The Pennsylvania State University
The Graduate School

EXAMINING GAMBLING AND SUBSTANCE USE: APPLICATIONS OF ADVANCED LATENT CLASS MODELING TECHNIQUES FOR CROSS-SECTIONAL AND LONGITUDINAL DATA

A Thesis in
Human Development and Family Studies
by
Bethany Cara Bray

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2007

The thesis of Bethany Cara Bray was reviewed and approved* by the following:

Linda M. Collins
Professor of Human Development and Family Studies
Thesis Advisor, Chair of Committee

Edward A. Smith
Senior Research Associate in the College of Health and Human Development

Jennifer L. Maggs
Associate Professor of Human Development and Family Studies

Joseph L. Schafer
Associate Professor of Statistics

Stephanie T. Lanza
Scientific Director and Research Associate, The Methodology Center
Special Member

Helene R. White
Professor of Sociology, Rutgers University
Special Signatory

Douglas M. Teti
Professor of Human Development and Family Studies
Graduate Professor In Charge

*Signatures are on file in the Graduate School.

Abstract

The purpose of the current project is to present three empirical studies that illustrate the application of advanced latent class modeling techniques for cross-sectional and longitudinal data to research questions about gambling and substance use. The first empirical study used latent class analysis and conditional latent class analysis to identify and predict types of college-student gamblers using data from a large northeastern university. Four types of gamblers were identified for men and women: non-gamblers, cards and lotto players, cards and games of skill players, and multi-game players. There were substantial gender differences in the latent class membership probabilities: (1) men were most likely to be cards and lotto players whereas women were most likely to be non-gamblers; and (2) men were more likely than women to be cards and games of skill and multi-game players, and less likely to be non-gamblers. Significant predictors of gambling latent class membership included: school year, living in off-campus housing, Greek membership, and past-year alcohol use. There were substantial gender differences in the predictive effects of Greek membership and past-year alcohol use: (1) the effects of Greek membership were in different directions for men and women; and (2) past-year alcohol use was more strongly related to gambling latent class membership for women.

The second empirical study used latent class analysis to identify types of adolescent and young adult gamblers and used latent class analysis for repeated measures to identify types of drinking trajectories using data from the National Longitudinal Study of Adolescent Health. Multivariable latent class modeling was used to examine the relation between gambling and drinking by linking specific types of gambling to specific types of drinking trajectories. Gambling and drinking were shown to be highly related: (1) consistent infrequent, light, or not intense drinkers were most likely to be non-gamblers; and (2) participants who were frequent, heavy, or intense drinkers at any time were most likely to gamble in all activities. Overall, drinking frequency appeared to be more predictive of gambling than was drinking quantity.

The third empirical study used latent transition analysis to identify types of adolescent smokers and types of drinkers, and to describe smoking and drinking development over time using data from the National Longitudinal Survey of Youth 1997. Multiprocess modeling was used to examine the relation between smoking and drinking by modeling the development of smoking and the development of drinking simultaneously. Three types of smokers and three types of drinkers were identified: non-smokers, light smokers, heavy smokers, non-drinkers, light drinkers, and heavy drinkers. The majority of participants were non-smokers and non-drinkers. The behavior of non-smokers, non-drinkers, heavy smokers, and heavy drinkers was relatively stable across time whereas the behavior of light smokers and light drinkers was variable. Linking smoking and drinking showed that: (1) knowing type of smoking provided limited information about type of drinking; (2) transitioning from non-drinking to heavy drinking was progressively more likely for more serious types of smoking; (3) transitioning from heavy drinking to non-drinking was progressively less likely for more serious types of smoking; and (4) transitioning from light drinking to non-drinking was most likely for non-smokers whereas transitioning from light drinking to heavy drinking was most likely for heavy smokers.

Table of Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I would like to thank my mentor, Dr. Linda M. Collins, for the time and energy that she has invested in my education and career and in our relationship over the past five years. Without her, none of this would have been possible. Her vision, guidance, and unfailing support provided the foundation on which I was able to begin building my research career. They continue to provide the foundation on which I am standing, looking ahead to a life of doing what I love.

I would also like to sincerely thank Dr. Stephanie T. Lanza, my mentor and friend, who worked closely with me on this project. Her encouragement and excitement about this area of research has kept me motivated during the past year, and I look forward to continuing to work with her in this area.

I would like to thank my entire committee for their outstanding training and mentorship through out my time here at Penn State. But, I would most like to thank Drs. Edward A. Smith, Joseph L. Schafer, and Jennifer L. Maggs for their belief in my potential. And, I would like to thank Dr. Helene R. White for providing me the opportunity to pursue my passion for gambling research by providing the data used in the first empirical study.

For their friendship, laughter, and good advice I would like to thank the members of our latent class project group over the years: Dr. Mildred Maldonado-Molina, Dr. Zhiqun Tang, Tammy Root, Karen Auerbach, Amanda Gottschall, and Dr. Michael Cleveland. I would like to extend a most special thank you to Mildred for always being there for me and for helping me navigate troubled waters.

I would also like to extend a most special thank you to Tammy for her friendship and unwavering support, for helping me out of so many sticky situations, and for being the sister I never knew I had.

For making the last five years so much easier than they should have been, and for always having a smile and an empty office chair, I would like to thank Tina Meyers. For helping me keep my life under control, I would like to thank Lisa Litz. I would like to thank Robert Lambert and Dr. David Lemmon for their constant technical support and help. Certainly, this project would not have been possible without you.

The Methodology Center provides the most supportive, engaging, and collaborative atmosphere one can imagine. My research would not be nearly so rich or fulfilling without the resources offered by The Methodology Center. Thus, thank you to the National Institute of Drug Abuse for funding both my center and my fellowship, without which my career would not be half so productive. My research is funded by NIDA Center Grant P50-DA-10075, and by NIDA Training Grant T32-DA-017629.

Thank you to my previous mentors who started me on my journey. Thomas Cotner, thank you for your enthusiasm, humor, mentorship, and friendship. But mostly, thank you for taking an interest in me all those years ago — your passion for teaching mathematics made me love it. Dr. John Gillespie, thank you for your passion for statistics and for your guidance in my research career, without you I would not be where I am today. But mostly, thank you for acting in *A Midsummer's Night's Dream* — your never-ending support, levity, and love of your students was never so clear as that night. You made me realize the kind of mentor I want to be. Dr. Susan Murphy, thank you so much for believing in me and for knowing that an academic career was right for me, despite all evidence to the contrary. Thank you for helping me to follow my heart, and for sending me to Linda.

To one of my oldest friends, Jonathan Finnegan, thank you for being a part of my life for the last 15 years. Who knew that my very first day of high school would bring me one of my closest friends and confidants. To Lori Chapman, thank you for always being there for me and for your continual support in life and in school. You are an amazing mathematician, thank you for helping me be one, too. To Dr. Martha Bleeker, thank you for being my friend, for taking care of me, and for laughing with me. You helped me remember how to not take myself so seriously, and showed me how to have fun with my research. To Brandon Hayes, thank you for being my best friend. The last 10 years have been filled with laughter and love because of you. Thank you, for everything. It is not possible to thank you all enough; this is your accomplishment as much as it is mine. Thank you for taking this journey with me.

Thank you to my extended family for their love and support over the last year and half. Thank you to Dr. James King, Jayne Cleveland, Patrick King, and Alex King for being there for me and cheering me on. Thank you, too, for making me laugh and for letting me nap when I visit. Most especially, thank you for Dave.

Thank you to Dave King for loving me and supporting me for the last year and a half. I am so lucky to have found you; you light my life. Thank you for your encouragement and for standing by me during the hard times. (Thank you, too, for putting up with me during those times.) Thank you for your friendship and your laughter. Thank you for helping me reach my dreams.

Finally, I want to most sincerely thank my family. They have been standing behind me, helping me reach my goals, all my life. They have been helping me make this happen since before I knew this is what I wanted. They have always believed in me. Thank you to my aunts and uncles: Mineko Bray, Philip Bray, Karen Bray, Marc Bray, Janice Hartley, and Kevin Hartley. I would like to especially thank my Aunt Mineko for always being so proud of me. Thank you to my grandmothers, Dolores Hartley and Francis Bray, for everything. I would like to especially thank my Grannie for all of the graduate school care packages. I would like to especially thank my Grams for all of the financial care packages. Thank you to my brother, Brandon Bray, for always being there for me and supporting me. To my mom and dad, John and Paula Bray, thank you so much for everything, always.

Dedication

For those teachers who helped this student realize how the infinite mystery of mathematics can be a part of everyone's life:

John Bray
Fred Palmer
Tom Cotner
John Gillespie
Susan Murphy
Linda Collins

CHAPTER 1

Introduction

1.1 Overview

The purpose of the current project is to demonstrate the use of loglinear models with latent variables to address developmental research questions. Loglinear modeling with latent variables provides a general framework for modeling relations among discrete latent variables in a variety of ways. These types of models provide the flexibility necessary for addressing complex research questions about development. Latent class and latent transition models, discussed in more detail below, are two examples of loglinear modeling with latent variables. The focus of the current project, however, is on more complex models like conditional latent class, multivariable latent class, and multiprocess models, which are extensions of latent class and latent transition models.

