators are treated the same as they are in the basic latent

class model and continuous indicators are accommodated by regressing them on the latent class variable. As with the categorical indicators in the latent class model, the continuous indicators are assumed to be conditionally independent given latent class membership. Just as the item-response probabilities ( $\rho$  parameters) characterize the relation between the latent variable and the categorical indicators, means and residual variances of the continuous indicators characterize the relation between the continuous indicators and the latent class variable. If the mean levels of a continuous indicator are different across the latent classes and the within-class residual variance is small, then that observed variable does a good job of discriminating among the classes.

#### **AREAS FOR FUTURE RESEARCH RELATED TO LCA**

One area where more work is needed is statistical power for testing the fit of complex latent class and latent transition models. In order to assess the statistical power of any test involving the  $G^2$ , the distribution of the  $G^2$  under the alternative hypothesis, (i.e., the noncentral distribution), is needed. Because the central distribution of the  $G^2$  is unknown when the contingency table is sparse, it follows that the noncentral distribution is also unknown. Of course, the usual factors that affect statistical power of a test statistic (namely, sample size, choice of alpha, and effect size) operate in this context, which in model selection is the overall difference between the true model and the model under consideration. A parametric bootstrap procedure can be used to empirically derive the distribution of the  $G^2$  test statistic under the alternative hypothesis, providing one basis for studying factors related to the power of detecting underlying latent classes.

In addition to statistical power considerations, there are other sample size considerations in LCA and LTA. Sparseness can impair parameter estimation. Several simulation studies (Collins and Wugalter, 1992; Collins et al., 1993; Collins & Tracy, 1997) have demonstrated that parameter estimation in LTA remains unbiased even in very sparse data tables. However, the standard errors of certain parameters can become unacceptably large. This is particularly true for the  $\tau$  parameters, because these transition probabilities are conditioned latent class membership at Time 1, and thus often based on considerably smaller  $N$ s than the other parameters. In addition, estimation of the effects of covariates on latent class membership and on transition probabilities can be increasingly problematic as sparseness increases. This is of particular concern when predicting latent class membership or transitions that are quite rare. Bayesian methods, specifically the inclusion of data-derived prior information, have been highly effective for solving some of the issues related to sparseness in LCA and LTA (e.g., Lanza et al., 2011).

Another area for future research is relating latent class membership to later outcomes. There are currently two dominant approaches, both involving a classify-analyze strategy. The first approach is to assign individuals to latent classes based on their maximum posterior probability, and then conduct the analysis (e.g., regressing the distal outcome on latent class assignment) as if latent class membership is known. This strategy does not take into account uncertainty associated with latent class membership; therefore, inference regarding the association between the latent class variable and the distal outcome may be biased. The second approach, referred to as the multiple pseudo-class draws technique, involves multiple (e.g., 20) random assignment to latent classes with probability equal to an individual's posterior probabilities (Wang, Brown, & Bandeen-Roche, 2005). This approach yields multiple datasets,

each containing a variable indicating class assignment based on that particular draw. The association between latent class membership and the distal outcome is calculated within each dataset, and then results are combined across datasets. An advantage of the first approach is that the technique is conceptually and computationally straightforward; the primary advantage of the pseudo-class draws approach is that it attempts to take uncertainty in latent class membership into account by averaging over multiple draws.

Another area for future consideration is the use of LCA as a measurement model to develop assessment tools. Work is being done on this topic in sociology with a focus on survey data (see Biemer, 2011), and much of this work may be relevant to research in psychology. Use of LCA as a measurement model is particularly important if one plans to assign people to latent classes; for example, in the context of adaptive treatment assignment. The latent class model is, in fact, a measurement model for an unobserved categorical variable, but in most applications it is used in an exploratory fashion, where the latent class solution provides a way to model heterogeneity in the data. By using LCA as a measurement model, however, psychologists would be able to identify indicators that are most helpful in measuring the underlying variable, as well as indicators that are not helpful and might be removed from future surveys. This approach could lead to the development of widely-used psychological assessment instruments for measuring underlying subgroups that cannot be directly observed.

## **CONCLUSIONS**

The usefulness of latent class models in psychological research will continue to grow as the model is extended. Important steps have already been taken to improve the utility of this

method, including extensions to longitudinal data, improvements in the assessment of model fit, and the incorporation of continuous predictors of latent class membership. Additional work in areas such as estimation with small sample sizes, Bayesian methods, concomitant variables and two-sequence models will further increase the power of latent class and related methods. LCA and LTA play an important role in understanding categorical latent constructs in psychology and their related processes of change over time.