1.1.1 Discrete Latent Variables in the Study of Development

There has been a long history of latent class models in the study of development. Many developmental phenomena have been profitably conceptualized as discrete latent variables (*i.e.*, discrete developmental processes). These include substance use (Collins & Wugalter, 1992; Collins et al., 2000; Reboussin & Anthony, 2001), temperament (Stern et al., 1995; Loken, 2004), teaching style (Aitkin et al., 1981), delinquency (Fergusson et al., 1994), sexual behavior (Tang, 2002), attitudes toward abortion (Chung, 2003), health-risk behavior (Reboussin et al., 1998; Martin et al., 1996; Velicer et al., 1996), eating disorders (Bulik et al., 2000), depression (Sullivan et al., 1998; Parker et al., 1999), and evaluation of the effectiveness of prevention programs for substance use and other problem behaviors (Collins et al., 1994), to name a few. In addition, latent class modeling can include covariates. For example, latent class modeling of substance use has included a number of risk factors as covariates, such as heavy caffeine use (Collins et al., 1997), pubertal timing (Lanza & Collins, 2002), and parental permissiveness (Hyatt & Collins, 2000). Latent class models not only provide an intuitive way to talk about development, but the identification of types of individuals may provide a way for researchers to design prevention and treatment programs to arrest the development of problem behavior or psychopathology.

The current project focuses on two areas of particular importance to developmentalists interested in prevention and treatment: substance use and gambling. The field of substance use research has particularly benefited from latent class models. Using latent class and latent transition models, researchers have found that a general pattern of substance use onset includes the following stages: no use, alcohol and/or tobacco use, and then use of marijuana and/or other illegal substances (Collins & Wugalter, 1992; Collins et al., 1997; Hyatt & Collins, 2000).

Gambling has yet to be conceptualized and analyzed as a discrete latent variable, but current beliefs about gambling suggest that conceptualizing it in this way will be beneficial to the field. For example, gambling and problems with gambling are almost universally treated as discrete by researchers already. Individuals are described as being particular types of gamblers: non-gamblers or non-problem gamblers, problem or at-risk gamblers, or pathological gamblers. This typology of gambling, however, relies on the clinical definition of the most severe form of gambling problem, pathological gambling. The clinical definition of pathological gambling is based on a sum of the number of endorsed diagnostic criteria. Because of this definition, it is possible to conceptualize gambling behavior as lying along a continuum of non-problem to pathological gambling behavior with a (more or less) linear escalation in gambling related problems between the two ends of the spectrum. The types of gamblers, then, are created by dividing this continuum into discrete categories.

Current thinking about gambling, however, is moving toward a more complex, multidimensional conceptualization of gambling and problems with gambling (Shaffer et al., 2004). For example, it is likely that among individuals traditionally identified as problem or pathological gamblers there are identifiable groups who share similar motivations, game and venue preference, and/or profiles of endorsed diagnostic criteria. Identifying groups based on similar patterns of behavioral or other characteristics may suggest different etiological processes at work, or may provide different implications for prevention and treatment. It seems, then, that conceptualizing gambling as a discrete latent variable and using appropriate methods to model gambling behavior may be particularly useful. In so doing, differences that may play an important role in the etiology and prevention of the development of problem and pathological gambling may be discovered.

1.1.2 Outline of the Current Discussion

This chapter provides an introduction to loglinear modeling with latent variables. Traditional loglinear modeling with manifest variables is discussed, followed by a brief general introduction to latent variable modeling. Then, a discussion of how loglinear modeling and latent variable modeling have been merged to model discrete latent variables is presented. A brief conceptual and statistical examination of five types of loglinear models with latent variables is provided to illuminate the variety of research questions that this approach may be used to address. The types include: latent class modeling, conditional latent class modeling, latent transition modeling, multivariable latent class modeling, and multiprocess modeling. It is important to note upfront that discussion is confined to discrete latent variables measured with discrete indicators, in order to limit the scope of this project.

In the chapters to follow, three empirical studies are presented. A brief substantive literature review proceeds each empirical study. These studies address developmental research questions about gambling and substance use empirically using the discussed modeling approaches. The empirical studies illustrate the practical application of sophisticated discrete latent variable models to concrete research questions about development. The final chapter provides an overall discussion of the current project as a whole.

1.2 Two Modeling Traditions

Two traditions of modeling variables are relevant to the current project. The first tradition, loglinear modeling, describes how contingency table cell counts depend on levels of discrete variables, and how the discrete variables are associated and interact (Agresti, 2002). The second tradition, latent variable modeling, is used to model constructs that are not directly observable. In latent variable models, a set of manifest indicators of the latent variable and a mathematical model are used to infer information about the latent variable. The discussion to follow briefly introduces loglinear modeling and latent variable modeling.

1.2.1 Loglinear Modeling

Loglinear models may be used to understand the ways in which discrete variables are associated and the ways in which they interact. Loglinear models with two, three, and four variables are briefly discussed below. For reference, more complete discussions of loglinear modeling may be found in many classic texts on categorical data analysis, including Agresti (2002) and Bishop et al. (1975). A discussion of loglinear modeling in developmental research may be found in von Eye and Clogg (1996; Part 3).

Consider a question on a survey asking, “What is your gender?” that has two response options “male” and “female.” In addition, consider a second question on the same survey asking, “Have you ever gambled?” that has three response options "yes," "no," and "I do not know." Cross-classifying respondents based on their responses to these two survey questions produces an I x J contingency table where $I = 2$ and $J = 3$. Respondents are classified based on their responses to two variables and the resulting contingency table is typically called a two-way table. In this example, the contingency table contains 6 cells; the proportion of respondents in each cell is denoted by π_{ij}, and the observed cell count is denoted by n_{ij}.

Now, consider the research question, "Are gender and gambling related?" This question may be addressed using the information contained in the 2 x 3 contingency table discussed above. When two variables are unrelated they are said to be independent; when two variables are related they are said to be dependent. That is, gender and gambling are independent when knowing an individual's gender provides no information about his or her response to the gambling question.

Typically, traditional hypothesis testing is used to test whether two variables are independent. In this case, the null hypothesis states that a model specifying independence between gender and gambling fits the observed data well. The alternative hypothesis states that a model specifying dependence between gender and gambling fits the observed data significantly better than the model specifying independence. What is being tested is whether the independence model should be rejected in favor of the dependence model. The null hypothesis may be tested using the X^2 or G^2 goodness-of-fit statistics. The goodness-of-fit statistics are

$$X^2 = \sum_{i=1}^{2}\sum_{j=1}^{3} \frac{(n_{ij} - \mu_{ij})^2}{\mu_{ij}}, \tag{1.1}$$

and

$$G^2 = 2\sum_{i=1}^{2}\sum_{j=1}^{3} n_{ij} \log \frac{n_{ij}}{\mu_{ij}}, \tag{1.2}$$

respectively, where $\mu_{ij} = n\pi_{ij}$ denotes the expected cell count based on the model assumed by the null hypothesis and n is the total sample size.

Both the X^2 and G^2 goodness-of-fit statistics are asymptotically distributed as a χ^2 with *degrees of freedom* (*df*) equal to $(I - 1)(J - 1)$. By comparing the X^2 or G^2 to the χ^2 distribution, the p-value for the test of independence between gender and gambling is obtained. A significant p-value results in the rejection of the null hypothesis in favor of the alternative hypothesis that the model of dependence between gender and gambling fits significantly better than the model of independence. Using this method it is possible to make conclusions about the relation between gender and gambling. This method works well only when the contingency table is not overly sparse; if the contingency table is sparse, the χ^2 distribution is not a good approximation to the null distribution of X^2 or G^2, and these statistics should not be used to asses goodness-of-fit.

If gender and gambling are independent, $\mu_{ij} = n\pi_{i+}\pi_{+j}$, where π_{i+} is the marginal probability of a level of i over all levels of j ($\pi_{i+} = \sum_{j=1}^{3} \pi_{ij}$), and π_{+j} is the marginal probability of a level of j over all levels of i ($\pi_{+j} = \sum_{i=1}^{2} \pi_{ij}$). Another way of thinking about independence, which may be more familiar to some readers, is that when gender and gambling are independent

$$\pi_{j|i} = \pi_{+j}, \tag{1.3}$$

$$\pi_{i|j} = \pi_{i+}, \tag{1.4}$$

and

$$\pi_{ij} = \pi_{i+}\pi_{+j}. \tag{1.5}$$

These expressions are direct results of the discussion above. An alternative way to express the simple multiplicative relation between gender and gambling described by $\mu_{ij} = n\pi_{i+}\pi_{+j}$ and Equation 1.5 is the loglinear expression of the independence model

$$\log \mu_{ij} = \lambda + \lambda_i^X + \lambda_j^Y, \tag{1.6}$$

where: (1) gender is denoted by X and gambling is denoted by Y; (2) λ is an intercept and a normalizing constant to ensure $\sum_{i=1}^{2}\sum_{j=1}^{3} \mu_{ij} = n$, λ_i^X is an effect of gender, and λ_j^Y is an effect of gambling; and (3) model identifiability requires constraints such as $\lambda_I^X = \lambda_J^Y = 0$. By taking the log of μ_{ij}, the expression of the independence model has an additive form.

The loglinear expression of the dependence model has the form

$$\log \mu_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}, \tag{1.7}$$

where λ_{ij}^{XY} is an interactive effect of gender and gambling that reflects a deviation from independence. In a two-way table, the dependence model is the most general model, called the saturated model. In general, the saturated model fits the data perfectly and includes all possible multi-way interactions. The dependence model, Model 1.7, is hierarchical in that it includes all lower-order terms (λ_i^X and λ_j^Y) composed from variables in a higher-order term (λ_{ij}^{XY}; Agresti, 2002). Typically, whenever higher-order effects are included in a loglinear model, all lower-order effects must also be included.

There is a shorthand notation for loglinear models that makes writing the models particularly convenient. In loglinear notation, variables separated by a comma are independent or conditionally independent, and variables written together with no comma are dependent or conditionally dependent. Using this notation, the independence model discussed above may be written (X, Y), and the dependence model may be written (XY). A summary of the model names, loglinear notation, and effects included in the models for the two-variable models discussed above and three-variable models discussed below is presented in Table 1.1.

The methods discussed above may be extended to three-way tables. Consider an additional survey question asking, "How many times have you drunk five or more drinks in one sitting?" that has response options "none," "once or twice," "three or four," and "five or more." Cross-classifying respondents based on their responses to all three survey questions produces an I x J x K contingency table where $K = 4$. Respondents are classified based on their responses to three variables and the resulting contingency table becomes a three-way table. In this example, the contingency table contains 24 cells; the proportion of respondents in each cell is denoted by π_{ijk}, and the observed cell count is denoted by n_{ijk}. In a three-way table, the formula expressing the saturated model has the form

$$\log \mu_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ}, \tag{1.8}$$

where: (1) λ_{ij}^{Z} is an effect of binge drinking; (2) λ_{ik}^{XZ} is an interactive effect of gender and binge drinking; (3) λ_{jk}^{YZ} is an interactive effect of gambling and binge drinking; and (4) λ_{ijk}^{XYZ} is a multi-way interactive effect of gender, gambling, and binge drinking.

Table 1.1. Summary of Two- and Three-variable Hierarchical Loglinear Models

Two-variable Models	Loglinear Notation	Included Effects
Independence	(X, Y)	X, Y
Dependence	(XY)	X, Y, XY

Three-variable Models	Loglinear Notation	Included Effects
Complete Independence	(X, Y, Z)	X, Y, Z
One-factor Independence	(XY, Z)	X, Y, Z, XY
	(XZ, Y)	X, Y, Z, XZ
	(YZ, X)	X, Y, Z, YZ
Conditional Independence	(XY, XZ)	X, Y, Z, XY, XZ
	(XY, YZ)	X, Y, Z, XY, YZ
	(XZ, YZ)	X, Y, Z, XZ, YZ
Homogeneous Association	(XY, XZ, YZ)	X, Y, Z, XY, XZ, YZ
Saturated	(XYZ)	X, Y, Z, XY, XZ, YZ, XYZ

There are a total of nine possible hierarchical models with three variables, which range in complexity from the independence model to the saturated model.[1] A summary of the model names, loglinear notation, and effects included in the models is presented in Table 1.1. In one-factor independence models, two variables are jointly independent of the third. For example, it may be hypothesized that gender and binge drinking are related, but that neither is related to gambling. This hypothesis corresponds to the one-factor independence model (XZ, Y). In conditional independence models, two variables are related only through their mutual association with the third. For example, it may be hypothesized that gender and binge drinking are related, and that gender and gambling are related, but that gambling and binge drinking are related only through their mutual association with gender. This hypothesis corresponds to the conditional independence model (XY, XZ). In the model of homogeneous association, the conditional relation between any two variables given the third is the same for each level of the third. For example, it may be hypothesized that all three variables are related but that a three-way interaction among gender, gambling, and binge drinking is not necessary to describe the observed relations among the variables. This hypothesis corresponds to the homogeneous association model (XY, XZ, YZ). The homogeneous association model may be plausible, for example, if the conditional relation between gambling and binge drinking given gender is the same for men and women.

The methods discussed above may be extended to n-way tables. When it is of interest to model four or more discrete variables, however, the set of possible models increases exponentially. In four-variable models, the set of possible hierarchical loglinear models includes: the complete independence model, models with all possible combinations of two-way interactions, models with all possible combinations of three-way interactions, and the saturated model.

In practice, when addressing research questions about the relations among three or more variables, it is often neither desirable nor practical to fit all possible models. Instead, a subset of models that correspond to theoretically logical, hypothesized relations among the variables should be selected. This set of relevant models for consideration should contain a reasonable number of models that will likely range

[1]Non-hierarchical loglinear models are sometimes used, but rarely (Rindskopf, 1990).

in complexity. Models within the set are then fit and compared using likelihood ratio tests (when the models are nested) or the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), in order to determine which is best at balancing fit and parsimony in its description of the observed relations among the variables. Once a model is selected, it may be used to make substantive conclusions about the nature of the relations among the variables.

1.2.2 Latent Variable Modeling

Latent variable models are used to model constructs that are not directly observable, such as temperament, depression, or substance use. Latent variable models are based on sets of manifest indicators used to measure the latent variables of interest. The manifest indicators are used to obtain information about characteristics or levels of the latent variable(s) for an individual. The general philosophy of measurement employed by latent variable models is that each manifest indicator is fundamentally error-prone. For example, this philosophy recognizes that responses to survey questions asking about gambling activity participation will never be error-free indicators of gambling behavior. Multiple manifest indicators for a single latent variable are preferred because this allows the error associated with each manifest indicator to be estimated and removed from other parameter estimates. In latent variable models, information from the manifest indicators is used in combination with a mathematical model to infer information about a latent variable, or about the relations among latent variables.

Factor analysis may be the most well-known latent variable methodology. Factor analysis is usually used to model continuous latent variables. The goal of factor analysis is to determine the underlying latent structure of a set of manifest indicators. In developmental research, factor analysis is often used to reduce a large number of manifest indicators to a smaller number of factors, select from a larger set of manifest indicators a smaller set that are performing well, or validate scales. An extension of factor analysis well-known in developmental research is structural equation modeling (SEM). SEM may be used to relate latent constructs to each other, and to relate latent constructs to covariates. More complete discussions of factor analysis and SEM may be found in many articles and books, including

Bollen (1989), Kim and Mueller (1978a, 1978b) and Kline (1998).

1.3 Loglinear Modeling with Latent Variables

Many latent constructs of interest to developmental researchers are profitability conceptualized as discrete latent variables, as discussed above. Methods appropriate for continuous latent variables, like factor analysis and SEM, are not appropriate when modeling the relations among several discrete latent variables using discrete manifest indicators. Merging the loglinear and latent variable approaches provides a way to model discrete latent variables, and the relations among discrete latent variables. Loglinear modeling with latent variables combines the features of loglinear modeling and latent variable modeling in order to examine: (1) the ways in which manifest indicators are probabilistically related to discrete latent variables, and (2) the ways in which discrete latent variables are related. The former is often referred to as the measurement model, and the latter as the structural model.

The discussion presented here considers the application of loglinear models with one, two, and four latent variables to developmental research questions. More complete discussions of loglinear modeling with latent variables and their application to developmental research may be found in Hagenaars (1993), Heinen (1996), Everitt and Dunn (1988), von Eye and Clogg (1996), and Vermunt (1996).

In the discussion below, research questions motivating the use of each model stem from an interesting research area that has not yet been adequately addressed: the relation between gambling behavior and substance use. Both gambling and substance use may be profitably conceptualized as discrete latent variables, and there are a variety of research questions that may be asked about the structure of each latent construct and the relation between the latent constructs. In the discussion to follow, gambling is labeled G and substance use is labeled S. All example variables and models discussed are hypothetical; no empirical data are used in this chapter.

1.3.1 Latent Class Modeling

Consider the research question, "Are there identifiable types of gamblers?" This is a question about the underlying categorical structure of gambling; gambling is conceptualized as a discrete latent variable. To address this question, a survey including three questions with responses of "yes" and "no" may be used to measure motivations to gamble; the survey questions are the manifest indicators of the latent variable, type of gambling.

To infer information about the underlying categorical structure of gambling, a mathematical model is needed to link the manifest indicators to the latent variable. The mathematical model may be provided in the form of a loglinear model. In this example there are one unobserved variable and three observed variables. Loglinear models that include both a latent variable and two or more manifest indicators are called latent class models.

The mathematical model used in traditional latent class modeling is (XG, YG, ZG), where X, Y, and Z represent the three survey questions measuring motivations to gamble. In this model, the three manifest indicators are directly related to the latent variable but not to each other; the assumption that the manifest indicators are not related to each other except through their relation to the latent variable is known as the assumption of local independence.

Latent class models divide a population into a set of mutually exclusive and exhaustive discrete latent classes (Goodman, 1974; Clogg, 1995; Lanza et al., 2003; Lazarsfeld & Henry, 1968). Typically, indicators having discrete response options are used for latent class modeling. These variables can be cross-tabulated to form a contingency table. This contingency table serves much the same role in latent class analysis (LCA) as that of a correlation matrix in factor analysis. LCA is used to identify latent classes in the contingency table data. For example, using the method described above, LCA may be used to identify different types of gamblers in the set of observed patterns of responses to the survey questions about gambling motivations. Identified classes may often be interpreted as "types" of respondents, as the classes represent groups of individuals who responded to the questions in a similar way.