**REFERENCES**

- Aitkin, M., Anderson, D., & Hinde, J. (1981). Statistical modelling of data on teaching styles. *Journal of the Royal Statistical Society (Series A)*, *144*, 419-461.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716-723.
- Asparouhov, T., & Muthén, B. (2008). Multilevel mixture models. In G. R. Hancock & K. M. Samuelson (Eds.), *Advances in latent variable mixture models* (pp. 27-51). Charlotte, NC: Information Age Publishing.
- Bandeen-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.
- Biemer, P. P. (2011). *Latent class analysis of survey error*. Hoboken, NJ: Wiley.
- Bozdogan, H. (1987). Model-selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, *52*, 345-370.
- Bray, B. C., Lanza, S. T., & Collins, L. M. (2010). Modeling relations among discrete developmental processes: A general approach to associative latent transition analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(4) 541-569.
- Coffman, D. L., Patrick, M. E., Palen, L. A., Rhoades, B. L., & Ventura, A. K. (2007). Why do high school seniors drink? Implications for a targeted approach to intervention. *Prevention Science*, *8*, 241-248.

- Collins, L. M., Fidler, P. L., Wugalter, S. E., & Long, J. D. (1993). Goodness-of-fit testing for latent class models. *Multivariate Behavioral Research, 28*, 375-389.
- Collins, L. M., Graham, J. W., Rousculp, S. S., Fidler, P. L., Pan, J., & Hansen, W. B. (1994). Latent transition analysis and how it can address prevention research questions. In L. M. Collins & L. A. Seitz (Eds.), *Advances in data analysis for prevention intervention research*. Rockville, MD: National Institute on Drug Abuse.
- Collins, L. M., Graham, J. W., Rousculp, S. S., & Hansen, W. B. (1997). Heavy caffeine use and the beginning of the substance use onset process: An illustration of latent transition analysis. In K. J. Bryant & M. Windle (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research*. Washington, DC: American Psychological Association.
- Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. New York: Wiley.
- Collins, L. M., Schafer, J. L., & Kam, C. (2001). A comparison of inclusive and restrictive missing-data strategies with ML and MI. *Psychological Methods, 6*, 330-351.
- Collins, L. M., & Tracy, A. J. (1997). Estimation in complex latent transition models with extreme data sparseness. *Kwantitatieve Methoden, 55*, 57-71.
- Collins, L. M., & Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. *Multivariate Behavioral Research, 27*, 131-157.

- Dayton, C. M., & Macready, G. B. (1988). Concomitant-variable latent-class models. *Journal of the American Statistical Association*, *83*, 173-178.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B*, *39*, 1-38.
- Erikson, E. H. (1950). *Childhood and society*. New York: W. W. Norton.
- Everitt, B. S. (1988). A Monte Carlo investigation of the likelihood ratio test for number of classes in latent class analysis. *Multivariate Behavioral Research*, *23*, 531-538.
- Evers, K. E., Harlow, L. L., Redding, C. A., & LaForge, R. G. (1998). Longitudinal changes in stages of change for condom use in women. *American Journal of Health Promotion*, *13*, 19-25.
- Flaherty, B. P. (2008a). Examining contingent discrete change over time with associative latent transition analysis. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (p. 77-103). Charlotte, NC: Information Age.
- Flaherty, B. P. (2008b). Testing the degree of cross-sectional and longitudinal dependence between two discrete dynamic processes. *Developmental Psychology*, *44*(2) 468-480.
- Freud, S. (1961). The Ego and the Id. *Standard Edition*, *19*, 3-66.

- Gebhardt, W. A., Dusseldorp, E., & Maes, S. (1999). Measuring transitions through stages of exercise: Application of latent transition analysis. *Perceptual and Motor Skills*, 88, 1097-1106.
- Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, 61, 215-231.
- Graham, J. W., Collins, L. M., Wugalter, S. E., Chung, N. K., & Hansen, W. B. (1991). Modeling transitions in latent stage-sequential processes: A substance use prevention example. *Journal of Consulting and Clinical Psychology*, 59, 48-57.
- Guo, J., Collins, L. M., Hill, K. G., & Hawkins, J. D. (2000). Developmental pathways to alcohol abuse and dependence in young adulthood. *Journal of Studies on Alcohol*, 61, 799-808.
- Hancock, G. R., & Samuelson, K. M. (2007). *Advances in latent variable mixture models*. Charlotte, NC: Information Age Publishing.
- Hansen, W. B., & Graham, J. W. (1991). Preventing alcohol, marijuana, and cigarette use among adolescents: Peer pressure resistance training versus establishing conservative norms. *Preventive Medicine*, 20, 414-430.
- Harris, K. M., Halpern, C. T., Whitsel, E., Hussey, J., Tabor, J., Entzel, P., & Udry, J. R. (2009). *The National Longitudinal Study of Adolescent Health: Research Design*. Retrieved from <http://www.cpc.unc.edu/projects/addhealth/design>.
- Hedeker, D., & Gibbons, R. D. (1994). A random-effects ordinal regression model for multilevel analysis. *Biometrics*, 50, 933-944.

- Hedeker, D., & Mermelstein, R. J. (1998). A multilevel thresholds of change model for analysis of stages of change data. *Multivariate Behavioral Research, 33*, 427-455.
- Henry, K. L., & Muthén, B. (2010). Multilevel latent class analysis: An application of adolescent smoking typologies with individual and contextual predictors. *Structural Equation Modeling: A Multidisciplinary Journal, 17*(2), 193-215.
- Hojtink, H. (1998). Constrained latent class analysis using the gibbs sampler and posterior predictive p-values: Applications to educational testing. *Statistica Sinica, 8*, 691-711.
- Hojtink, H., & Molenaar, I. W. (1997). A multidimensional item response model: Constrained latent class analysis using the Gibbs sampler and posterior predictive checks. *Psychometrika, 62*, 171-189.
- Humphreys, K. (1998). A latent Markov chain with multivariate random effects. Sociological Methods and Research, 26, 269-299.
- Humphreys, K. & Janson, H. (2000). Latent transition analysis with covariates, nonresponse, summary statistics and diagnostics: Modelling children's drawing development. *Multivariate Behavioral Research, 35*, 89-118.
- Hyatt, S. L., & Collins, L. M. (1998). *Estimation in latent transition models with missing data*. (Technical Report Number 98-22). University Park: The Methodology Center, Penn State.

- Hyatt, S. L., & Collins, L. M. (1999). *Pubertal timing and the onset of substance use: Detailed results based on data augmentation*. (Technical Report Number 99-35). University Park: The Methodology Center, Penn State.
- Hyatt, S. L., & Collins, L. M. (2000). Using latent transition analysis to examine the relationship between parental permissiveness and the onset of substance use. In J. Rose, L. Chassin, C. Presson, & S. J. Sherman (Eds.), *Multivariate applications in substance use research: New methods for new questions*. Hillsdale, NJ: Erlbaum, pp. 259-288.
- Hyatt, S. L., Collins, L. M., & Schafer, J. L. (1999). *Using data augmentation to conduct hypothesis testing in LTA*. (Technical Report Number 99-34). University Park: The Methodology Center, Penn State.
- Jöreskog, K., & Sörbom, D. (1996). *Lisrel 8: User's reference guide*. Chicago: Scientific Software International.
- Kagan, J. (1989). Temperamental contributions to social behavior. *American Psychologist*, *44*, 668-674.
- Kandel, D. (1975). Stages in adolescent involvement in drug use. *Science*, *190*, 912-914.
- Kandel, D. (2002). *Stages and pathways of drug involvement: Examining the gateway hypothesis*. Cambridge, MA: University Press.
- Kim, J., & Böckenholt, U. (2000). Modeling stage-sequential change in ordered categorical responses. *Psychological Methods*, *5*, 380-400.

- Kohlberg, L. (1966). Cognitive stages and preschool education. *Human Development, 9*, 5-17.
- Kolb, R. R., & Dayton, C. M. (1996). Correcting for nonresponse in latent class analysis. *Multivariate Behavioral Research, 31*, 7-32.
- Kuo, P.-H., Aggen, S. H., Prescott, C. A., Kendler, K. S., & Neale, M. C. (2008). Using a factor mixture modeling approach in alcohol dependence in a general population sample. *Drug and Alcohol Dependence, 98*, 105-114.
- Langeheine, R. (1994). Latent variables Markov models. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research*. Thousand Oaks, CA: Sage.
- Langeheine, R., Pannekoek, J., & van de Pol, F. (1996). Bootstrapping goodness-of-fit measures in categorical data analysis. *Sociological Methods and Research, 24*, 492-516.
- Langeheine, R., & Van de Pol, F. (1994). Discrete-time mixed Markov latent class models. In A. Dale & R. Cavies (Eds.), *Analyzing social and political change: A casebook of methods* (pp. 170-197). London: Sage.
- Lanza, S. T., & Collins, L. M. (2008). A new SAS procedure for latent transition analysis: Transitions in dating and sexual risk behavior. *Developmental Psychology, 44*(2), 446-456.
- Lanza, S. T., & Collins, L. M. (2006). A mixture model of discontinuous development in heavy drinking from ages 18 to 30: The role of college enrollment. *Journal of Studies on Alcohol, 67*, 552-561.

- Lanza, S. T., Collins, L. M., Schafer, J. L., & Flaherty, B. P. (2005). Using Data Augmentation to Obtain Standard Errors and Conduct Hypothesis Tests in Latent Class and Latent Transition Analysis. *Psychological Methods, 10*, 84-100.
- Lanza, S. T., Dziak, J. J., Huang, L., Xu, S., & Collins, L. M. (2011). *Proc LCA & Proc LTA users' guide* (Version 1.2.6). University Park: The Methodology Center, Penn State. Available from <http://methodology.psu.edu>.
- Lanza, S. T., & Rhoades, B. L. (in press). Latent class analysis: An alternative perspective on subgroup analysis in prevention and treatment. *Prevention Science*.
- Lanza, S. T., Rhoades, B. L., Nix, R. L., Greenberg, M. T., & The Conduct Problems Prevention Research Group. (2010). Modeling the interplay of multilevel risk factors for future academic and behavior problems: A person-centered approach. *Development and Psychopathology, 22*, 313-335.
- Lanza, S. T., Savage, J., & Birch, L. (2010). Identification and prediction of latent classes of weight loss strategies among women. *Obesity, 18*(4), 833-840.
- Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Boston, MA: Houghton Mifflin.
- Lin, T. H., & Dayton, C. M. (1997). Model selection information criteria for non-nested latent class models. *Journal of Educational and Behavioral Statistics, 22*, 249-264.