In LCA, two sets of parameters are estimated. One set, the latent class membership probabilities, are the unconditional probabilities of membership in the latent classes. These can be interpreted as the proportions of the population expected to belong to each latent class. Often referred to as the γ parameters, these comprise what is termed the structural part of the latent class model. The other set, the item-response probabilities, are the probabilities of giving particular responses to particular indicators conditional on latent class membership. Often referred to as the ρ parameters, these comprise what is termed the measurement part of the latent class model. The item-response probabilities describe the probabilistic relation between the manifest indicators and the latent variable. Additional details of the latent class model, including the original parameterization using conditional response probabilities, may be found in a variety of resources including Goodman (1974), Clogg (1995), and Lanza et al. (2007).

LCA is the "standard operating procedure" for modeling discrete latent variables. As discussed above, LCA has been applied to a myriad of research questions about development. All of the models discussed below are extensions of traditional LCA that allow developmental researchers to address a broader array of research questions. A summary of the modeling approaches for loglinear models with latent variables selected for discussion is shown in Table 1.2. This table includes information about the name of the method, acronym, study design (cross-sectional or longitudinal), and whether or not covariates are included. The names and acronyms presented in the table and used throughout, are conventions adopted for the current project, they are not universally recognized.

Once the underlying categorical structure of gambling is identified using LCA, more complex research questions may be addressed using more general latent class modeling approaches (Lanza et al., in press). For example, consider the research question, "Do the gambling latent class membership probabilities vary by gender?" A traditional latent class model with gender as a grouping variable may be used to address this question.

When a grouping variable like gender is included in the standard latent class model, the latent class membership probabilities and item-response probabilities can be conditioned on group membership. In this case, the probabilities of gambling latent class membership may be different for men and women; so too, may the item-response probabilities.

Table 1.2. Summary of Selected Modeling Approaches for Loglinear Models with Latent Variables

Acronym	Method	Design	Covariates
LCA	Latent Class Analysis	Cross-sectional	No
RMLCA	Latent Class Analysis for Repeated Measures	Longitudinal	No
CLCA	Conditional Latent Class Analysis	Cross-sectional	Yes
LTA	Latent Transition Analysis	Longitudinal	No
CLTA	Conditional Latent Transition Analysis	Longitudinal	Yes
MVLCM	Multivariable Latent Class Modeling	Cross-sectional	No
MPM	Multiprocess Modeling	Longitudinal	No

Note: The names and acronyms are conventions adopted for the current project, they are not universally recognized.

1.3.1.1 Latent Class Modeling for Repeated Measures

Typically, LCA uses manifest indicators that were all measured at the same time. In LCA for repeated measures (RMLCA), one or more indicators are drawn from multiple points in time. The identified classes may often be interpreted as "trajectory classes," as the classes represent groups of individuals who have similar trajectories or profiles of behavior over time. The technical and mathematical details of RMLCA are the same as those of LCA but the slightly different interpretation of the latent classes stems from the indicators included in the model. See Lanza and Collins (2006) for an example of RMLCA.

1.3.1.2 Conditional Latent Class Modeling

Next, consider the research question "Does income predict gambling latent class membership?" To address this question, manifest covariates may be included in the traditional latent class model (Chung et al., 2006). For the purposes of the current project, latent class models that include covariates are called conditional latent class models. In conditional latent class models, continuous or categorical manifest variables are added to the model as covariates to predict latent class membership.

In addition to the parameters discussed above, conditional LCA (CLCA) involves a third set of parameters: multinomial logistic regression coefficients quantifying the effect of the covariate on latent class membership. For example, in a latent class model including income as a covariate, the effect of income on gambling latent class membership is expressed by a set of odds ratios based on the logistic regression coefficients.

Covariates may be included in models with or without a grouping variable. When a grouping variable is simultaneously included in the model, the logistic regression coefficients are conditional on group membership and may vary across groups. Examples of possible applications and the details of the mathematical models for latent class models with a grouping variable and/or covariates may be found in a variety of resources including Chung et al. (2006) and Lanza et al. (in press).

1.3.2 Latent Transition Modeling

In addition to identifying and predicting types of gamblers, when longitudinal data are collected, research questions about developmental transitions in gambling behavior may be posed. Consider the research question, "How does gambling develop over time?" This can be framed as a question about change over time in gambling latent class membership.

If gambling is measured at two times, by three indicators at each time, there are two unobserved and six observed variables included in the loglinear model with latent variables; if gambling is measured at three times there are three unobserved and nine observed variables, and so on. Similar to the way in which the discussion above about loglinear models with manifest variables was extended from two-way to multi-way tables, loglinear modeling with latent variables provides a flexible framework that allows the one-latent-variable loglinear model to be easily extended to multiple discrete latent variables. Loglinear models that include one construct measured over time are called latent transition models; latent transition models describe and test models of longitudinal change within a developmental process.

In this case, to infer information about the underlying categorical structure of gambling and the way in which gambling at time t is related to gambling at time $t+1$, a mathematical model is needed to: (1) link the manifest indicators to the latent variable at each time, and (2) relate gambling at time t to gambling at time $t+1$. The mathematical model may again be provided in the form of a loglinear model. Loglinear modeling with multiple latent variables may be called by many different names depending on the nature of the research question it is being used to address. The mathematical model used in latent transition modeling is $(X_1G_1, Y_1G_1,\ Z_1G_1, X_2G_2, Y_2G_2, Z_2G_2, G_1G_2,)$, where the subscripts denote time; local independence is again assumed.

Latent transition analysis (LTA) is a reparameterization of LCA that models change over time in a discrete developmental process (Collins & Wugalter, 1992; Lanza et al., 2003; Collins et al., 1994). In LTA, latent class membership is dynamic; participants may transition between classes over time. These dynamic latent classes are called latent statuses. In addition to latent status membership probabilities at each time and item response probabilities linking manifest indicators to latent statuses at each time, LTA estimates transition probabilities. The transition probabilities are the probabilities of latent status membership at time $t+1$ conditional on latent status membership at time t. The transition probabilities are often referred to as the τ parameters. In latent transition models, the γ and τ parameters comprise the structural model and the ρ parameters comprise the measurement model.

It is important to note how LTA differs from RMLCA. RMLCA uses manifest indicators from multiple time points in a traditional latent class model to identify classes that describe patterns of change over three or more times. For example, RMLCA may identify classes with different longitudinal patterns of drinking behavior; classes may include "consistent heavy drinkers" who drink heavily at three different points in time, and "late starters" who drink heavily at the third time but who did not drink heavily at the first two times. Comparatively, LTA describes longitudinal change in latent status membership between a pair of times. For example, LTA provides a way to describe the proportion of individuals who transition from non-drinking to heavy drinking between any two times whereas RMLCA does not.

Similar to latent class models, latent transition models may be extended to include a grouping variable. When a grouping variable is included, all parameters, including the τ parameters, can be conditioned on group membership.

1.3.2.1 Conditional Latent Transition Modeling

In conditional latent transition analysis (CLTA), continuous or categorical manifest variables are added to the model as covariates to predict latent status membership at time 1, and/or to predict transitions in latent status membership between time t and time $t+1$. Examples of possible applications of and the details of the mathematical models for latent transition modeling with a grouping variable and conditional latent transition modeling may be found in a variety of resources including Chung et al. (2005) and Lanza and Collins (under review).

1.3.3 Multivariable Latent Class Modeling

For the purposes of the current study, loglinear models that include multiple latent variables that measure multiple unique constructs (as opposed to multiple latent variables that measure one construct over time as in LTA) are called multivariable latent class models. Multivariable latent class models describe and test relations among multiple discrete latent constructs. Research questions like "What is the relation between gambling and substance use?" may be addressed using multivariable latent class modeling (MVLCM).

Interestingly, this research question brings us back to the discussion of two-way tables. There are only two possibilities for the structural part of the model in loglinear models with two latent variables: gambling and substance use are independent, or gambling and substance use are dependent. Leaving off the notation for the measurement model, the loglinear notation for the independence model is (G, S) and the loglinear notation for the dependence model is (GS). Thus, the information presented in Table 1.1 is directly related to the structural part of loglinear models with two and three latent variables. In other words, the structural part of multivariable latent class models are direct extensions of traditional loglinear models with manifest variables. The difference is that in loglinear modeling with manifest variables there is no measurement model estimated whereas in loglinear modeling with latent variables measurement error is estimated and removed from other parameter estimates. The measurement part of multivariable latent class models are direct extensions of the measurement models in latent class and latent transition models.

The approach to addressing a research question using MVLCM is the same as the one taken for loglinear models with manifest variables. First, a set of relevant models is selected for consideration. Models included in the set should correspond to theoretically logical, hypothesized relations among the variables. This set of models should contain a reasonable number of models that will likely range in complexity. Models within the set are then fit and compared to determine which is best at balancing fit and parsimony in its description of the observed relations among the variables. Once a model is selected, it may be used to make substantive conclusions about the nature of the relations among the latent variables.

1.3.4 Multiprocess Modeling

For the purposes of the current study, multiprocess models refer to models that examine change over time in two or more discrete developmental processes. For example, a research question fundamental to the study of gambling development is, “How do developmental transitions in gambling relate to substance use over time?” Such research questions may be addressed using multiprocess modeling (MPM). Early work with multiprocess models used an approach which has been referred to as associative latent transition analysis (ALTA; Flaherty & Collins (1999); Tang (2002).