- Lubke, G.H., & Spies, J. (2008). Choosing a 'correct' factor mixture model: Power, limitations, and graphical data exploration. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models*. Charlotte, NC: Information Age Publishing, 343-362.
- Macready, G. B., & Dayton, C. M. (1994). Latent class models for longitudinal assessment of trait acquisition. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research*. Thousand Oaks, CA: Sage Publications.
- Maldonado-Molina, M. M., & Lanza, S. T. (2010). A framework to examine gateway relations in drug use: An application of latent transition analysis. *Journal of Drug Issues, 40*, 901-924.
- Martin, R. A., Velicer, W. F., & Fava, J. L. (1996). Latent transition analysis to the stages of change for smoking cessation. *Addictive Behaviors, 21*, 67-80.
- McLachlan, G., & Peel, D. (2000). *Finite Mixture Models*. New York: Wiley.
- Mislevy, R. J., Levy, R., Kroopnick, M., & Rutstein, D. (2008). Evidentiary foundations of mixture item response theory models. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (pp. 149-175). Charlotte, NC: Information Age Publishing.
- Moustaki, I. (1996). A latent trait and a latent class model for mixed observed variables. *British Journal of Mathematical and Statistical Psychology, 49*, 313-334.
- Muthen, B., & Asparouhov, T. (2009). Multilevel regression mixture analysis. *Journal of the Royal Statistical Society, Series A, 172*, 639-657.

Muthén, L.K., & Muthén, B.O. (1998-2010). *Mplus user's guide* (6<sup>th</sup> ed.). Los Angeles, CA:

Muthén & Muthén.

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

Nagin, D. (2005). *Group-based modeling of development*. Cambridge: Harvard University Press.

Nagin, D., & Tremblay, R. E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development*, *70*, 1181-1196.

Nylund, K. L., Asparouhov, T., & Muthen, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling*, *14*, 535-569.

Parker, G., Wilhelm, K., Mitchell, P., Roy, K., & Hadzi-Pavlovic, D. (1999). Subtyping depression: Testing algorithms and identification of a tiered model. *The Journal of Nervous and Mental Disease*, *187*, 610-617.

Pfeffermann, D., Skinner, C., & Humphreys, K. (1998). The estimation of gross flows in the presence of measurement error using auxiliary variables. *Journal of the Royal Statistical Society, Series A*, *161*, 13-32.

Piaget, J. (1973). *The child and reality*. New York: Grossman.

- Posner, S. F., Collins, L. M., Longshore, D., & Anglin, M. D. (1996). The acquisition and maintenance of safer sexual behaviors among injection drug users. *Substance Use and Misuse, 31*, 1995-2015.
- Qu, Y., Tan, M., & Kutner, M. H. (1996). Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics, 52*, 797-810.
- Raudenbush, S. W., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2<sup>nd</sup> Ed.). Newbury Park, CA: Sage.
- Reboussin, B. A., Reboussin, D. M., Liang, K. Y., & Anthony, J. C. (1998). A latent transition approach for modeling progression of health-risk behavior. *Multivariate Behavioral Research, 33*, 457-478.
- Rindskopf, D. (1983). A general framework for using latent class analysis to test hierarchical and nonhierarchical learning models. *Psychometrika, 48*, 85-97.
- Rindskopf, D. (2006). Heavy alcohol use in the "Fighting Back" survey sample: Separating individual and community level influences using multilevel latent class analysis. *Journal of Drug Issues, 36*(2), 441-462.
- Rubin, D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *The Annals of Statistics, 12*, 1151-1172.
- Rubin, D. B., & Schenker, N. (1987). Logit-based interval estimation for binomial data using the Jeffreys prior. *Sociological Methodology, 17*, 131-144.

- Rubin, D. B., & Stern, H. S. (1994). Testing in latent class models using a posterior predictive check distribution. In Von Eye, A., & Clogg, C. C. (Eds.), *Latent variable analysis: Applications for developmental research*. Thousand Oaks, CA: Sage.
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph (Suppl. No. 17)*, 100.
- Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147-177.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.
- Sclove, L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, 52, 333-343.
- Spiel, C. (1994). Latent trait models for measuring change. In A. von Eye, & C. C. Clogg (Eds.), *Latent variable analysis: Applications for developmental research*. Thousand Oaks, CA: Sage.
- Stern, H. S., Arcus, D., Kagan, J., Rubin, D. B., & Snidman, N. (1995). Using mixture models in temperament research. *International Journal of Behavioral Development*, 18, 407-423.
- Sullivan, P. F., Kessler, R. C., & Kendler, K. S. (1998). Latent class analysis of lifetime depressive symptoms in the national comorbidity survey. *American Journal of Psychiatry*, 155, 1398-1406.

- Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82, 528-540.
- Tracy, A.J., Collins, L.M., & Graham, J.W. (1997). *Exposure to adult substance use as a risk factor in adolescent substance use onset* (Tech. Rep. No. 97-13). University Park: The Methodology Center, Penn State.
- van der Heijden, Dessens, J. & Bockenholt, U. (1996). Estimating the concomitant-variable latent-class model with the EM algorithm. *Journal of Educational and Behavioral Statistics*, 21, 215-229.
- Van Horn, M. L., Fagan, A. A., Jaki, T., Brown, E. C., Hawkins, D., Arthur, M. W.,..., Catalano, R. F. (2008). Using multilevel mixtures to evaluate intervention effects in group randomized trials. *Multivariate Behavioral Research*, 43, 289-326.
- Velicer, W. F., Martin, R. A., & Collins, L. M. (1996). Latent transition analysis for longitudinal data. *Addiction*, 91, S197-S209.
- Vermunt, J. K. (1997a). *LEM: A general program for the analysis of categorical data*. Department of Methodology and Statistics, Tilburg University.
- Vermunt, J. K. (1997b). *Log-linear models for event histories*. Thousand Oaks, CA: Sage.
- Vermunt, J. K. (2003). Multilevel latent class models. *Sociological Methodology*, 33(1), 213-239.

Vermunt, J. K. (2008). Latent class and finite mixture models for multilevel datasets. *Statistical Methods in Medical Research, 17*(1), 35-51.

Vermunt, J. K., and Magidson, J. (2005). *Latent GOLD 4.0 user's guide*. Belmont, MA: Statistical Innovations.

Wang, Brown, & Bandeen-Roche, 2005 Wang, C. P., Brown, C. H., & Bandeen-Roche, K. (2005). Residual diagnostics for growth mixture models. *Journal of the American Statistical Association, 100*, 1054–1076.

Witkiewitz, K., & Villarroel, N. A. (2009). Dynamic association between alcohol lapses and negative affect following alcohol treatment. *Journal of Consulting and Clinical Psychology, 77*(4), 633-644.

Table 1

*Frequency Distributions for Eight Observed Variables from the Add Health Feelings Scale:  
Percent of Students who Reported Symptom Sometimes/A lot/Most/All of the time During the  
Past Week*

	Time 1: 11 <sup>th</sup> -Grade Frequency (Valid %)		Time 2: 12 <sup>th</sup> -Grade Frequency (Valid %)	
	Male	Female	Male	Female
<i>Sad Indicators</i>				
Could not shake blues	147 (27.3)	191 (37.8)	110 (24.1)	165 (37.2)
Felt depressed	197 (36.6)	248 (49.1)	153 (33.6)	200 (45.1)
Felt lonely	187 (34.8)	222 (44.0)	148 (32.5)	184 (41.4)
Felt sad	230 (42.8)	286 (56.6)	175 (38.4)	238 (53.6)
<i>Failure Indicators</i>				
Thought life was a failure	74 (13.8)	74 (14.7)	62 (13.6)	75 (16.9)
Felt life not worth living	53 (9.9)	56 (11.1)	40 (8.8)	48 (10.8)
<i>Disliked Indicators</i>				
People unfriendly to you	178 (33.1)	179 (35.5)	154 (33.8)	150 33.78
People disliked you	163 (30.3)	185 (36.7)	137 (30.0)	146 (32.9)

*Note:* All indicators were coded 1 = never/rarely during the past week, 2 = sometimes/a lot/most/all of the time during the past week.

*Note:*  $N = 1043$  at Time 1,  $N = 900$  at Time 2.

Table 2

*Goodness-of-Fit Criteria for Various Latent Class Models at Time 1 (11<sup>th</sup> Grade)*

Number of Classes	Degrees of Freedom	$G^2$	AIC	BIC	CAIC	a-BIC	BLRT	Entropy $R^2$	Solution %
1	247	1837.5	1853.5	1893.1	1901.1	1867.7	---	1.0	100
2	238	557.6	591.6	675.7	692.7	621.7	.001	.78	100
3	229	400.4	452.4	581.1	607.1	498.5	.001	.73	35.3
4	220	286.4	356.4	529.7	564.7	418.5	.001	.72	99.4
<b>5</b>	<b>211</b>	<b>240.6</b>	<b>328.6</b>	<b>546.4</b>	<b>590.4</b>	<b>406.7</b>	<b>.001</b>	<b>.76</b>	<b>47.7</b>
6	202	210.6	316.6	578.9	631.9	410.6	.007	.77	14.7
7	193	179.2	303.2	610.1	672.1	413.2	.002	.76	25.0
8	184	158.6	300.6	652.1	723.1	426.6	.102	.75	34.9

*Note:* Solution % is the percentage of times solution was selected out of 1000 random sets of starting values. Dashes indicate criterion was not calculated for the model. Bold indicates selected model.