The term “multiprocess models” refers to loglinear models with latent variables that may be used to model change over two or more times in two or more discrete developmental processes simultaneously. Thus, multiprocess models express relations among a minimum of four discrete latent variables. As discussed above, the complete set of possible four-variable models is very large. Fortunately, only a subset of the possible models are multiprocess models because they include development in each process. In other words, gambling at time 1 and time 2 (G_1 and G_2) would be assumed to be dependent, just as substance use at time 1 and time 2 (S_1 and S_2), so only loglinear models including the terms G_1G_2 and S_1S_2 would be considered. The discussion here is restricted to multiprocess models with two developmental processes measured at two times. The methods, however, may be extended to additional processes or additional times.

Similar to multivariable latent class models, the structural part of multiprocess models are direct extensions of traditional loglinear models with manifest variables. And, the approach to addressing research questions using MPM is the same as the one taken in MVLCM. In MPM, three types of relations would typically be included in all models that are part of the set of relevant models for consideration: (1) development in the first process, (2) development in the second process, and (3) hypothesized relations between the two processes. For example, to address the research question posed above, the three types of relations that should be included in all models within the set of possible models are: (1) development of gambling, (2) development of substance use, and (3) hypothesized relations between gambling and substance use.

To introduce the MPM methodological framework, consider the research question posed above about the relation between gambling and substance use. Suppose three possible sets of relations between gambling and substance use are hypothesized. (A larger set of possible types of relations and their corresponding multiprocess models are discussed in more depth in Chapter 4.) Theory and logic suggest that: (1) gambling development may be unrelated to substance use development; (2) gambling and substance use may be correlated at each time but previous substance use does not have a direct effect on current gambling; or (3) gambling may be affected by an interaction between previous gambling and substance use. Depending on the problem at hand and previous theory and research, different numbers of possible types of relations and different types of relations than those hypothesized here may be appropriate.

Based on these three hypothesized types of relations, three corresponding multiprocess models may be developed. The first hypothesized relation suggests that gambling and substance use both develop over time, but that they are independent. In the corresponding multiprocess model: (1) time 2 substance use depends on time 1 substance use, (2) time 2 gambling depends on time 1 gambling, and (3) gambling and substance use are independent. Using loglinear modeling notation, the structural part of this model is written (S_1S_2, G_1G_2); this multiprocess model is labeled Model 1. A summary of the loglinear notation for the structural part of the model and effects included in Model 1 are shown in Table 1.3. Table 1.3 also summarizes Models 2 and 3 discussed below.

Model 2 corresponds to the second hypothesized relation that not only do gambling and substance use develop over time but that there is also a cross-sectional relation between the two processes at each time point. In Model 2: (1) time 2 substance use depends on time 1 substance use, (2) time 2 gambling depends on time 1 gambling, (3) gambling and substance use are related at time 1, and (4) gambling and substance use are related at time 2. The structural part of Model 2 may be written $(S_1S_2, G_1G_2, S_1G_1, S_2G_2)$.

Model 3 corresponds to the third hypothesized relation that not only do gambling and substance use develop over time and that there is a cross-sectional relation between the two processes at each time point, but that there is also an interactive effect of past gambling and past substance use on gambling. In Model 3: (1) time 2 substance use depends on time 1 substance use, (2) time 2 gambling

Table 1.3. Summary of Selected Multiprocess Models

	Loglinear Notation	Included Effects
Model 1	(S_1S_2, G_1G_2)	$S_1, G_1, S_2, G_2, S_1S_2,$ G_1G_2
Model 2	$(S_1S_2, G_1G_2, S_1G_1, S_2G_2)$	$S_1, G_1, S_2, G_2, S_1S_2,$ G_1G_2, S_1G_1, S_2G_2
Model 3	$(S_1G_1G_2, S_1S_2, S_2G_2)$	$S_1, G_1, S_2, G_2, S_1S_2,$ $G_1G_2, S_1G_1, S_2G_2,$ $S_1G_2, S_1G_1G_2$

depends on time 1 gambling and substance use jointly, and (3) gambling and substance use are related at time 2. The structural part of Model 3 may be written $(S_1G_1G_2, S_1S_2, S_2G_2)$.

1.4 Purpose

The purpose of the current project is to present a collection of sophisticated applications of the models discussed above in order to illustrate the usefulness and feasibility of these models in practical developmental applications. Chapters 2, 3, and 4 provide three empirical studies employing some of the more complex models discussed here to address concrete research questions about development. Each of the three empirical studies focuses on: (1) the selection of an appropriate model to address the research question, and (2) the interpretation of the selected model to make substantive conclusions. The empirical studies are presented in order of increasing model complexity; the first example illustrates the use of LCA, LCA with a grouping variable, and CLCA with a grouping variable; the second illustrates the use of LCA, RMLCA, and MVLCM; the third illustrates the use of LTA and MPM. It is hoped that the clarification and dissemination of how to practically apply these models in developmental science will inspire new applications in areas that have not previously used these types of models.

The first two empirical studies address research questions from the emerging area of gambling and problem gambling research. In the first study, LCA is used to identify types of college-student gamblers. Then, CLCA with a grouping variable is used to predict gambling latent class membership from a variety of demographic and substance use variables that have been shown previously to be associated with gambling behavior. In the second empirical study, LCA is used to identify types of adolescent and young adult gamblers. RMLCA is used to identify latent trajectory classes of drinking frequency, quantity, and intensity. Then, MVLCM is used to describe the relation between gambling and longitudinal patterns of drinking. The second study is designed to explore the extent to which longitudinal patterns of drinking predict gambling latent class membership.

The third empirical example addresses a research question about the relation between drinking and smoking development. In this example, LTA is used to model adolescent drinking development over time and smoking development over time as separate processes. Then, MPM is used to describe the nature of the relation between drinking and smoking development. This example is designed to explore the extent to which developmental transitions in drinking vary by smoking behavior.

Finally, Chapter 5 presents an overall discussion of the current project. It compares and contrasts the models used in the three empirical studies, and discusses the general use of loglinear models with latent variables including some advantages and limitations. Finally, a few alternatives to discrete latent variable models are mentioned, and some future directions for the current research are discussed.

CHAPTER 2

Empirical Study #1: Identifying and Predicting Types of College-Student Gamblers

2.1 Introduction

Gambling and problems with gambling are increasingly being recognized as emerging public health concerns as the availability of legalized gambling opportunities in the United States continues to expand at an unprecedented rate (Korn & Shaffer, 1999; National Gambling Impact Study Commission, 1999). The most recent national surveys of gambling estimate that approximately 85% of the general adult population has gambled and 60-85% has done so in the past year (National Gambling Impact Study Commission, 1999; Welte et al., 2001; Volberg, 2007). Currently, it is estimated that approximately 1-2% of the general adult population has experienced clinical-level problems with gambling and an additional 3-5% has experienced sub-clinical problems with gambling, rates that appear to have increased over the years (Shaffer et al., 2004, 1999; Shaffer & Hall, 2001; National Gambling Impact Study Commission, 1999). Clinical and sub-clinical problems with gambling appear to be even more prevalent among today's youth; approximately 4-8% and 10-14% have experienced clinical-level and sub-clinical problems with gambling, respectively (Shaffer & Hall, 1996; Derevensky et al., 2003).

Gambling is also receiving increasing attention from researchers in prevention and treatment of a variety of problem behaviors. This is mostly due to an ever-growing body of literature making it increasingly clear that individuals who engage in problematic gambling are also at increased risk of experiencing a host of other negative outcomes including: substance use, abuse, and dependence; problems with family, education, and work; delinquency and arrest; mood and anxiety disorders; suicide ideation and attempts; and severe financial difficulties (Petry, 2005; National Research Council, 1999; National Gambling Impact Study Commission, 1999; Ladd & Petry, 2003; Rosenthal & Lorenz, 1992; Specker et al., 1996; Phillips et al., 1997; Miller & Westermeyer, 1996; Frank et al., 1991; Lesieur, 1979; Lesieur et al., 1991; Neighbors & Larimer, 2004).

2.1.1 Demographic Characteristics of Gamblers

Many demographic characteristics have been shown to be associated with gambling. In particular, gambling behavior appears to be substantively different between men and women, and among different racial/ethnic groups.

Gender differences in gambling behavior appear to be particularly pronounced. Men are more involved with gambling in general and have higher rates of gambling problems than women (Winters et al., 1993; Wallisch, 1993; Zitzow, 1996; Stinchfield et al., 1997; National Research Council, 1999; Welte et al., 2001; Lesieur et al., 1986; Ibáñez et al., 2003; Welte et al., 2004; Lesieur & Rosenthal, 1991; Raylu & Oei, 2002; Volberg, 1996; LaBrie et al., 2003; National Gambling Impact Study Commission, 1999; Ladd & Petry, 2003). In general, men are more likely to be exposed to gambling during adolescence, place their first bets at a younger age, and begin gambling regularly at earlier ages (Ibáñez et al., 2003; Tavares et al., 2003; Petry, 2002). Women, however, tend to have a faster progression to problem gambling than men, which is sometimes referred to as the "telescoping effect"

(Ibáñez et al., 2003; Tavares et al., 2003, 2001; Potenza et al., 2001).