Table 3

*Parameter Estimates and Standard Errors for Latent Class Model at Time 1: Latent Classes of Depression in 11<sup>th</sup> Grade*

	1 Non- depressed	2 Sad	3 Disliked	4 Sad + Disliked	5 Depressed
<i>Latent Class Prevalences</i>					
	.39 (.04)	.23 (.02)	.17 (.04)	.15 (.03)	.07 (.02)
Observed Variable	<i>Item-Response Probabilities<sup>a</sup></i>				
Could not shake blues	.03 (.01)	<b>.54</b> (.05)	.17 (.06)	<b>.66</b> (.06)	<b>.90</b> (.05)
Felt depressed	.06 (.02)	<b>.73</b> (.05)	.24 (.07)	<b>.86</b> (.06)	<b>1.00</b> (.00)
Felt lonely	.07 (.02)	<b>.58</b> (.04)	.33 (.08)	<b>.77</b> (.05)	<b>.92</b> (.05)
Felt sad	.14 (.03)	<b>.80</b> (.04)	.38 (.08)	<b>.87</b> (.04)	<b>.94</b> (.04)
Thought life was a failure	.01 (.01)	.11 (.03)	.10 (.04)	.23 (.06)	<b>.90</b> (.09)
Felt life not worth living	.00 (.01)	.06 (.02)	.11 (.04)	.10 (.06)	<b>.85</b> (.15)
People unfriendly to you	.13 (.03)	.17 (.04)	<b>.64</b> (.08)	<b>.67</b> (.04)	<b>.70</b> (.06)
People disliked you	.04 (.03)	.00 (.00)	<b>.68</b> (.15)	<b>1.00</b> (.00)	<b>.77</b> (.06)

<sup>a</sup> Probability of reporting symptom “sometimes/a lot/most/all of the time” during the past week.

*Note:* Item-response probabilities greater than .50 appear in bold to facilitate interpretation.

Table 4

*Parameter Estimates and Standard Errors by Gender for Time 1 (11<sup>th</sup> Grade) Five-Class Latent Class Model, Item-Response Probabilities Constrained Equal Across Gender*

Gender	1 Non- Depressed*	2 Sad*	3 Disliked	4 Sad + Disliked*	5 Depressed
<i>Latent Class Prevalences</i>					
Male	.44 (.04)	.18 (.03)	.20 (.05)	.12 (.03)	.06 (.02)
Female	.30 (.04)	.25 (.05)	.16 (.05)	.22 (.04)	.06 (.02)
Observed Variable	<i>Item-Response Probabilities<sup>a</sup></i>				
Could not shake blues	.03 (.01)	<b>.54</b> (.05)	.14 (.05)	<b>.66</b> (.06)	<b>.89</b> (.05)
Felt depressed	.05 (.02)	<b>.74</b> (.05)	.26 (.07)	<b>.83</b> (.05)	<b>1.00</b> (.01)
Felt lonely	.07 (.02)	<b>.57</b> (.04)	.30 (.07)	<b>.79</b> (.05)	<b>.92</b> (.05)
Felt sad	.13 (.03)	<b>.80</b> (.04)	.37 (.07)	<b>.87</b> (.04)	<b>.94</b> (.04)
Thought life was a failure	.02 (.01)	.10 (.03)	.08 (.04)	.24 (.06)	<b>.93</b> (.09)
Felt life not worth living	.00 (.01)	.05 (.02)	.10 (.03)	.12 (.05)	<b>.85</b> (.13)
People unfriendly to you	.11 (.03)	.13 (.07)	<b>.62</b> (.08)	<b>.68</b> (.06)	<b>.69</b> (.07)
People disliked you	.04 (.03)	.02 (.07)	<b>.61</b> (.12)	<b>.93</b> (.10)	<b>.77</b> (.06)

\* Hypothesis test for gender difference in latent class prevalence rates statistically significant at the .05 level.

<sup>a</sup> Probability of reporting symptom “sometimes/a lot/most/all of the time” during the past week

*Note:* Item-response probabilities greater than .50 appear in bold to facilitate interpretation.

Table 5

*Parameter Estimates and Odds Ratios for Effect of Average Grade on 11<sup>th</sup> Grade Latent Class Membership*

	1 Non- depressed	2 Sad	3 Disliked	4 Sad + Disliked	5 Depressed
Beta (SE)	---	-.19 (.14)	-.53 (.18)	-.07 (.20)	-.78 (.20)
Odds Ratio (CI)	---	.82 [.63, 1.08]	.59 [.41, .84]	.93 [.63, 1.38]	.46 [.31, .68]
Inverse Odds Ratio	---	1.22 [.93, 1.59]	1.69 [1.19, 2.44]	1.08 [.72, 1.59]	2.17 [1.47, 3.23]

*Note:* Dashes indicate reference latent class. Average grade was a significant predictor of latent class membership ( $p < .01$ ).

*Note:*  $N = 1035$ .

Table 6

*Parameter Estimates for Time 1 (11<sup>th</sup> Grade) to Time 2 (12<sup>th</sup> Grade) Five-Class Latent**Transition Model, Item-Response Probabilities Constrained Equal Across Time*

		1	2	3	4	5
		Non- Depressed	Sad	Disliked	Sad + Disliked	Depressed
<i>Latent Class Prevalences</i>						
Time 1:	11 <sup>th</sup> grade	.32	.26	.19	.14	.08
Time 2:	12 <sup>th</sup> grade	.37	.23	.18	.12	.10
<i>Item-response Probabilities: Probability of Reporting Symptom "Sometimes/A lot/Most/All of the time" During the Past Week</i>						
Observed Variable						
Could not shake blues		.03	<b>.54</b>	.09	<b>.60</b>	<b>.85</b>
Felt depressed		.05	<b>.71</b>	.13	<b>.80</b>	<b>.96</b>
Felt lonely		.06	<b>.55</b>	.23	<b>.78</b>	<b>.85</b>
Felt sad		.11	<b>.78</b>	.26	<b>.87</b>	<b>.92</b>
Thought life was a failure		.01	.12	.05	.19	<b>.88</b>
Felt life not worth living		.01	.04	.05	.09	<b>.77</b>
People unfriendly to you		.07	.17	<b>.60</b>	<b>.81</b>	<b>.63</b>
People disliked you		.00	.11	<b>.55</b>	<b>.96</b>	<b>.77</b>
<i>Transition Probabilities</i>						
Time 1 Latent Class		Time 2 Latent Class				
Non-depressed		<b>.75</b>	.13	.09	.02	.01
Sad		.32	<b>.46</b>	.08	.09	.06
Disliked		.25	.05	<b>.50</b>	.15	.06
Sad + Disliked		.04	.24	.21	<b>.39</b>	.12
Depressed		.00	.21	.04	.09	<b>.65</b>

*Note:* Item-response probabilities greater than .50 and transition probabilities on the diagonal of the matrix appear in bold to facilitate interpretation.

*Note:* N=1044.

Table 7

*Parameter Estimates by Gender for Time 1 (11<sup>th</sup> Grade) to Time 2 (12<sup>th</sup> Grade) Five-Class**Latent Transition Model*

		1	2	3	4	5
		Non- depressed	Sad	Disliked	Sad + Disliked	Depressed
<i>Latent Class Prevalences</i>						
Time 1:	Male	.37	.21	.23	.10	.09
	Female	.26	.29	.17	.19	.10
Time 2:	Male	.39	.16	.27	.08	.09
	Female	.35	.25	.12	.17	.12
<i>Transition Probabilities</i>						
Time 1 Latent Class		Time 2 Latent Class				
		Male				
Non-Depressed		<b>.72</b>	.13	.13	.00	.02
Sad		.33	<b>.42</b>	.16	.07	.02
Disliked		.22	.02	<b>.60</b>	.12	.04
Sad + Disliked		.03	.24	.39	<b>.29</b>	.05
Depressed		.05	.00	.04	.12	<b>.79</b>
		Female				
Non-Depressed		<b>.79</b>	.13	.05	.04	.00
Sad		.32	<b>.45</b>	.04	.13	.06
Disliked		.27	.09	<b>.38</b>	.19	.06
Sad + Disliked		.03	.23	.11	<b>.43</b>	.20
Depressed		.00	.31	.09	.07	<b>.54</b>

*Note:* Measurement invariance across time (i.e., Time 1 and Time 2) and across groups (i.e., male and female) imposed.

*Note:* Transition probabilities on the diagonal of the matrix appear in bold to facilitate interpretation.

Figure 1

*Latent Variable Frameworks*

		Latent Variables	
		Categorical	Continuous
Observed Variables	Categorical	Latent Class Analysis	Latent Trait Analysis
	Continuous	Latent Profile Analysis	Covariance Structure Analysis