In addition, women are more likely to have gambling behavior triggered by negative emotional feelings, and their motivation to continue gambling is often inherent enjoyment in the game itself (Ibáñez et al., 2003; Tavares et al., 2003; Brown & Coventry, 1997; Lesieur & Blume, 1991; Lesieur & Rosenthal, 1991). In contrast, men are more likely to have gambling behavior triggered by possible gain, and their motivation to continue gambling is often the potential profit (Ibáñez et al., 2003; Brown & Coventry, 1997; Lesieur & Blume, 1991; Lesieur & Rosenthal, 1991). There also appear to be differences in the types of games men and women prefer. Women tend to prefer playing bingo, video bingo, and at video lottery terminals whereas men seem to have more eclectic preferences including slots, video bingo/lottery, cards, and combinations of these and other games (Ibáñez et al., 2003; Tavares et al., 2003).

There has been relatively little research investigating patterns of gambling in different ethnic and cultural groups (Raylu & Oei, 2002). Research has consistently shown that ethnic minorities, including African Americans, Hispanics, and Asians, have higher rates of problem gambling than whites (*e.g.*, Lesieur & Rosenthal, 1991; Welte et al., 2004; Petry et al., 2005; Blaszczynski et al., 1998). It has been hypothesized that one reason for this may be because minorities in the United States tend to have much lower net worths than whites, even at the same income levels (Welte et al., 2004). Because of these decreased financial resources, it may be more likely for minorities to endorse pathological gambling criteria that inquire about financial problems related to gambling. It has also been hypothesized that another reason for these increased rates of problem gambling may be factors like cultural values and beliefs, acculturation effects, and treatment-seeking behaviors of different cultures (Raylu & Oei, 2002).

2.1.2 Substance Use of Gamblers

One of the most important correlates of gambling and problems with gambling is the use of alcohol and other substances. Research has clearly established that gambling and alcohol use, and in particular problem gambling and problem alcohol use, are positively correlated (Smart & Ferris, 1996; Grant et al., 2002; Feigelman et al., 1998; National Gambling Impact Study Commission, 1999; Welte et al., 2001; Ramirez et al., 1984; Lesieur et al., 1986). For example, studies conducted on the general adult population have found that approximately 25-50% of individuals identified as problem gamblers have also experienced alcohol abuse/dependency or other drug use. Welte et al. (2001) found that the prevalence of past-year problems with gambling increased with higher levels of alcohol consumption. In addition, gambling problems have been shown to be correlated with marijuana, cocaine, heroin, and other illicit drug use (Lesieur et al., 1986; Dube et al., 1996).

The relation between gambling problems and problems with alcohol appears to be particularly pronounced for college students. For example, in a nationally representative study of college students, LaBrie et al. (2003) found that alcohol-related behaviors appeared to be the strongest correlates of problem gambling. In addition, they found that lifetime alcohol use, past-year alcohol use, past-month alcohol use, binge drinking, past-year and past-month cigarette use, past-year marijuana use, and past-year and past-month use of other illicit drugs were all significantly related to gambling participation during college.

Although it is known that problems with gambling tend to co-occur with substance use problems, it is unclear exactly how substance use during college is related to specific types of gambling. Many researchers in the problem gambling field have noted that gambling and substance use are highly related and that the nature of the predictive relation between the two processes should be investigated, including Barnes et al. (1999, 2002), Griffiths and Sutherland (1998), Kassinove et al. (2000), Nower et al. (2004), Vitaro et al. (1998, 2001), and Winters and Anderson (2000).

2.1.3 Identifying Types of Gamblers

As prevention and treatment efforts for gambling problems become priorities for public health agencies, it is important that the distribution, correlates, and determinants of gambling are understood. In addition, it is essential that researchers provide ways to identify individuals who may be at increased risk for developing gambling problems.

Historically, the identification of individuals who may benefit from prevention and treatment programs has been conducted using diagnostic criteria to categorize gamblers as being "non-problem," "problem," or "pathological." This approach is based on the number of diagnostic criteria met, rather than which criteria are met. Gamblers who endorse none or very few of the diagnostic criteria are not considered to have any problems with gambling. Gamblers who endorse enough diagnostic criteria to meet the clinical definition of pathological gambling are considered to be pathological gamblers; these gamblers are often referred to treatment programs. Gamblers who endorse some moderate number of diagnostic criteria, but not enough to meet the clinical definition of pathological gambling, are considered to be problem gamblers who are at high risk for developing pathological gambling; these gamblers are often referred to prevention programs.

Pathological gambling diagnostic criteria address several dimensions of problematic gambling behavior including: preoccupation, tolerance, loss of control, withdrawal, escape, chasing, lying, illegal acts, risking significant relationship(s), and bailout. Using the approach to identifying types of gamblers described above, individuals primarily endorsing criteria about intra- and interpersonal problems stemming from gambling, and individuals primarily endorsing criteria about financial problems stemming from gambling may be identified as the same type of gambler if their total number of endorsed criteria is the same. However, as discussed by Shaffer et al. (2004) this

> "unidimensional additive scoring ... is inadequate to represent a multidimensional latent state. The method of summing endorsed characteristics assumes that all dimensions exist on the same additive continuum, and that all dimensions equally predict gambling disorders ... This equivalence is highly unlikely and misleading."

It is likely that there are types of individuals who are at lower or higher risk of developing gambling problems, and that these individuals may be identified long before they are at the point of meeting diagnostic criteria. For example, substance use researchers identify youth at high risk for developing substance abuse and dependence long before youth are likely to meet diagnostic criteria. Furthermore, it is likely that types of gamblers may be identified by using information about gambling behavior that is not included in the diagnostic criteria, for example low-level types of gambling.

The identification of types of gamblers based on similarities in behavioral characteristics should be considered and investigated as an important part of understanding gambling and the development of problems with gambling. One way to do this is to consider the activities in which gamblers engage. It may be that some gamblers only gamble at casinos, some only gamble on the Internet, and others participate in multiple activities on a regular basis. By using methods that allow the multi-dimensional nature of behavior to be examined, it may be possible to identify types of gamblers in ways that are more helpful for prevention and treatment researchers and providers. In addition, once types of gamblers have been identified, it is possible to explore how individual and contextual characteristics are related to type in order to reach a more nuanced understanding of gambling and the development of problems with gambling.

2.1.4 The Current Study: Types of Gambling in a College-student Population

This first empirical study uses data from a large northeastern university to explore the underlying discrete latent structure of gambling and to explore whether demographic characteristics and substance use are predictive of that structure. College students may be a particularly interesting population in which to study gambling. Rates of gambling problems appear to be somewhere in between rates for youth and for adults; approximately 4-8% of college students report clinical-level problems with gambling (specific percentages from a variety of studies can be found in LaBrie et al., 2003). College appears to be a time when many individuals are willing to take risks in a variety of domains. During this time individuals may initiate or expand their gambling behavior as they approach an age at which gambling activities become legal. It may also be a time when some individuals are first beginning to develop problems with gambling. College students, then, may be a particularly good population on which to try out new ways of identifying types of gamblers.

The current study addresses the question, “Are there identifiable types of gamblers based on the activities in which they participate, and do these types vary between men and women?” In addition, the current study also addresses the question, "Do demographic characteristics and substance use predict gambling type, and do the predictive effects vary between men and women?" These questions are addressed by: (1) identifying types of gamblers using LCA with gender as a grouping variable; and (2) predicting gambling latent class membership from a variety of demographic characteristics and substance use behaviors using CLCA with gender as a grouping variable.

2.2 Methods

2.2.1 Participants

Data are from a study conducted at a large northeastern university designed to collect information about risk-taking behaviors and motivations from undergraduate college students and student-athletes (Yusko et al., under review). The original student sample was drawn from introductory psychology and communications courses. The original student-athlete sample was drawn from varsity student athletes on 17 athletic teams participating in a mandatory alcohol education seminar. There were no inclusion/exclusion criteria except that participants had to be between 18 and 26 years of age. Data were collected in Spring 2005, Winter 2005, and Spring 2006. The sample for the current study consisted of 507 student participants (37% male, 63% female) who reported their gender and who responded to at least one question about gambling activity participation. (Student-athletes were not included in sample for the current study.)

2.2.1.1 Procedure

Students were invited to participate in a voluntary research study. The nature of the research was explained and an information form was reviewed by a trained research assistant. Participants gave verbal assent prior to survey completion. The original study was approved by the university's Institutional Review Board. The survey took approximately 30 minutes to complete, and students received research credit for their participation.

2.2.2 Measures

2.2.2.1 Gambling Activity Participation

Gambling activity participation was measured using 13 binary indicators corresponding to 13 types of gambling activities. Participants were considered to be "players" of an activity if they reported having engaged in the activity at least once during the past 12 months; participants were considered to be "non-players" of an activity if they reported not having engaged in the activity during the past 12 months. The 13 gambling activities, labels for each gambling activity, and the proportion of participants playing each activity are shown in Table 2.1.

2.2.2.2 Demographic Characteristics

Six demographic characteristics shown to be associated with gambling in previous research were examined to determine whether they were predictive of gambling latent class membership. Table 2.2 shows the distribution of the demographic characteristics.

2.2.2.3 Substance Use

A variety of types of substance use have been shown to be associated with gambling in previous research. Four types of substance use were examined to determine whether they were predictive of gambling latent class membership. Table 2.3 shows the distribution of the substance use behaviors.

Past-semester cigarette use was measured using a question asking, "On average, how many cigarettes do you usually smoke a day during the current semester?" Participants responded with the total number of cigarettes. Based on the frequency distribution of responses, the majority of participants reported no smoking and the distribution of the remaining responses was highly skewed. Therefore, it was determined that the occurrence of smoking, rather than the frequency, was important and a dichotomous indicator of past-semester smoking was created.