Figure 2

*Stages of Change Construct from the Transtheoretical Model*

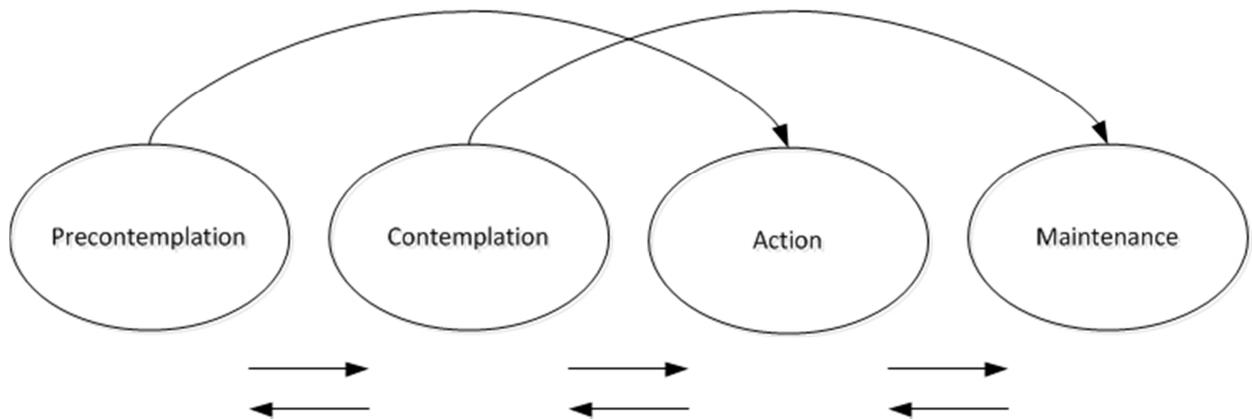


Figure 3

*Probability of Reporting Symptom Sometimes/A Lot/Most/All of The Time Conditional on Depression Latent Class*

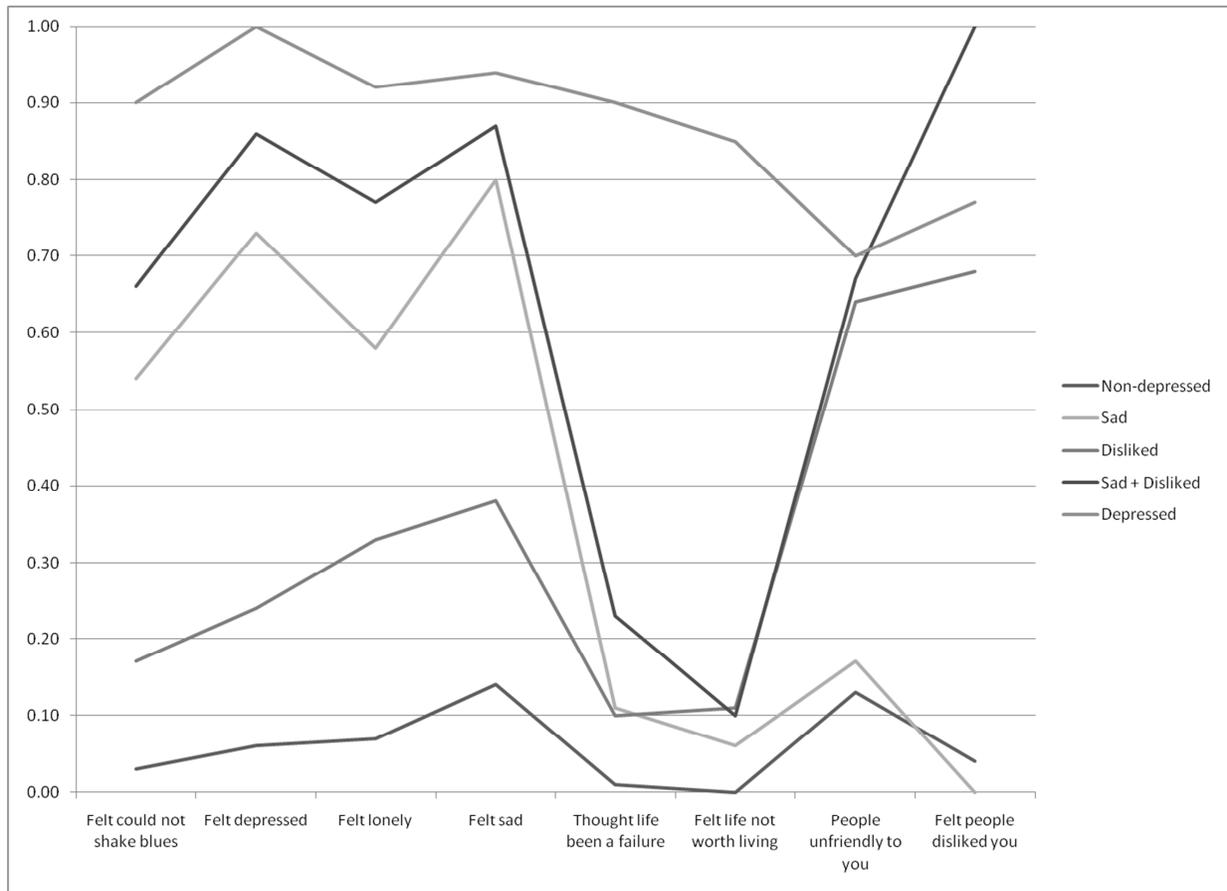


Figure 4

*Probability of Membership in Each Depression Latent Class by Gender*