Past-year alcohol use was measured using a question asking, "How often did you drink alcohol during the past year?" This question employed an eight-point response scale to determine the frequency of past-year alcohol use; response options ranged from "I did not drink at all" to "once a day or more." Based on the

Table 2.1. Gambling Activities

Activity	Label	Percent Playing
Played cards or board games for money with family or friends	cards	49.9
Played table games at a commercial card parlor or casino	casino	19.5
Bet on games of personal skill	skill	30.0
Played the stock or commodities market	stock	6.4
Played commercial BINGO	bingo	7.1
Shot dice or played craps	dice	10.9
Wagered on the Internet (on casino or other games)	inter	10.5
Bet on sports cards, football pools, or parlays	sports	15.5
Bet on horse or dog races	horses	6.5
Wagered on intercollegiate games with a bookie (on- or off-campus)	bookie	4.2
Bought lottery tickets	lotto	45.0
Played slot or electronic poker machines	slots	19.8
Engaged in some other type of gambling	other	18.4

Table 2.2. Demographic Characteristics

Variable	Label	Percent
Gender	Men	36.9
	Women	63.1
Year in School	First Year	10.2
	Sophomore	41.1
	Junior	41.5
	Senior	5.5
	Fifth Year or Above	1.6
Ethnic Background	White	58.9
	Asian	17.6
	Black or Hispanic	14.0
	Other	9.4
Current-semester Living Arrangement	Dorm	49.6
	Off-campus	33.9
	Other	16.5
Member of Greek Organization	No	87.6
	Yes	12.4
Grade Point Average	Mean = 3.00	
	SD = 0.51	

Note: Percentages may not sum to 100% for a particular variable due to rounding error.

Table 2.3. Substance Use Behaviors

Variable	Label	Percent
Past-semester Cigarette Use	No	77.5
	Yes	22.5
Past-year Alcohol Use	None or Low	23.6
	Moderate	34.3
	Heavy	42.2
Past-year Binge Drinking	No	20.9
	Yes	79.1
Past-year Marijuana Use	None	52.1
	Moderate	25.7
	Heavy	22.2

Note: Percentages may not sum to 100% for a particular variable due to rounding error.

frequency distribution of responses, the eight-point scale did not approximate a normally distributed continuous variable, and there were not enough participants using each response option to justify an eight-level categorical variable. Based on the frequency distribution, three categories of past-year alcohol use were created: no or low use was defined as drinking less than once per month; moderate use was defined as drinking once to three times per month; heavy use was defined as drinking once per week or more often. These categories were selected to create groups large enough to compare while capturing most of the variation in drinking frequency.

Past-year binge drinking for men was measured using a question asking, "How many times in the past year have you drank five or more drinks in one sitting?"; for women the question was modified to "four or more drinks." Participants responded with the total number of days. Based on the frequency distribution of responses, the majority of participants reported binge drinking but one-fifth reported no binge drinking and the distribution of the remaining responses was highly skewed. Therefore, it was determined that the occurrence of binge drinking, rather than the frequency, was important and a dichotomous indicator of past-year binge drinking was created. The dichotomization of binge drinking was similar to the approach taken by Auerbach and Collins (2006).

Past-year marijuana use was measured using a question asking, "How often have you used Marijuana or Hashish in the last year?" This question employed an eight-point response scale to determine the frequency of past-year marijuana use; response options ranged from "no use in the last year" to "everyday or nearly everyday." Based on the frequency distribution of responses, the eight-point scale did not approximate a normally distributed continuous variable, and there were not enough participants using each response option to justify an eight-level categorical variable. Based on the frequency distribution, three categories of past-year marijuana use were created: no use was defined as not smoking marijuana; moderate use was defined as smoking marijuana once per month or less often; heavy use was defined as smoking marijuana more than once per month. These categories were selected to create groups large enough to compare while capturing most of the variation in the frequency of marijuana use.

2.2.3 Overview of Models Fit in the Current Study

The first purpose of the current study is to identify types of college-student gamblers, and to see if these types differ by gender. To address this objective, LCA with gender as a grouping variable was used to develop and select a model identifying types of gamblers for men and women based on the indicators of gambling activity participation discussed above. Gender was included as a grouping variable to allow for the possibility of gender differences in the measurement model. Using LCA with a grouping variable, a variety of models with different numbers of latent class were explored. The models were evaluated based on the G^2 fit statistic and the plausibility of their interpretations given the observed data and theory. The AIC and BIC were used to aid in model selection where appropriate.

The second purpose of the current study is to predict gambling latent class membership from a variety of demographic characteristics and substance use behaviors. To address this objective, CLCA with a grouping variable was used. Using CLCA with a grouping variable, each of the remaining five demographic characteristics show in Table 2.2 (gender remained the grouping variable) and each of the four substance use behaviors shown in Table 2.3 was added to the selected latent class model as a predictor of latent class membership separately. That is, each of the predictors was entered one at a time into the selected latent class model as a covariate; there are nine predictors, so a total of nine conditional latent class models with a grouping variable were estimated. Predictors were examined individually because these analyses were exploratory and to avoid issues of model estimation that can arise due to sparseness. Each model was examined to determine whether the predictor was predictive of gambling latent class membership.

All latent class and conditional latent class models were estimated using PROC LCA. PROC LCA is a new SAS procedure for LCA developed by the Methodology Center at Penn State. This procedure was developed for Version 9.1 of the SAS System for Windows.[1] PROC LCA is available by free download from the Methodology Center's website at `http://methodology.psu.edu/`. Examples of the SAS programming code used to fit the models discussed above may be found in Appendix A.

2.3 Results

2.3.1 Latent Class Models

Preliminary analyses were conducted for men and women separately to determine the extent to which the latent class structure differed by gender. In the preliminary analyses, models with 2 to 6 latent classes were fit for both men and women. The identification of all models was checked using 10 sets of random starting values. For models in which all sets of starting values did not converge to the same solution, the solution with the smallest G^2 that was replicated was examined; for models that did not have a G^2 value replicated across the sets of starting values, the solution with the smallest G^2 was examined in combination with the other solutions.

The preliminary analyses suggested that models with 3 or more latent classes required the use of heavily restricted measurement models and starting values to help with model identification. In addition, these analyses suggested that a model with either 3 or 4 latent classes was appropriate for men, and a model with either 3 or 4 latent classes was appropriate for women. The 3- and 4-class models were chosen based on the G^2 fit statistic, AIC, BIC, and the plausibility of their interpretations given the observed data and theory. Fit statistics for the 2- to 6-class models fit for men and women separately are shown in Table 2.4. The fit statistics shown in Table 2.4 are from the solutions selected based on the criteria described above.

The 3- and 4-class models for men and women were closely examined to determine the extent to which the latent class structure differed by gender in order to determine how to fit the latent class model with gender as a grouping variable. The 3-class models did not appear to have classes defined in the same way between men and women. In the 4-class models, the examination revealed that although men and women did have somewhat different item-response probabilities for some of the indicators given latent class membership, the overall pattern was the same across genders. In other words, the classes in the 4-class models appeared to be defined by "high" and "low" item-response probabilities on the same subsets of indicators. For both men and women the 4-class solution included: (1) a class characterized by very low endorsement of all activities, with just slightly higher probabilities for playing cards and playing lotto; (2) a class characterized by high endorsement of

Table 2.4. Fit Statistics for Models Fit for Men and Women Separately

Men	Model	G^2	*df*	AIC	BIC
	2-class	512.86	8164	566.86	654.10
	3-class	**441.45**	**8150**	**523.45**	**655.92**
	4-class	**404.54**	**8136**	**514.54**	**692.25**
	5-class	374.00	8122	512.00	734.95
	6-class	351.23	8108	517.23	785.41
Women	**Model**	G^2	*df*	**AIC**	**BIC**
	2-class	547.38	8164	601.38	703.13
	3-class	**466.13**	**8150**	**548.13**	**702.63**
	4-class	**416.46**	**8136**	**526.46**	**733.72**
	5-class	379.65	8122	517.65	777.67
	6-class	341.30	8108	507.30	820.07

Note: Boldface type indicates the selected models.

playing cards and playing the lotto, and very low endorsement of seven other activities; (3) a class characterized by high endorsement of playing cards and games of skill, with lower and variable endorsement of the other activities; and (4) a class characterized by high endorsement of playing cards, table games at a casino, games of skill, lotto, slots, and other (item-response probabilities for the other activities varied between men and women).

Based on the preliminary analyses, latent class models with 3 and 4 classes that included gender as a grouping variable were selected to be evaluated and compared. Because the 3-class model showed much greater gender differences than the 4-class model, parameter restrictions were handled differently in the two models. The 3-class model was restricted within gender using the patterns of restrictions suggested by the corresponding 3-class models fit for men and women separately. None of the item-response probabilities for men were restricted to be equal to any of the item-response probabilities for women.

The 4-class model was restricted across gender using a pattern of restrictions reflecting the similarities in latent classes between men and women described above. That is, certain item-response probabilities for men were restricted to be equal to the corresponding item-response probabilities for women in order to define the classes in similar ways for both genders. The restrictions on the measurement model were chosen such that the four classes were restricted to be characterized by the subsets of indicators of particular importance for each class. It was clear that measurement invariance across genders would not hold for each indicator in each latent class, but that it was possible to restrict subsets of indicators across genders to ensure consistent interpretation of the classes across genders. This complex pattern of restrictions was desirable because: (1) the latent classes suggested by the 4-class model were defined in similar ways for men and women by the same primary behaviors, allowing a cleaner interpretation of the effects of the predictors later; and (2) it allowed enough flexibility in the definitions of the latent classes to capture the major differences between men and women.

A comparison of the 3- and 4-class models based on the G^2 fit statistic, AIC, BIC, plausibility of their interpretations given the observed data and theory, and stability of the solutions suggested that the 4-class model was most appropriate. Thus, the 4-class model was selected as the latent class model describing types of gamblers. Fit statistics for the 3- and 4-class models with gender as a grouping variable are shown in Table 2.5.

Latent class membership probabilities and item-response probabilities for playing particular activities given latent class membership for the selected model are shown in Table 2.6. Item-response probabilities restricted to be equal are denoted with the same letter. In this model, the first latent class was labeled "non-gamblers" because members of the latent class had very low (0.0025) probabilities of playing eleven of the gambling activities and fairly low (0.1652) probabilities of playing cards and lotto. The second latent class was labeled "cards and lotto" because members had very low (0.0567) probabilities of playing seven of the gambling activities and high (0.6637) probabilities of playing cards and lotto; the probabilities of playing the other four activities varied between men and women. The "cards and skill" and "multi-game" latent classes were labeled using a similar method.

Table 2.6 shows that there were substantial gender differences in the proportions of men and women who were non-gamblers, cards and games of skill players, and multi-game players. Women were more likely to be non-gamblers, and men were more likely to be cards and games of skill players, and multi-game players.

2.3.2 Conditional Latent Class Models

Once a model describing types of gamblers was selected and interpreted, it was possible to investigate the predictive effects of the demographic characteristics and substance use behaviors on latent class membership. Each of the remaining five demographic characteristics (gender remained a grouping variable) and each of the four substance use behaviors was added to the selected latent class model separately using CLCA with a grouping variable. Predictors were examined individually because these analyses were exploratory and to avoid issues of model estimation that can arise due to sparseness.

In each of the nine conditional latent class models, non-gamblers were used as the reference class for the multinomial logistic regression that estimated the effect of a covariate on the odds of membership in the other latent classes relative to the reference class. That is, the effect of a significant predictor may be interpreted as the effect of a covariate on the odds of membership as a cards and lotto player,

Table 2.5. Fit Statistics for Models with Gender as a Grouping Variable

Model	G^2	*df*	AIC	BIC
3-class	913.71	16310	1059.71	1368.39
4-class	**909.05**	**16327**	**1021.05**	**1257.84**

Note: Boldface type indicates the selected model.

Table 2.6. Latent Class Membership and Item-response Probabilities*

Men	Latent Class Membership			
	Non-gamblers (0.295)	Cards & Lotto (0.365)	Cards & Skill (0.224)	Multi-game (0.115)
cards	0.1652[a]	0.6637[c]	0.9774[e]	0.9096[f]
casino	0.0025[b]	0.2126	0.3780	0.9096[f]
skill	0.0025[b]	0.4485	0.9774[e]	0.9096[f]
stock	0.0025[b]	0.0567[d]	0.2137	0.1872
bingo	0.0025[b]	0.0567[d]	0.0461	0.2023
dice	0.0025[b]	0.0567[d]	0.3516	0.6204
inter	0.0025[b]	0.0567[d]	0.4112	0.6237
sports	0.0025[b]	0.0567[d]	0.6877	0.7165
horses	0.0025[b]	0.0567[d]	0.0967	0.3589
bookie	0.0025[b]	0.0567[d]	0.1656	0.3224
lotto	0.1652[a]	0.6637[c]	0.4275	0.9096[f]
slots	0.0025[b]	0.3110	0.0000	0.9096[f]
other	0.0025[b]	0.3081	0.3597	0.9096[f]
Women	**Latent Class Membership**			
	Non-gamblers (0.525)	Cards & Lotto (0.347)	Cards & Skill (0.075)	Multi-game (0.054)
cards	0.1652[a]	0.6637[c]	0.9774[e]	0.9096[f]
casino	0.0025[b]	0.2613	0.0889	0.9096[f]
skill	0.0025[b]	0.2106	0.9774[e]	0.9096[f]
stock	0.0025[b]	0.0567[d]	0.1710	0.4165
bingo	0.0025[b]	0.0567[d]	0.3900	0.4832
dice	0.0025[b]	0.0567[d]	0.2624	0.7332
inter	0.0025[b]	0.0567[d]	0.0000	0.5362
sports	0.0025[b]	0.0567[d]	0.4433	0.6020
horses	0.0025[b]	0.0567[d]	0.1800	0.3636
bookie	0.0025[b]	0.0567[d]	0.0366	0.2470
lotto	0.1652[a]	0.6637[c]	0.6387	0.9096[f]
slots	0.0025[b]	0.3476	0.2192	0.9096[f]
other	0.0025[b]	0.1457	0.4066	0.9096[f]

Note: Entries denoted by the same letter were restricted to be equal during estimation.
* Probability of playing an activity given latent class membership.

cards and games of skill player, or multi-game player relative to membership as a non-gambler using odds ratios. The significance tests and odds ratios for men and women, for each of the demographic characteristics and substance use behaviors, are shown in Table 2.7.

School year, living in off-campus housing, Greek membership, and moderate and heavy past-year alcohol use were significant predictors of gambling latent class membership at the $\alpha = 0.05$ level. The strongest effect of off-campus housing appears to be on the odds of membership as a multi-game player relative to membership as a non-gambler. For men, the odds of membership as a multi-game player relative to membership as a non-gambler was almost 3 times higher for college students living in off-campus housing relative to those living in a dorm (OR = 2.9); for women it was almost 9 times higher (OR = 8.9).

Gender differences in the effects of school year, Greek membership, and alcohol use are particularly interesting. For example, for men it appears that the odds of membership as a multi-game player relative to membership as a non-gambler was 1.3 times higher for each one-year increase in school year. Comparatively, for women the odds of membership as a multi-game player relative to membership as a non-gambler was almost 3 times higher for each one-year increase in school year (OR = 2.8).

Greek membership had a variable effect on gambling latent class membership. For men, the odds of membership as a cards and lotto player relative to membership as a non-gambler was 2.5 times higher for Greek members relative to non-members. The odds of membership as a cards and games of skill player relative to membership as a non-gambler was 1.8 times higher for Greek members relative to non-members. Perhaps surprisingly, however, the odds of membership as a non-gambler relative to membership as a multi-game player was 9.8 times higher (this is the inverse of the odds ratio reported in Table 2.7) for Greek members relative to non-members. Comparatively, for women the odds of membership as a non-gambler relative to membership as a cards and games of skill player or a multi-game player were 58.1 and 67.1 times higher for Greek members relative to non-members, respectively.

Finally, the strongest effects of moderate and heavy past-year alcohol use appear to be on the odds of membership as a multi-game player relative to membership as a non-gambler. Although this is true for both men and women, the effects

Table 2.7. Odds Ratios and Significance Tests for Demographic Characteristics and Substance Use Behaviors

Predictor		Latent Class Membership: Non-gamblers	Cards & Lotto	Cards & Skill	Multi-game	$p-$ value
School Year	Men	Ref	0.5333	0.5190	1.2454	< 0.01
	Women	Ref	0.9762	0.6560	2.8393	
Asian	Men	Ref	2.2467	1.7483	1.2546	0.14
	Women	Ref	0.5602	1.9120	0.0134	
Black or Hispanic	Men	Ref	0.8272	0.8596	0.0794	0.61
	Women	Ref	0.7049	0.5246	0.6319	
Grade Point Average	Men	Ref	0.9445	1.0593	0.8858	0.60
	Women	Ref	0.9176	0.5917	0.9141	
Off-campus Housing	Men	Ref	0.7775	0.6182	2.8954	< 0.01
	Women	Ref	1.4234	0.6081	8.9374	
Greek Membership	Men	Ref	2.4388	1.8420	0.1026	0.02
	Women	Ref	0.6625	0.0172	0.0149	
Current-semester Cigarette Use	Men	Ref	1.5349	0.7820	1.3205	0.09
	Women	Ref	2.0825	0.2928	4.3358	
Past-year Alcohol Use:						
Moderate	Men	Ref	2.6764	2.8108	7.5368	< 0.01
	Women	Ref	2.7363	0.1988	48.6764	
Heavy	Men	Ref	1.4290	1.5334	9.6734	0.01
	Women	Ref	1.5462	0.5691	85.6294	
Binge Drinking	Men	Ref	2.1872	1.3427	3.1115	0.67
	Women	Ref	1.1835	1.2187	2.3960	
Past-year Marijuana Use:						
Moderate	Men	Ref	1.2076	1.0553	2.9760	0.37
	Women	Ref	1.0489	0.2173	1.9337	
Heavy	Men	Ref	1.2937	1.8501	1.8814	0.31
	Women	Ref	0.6761	1.0123	3.7914	

appear to be especially strong for women. For men, the odds of membership as a multi-game player relative to membership as a non-gambler was 7.5 times higher for moderate drinkers relative to non-drinkers; for women it was 48.7 times higher for moderate drinkers relative to non-drinkers. For heavy drinkers the effect appears to be even stronger; for men the odds of membership as a multi-game player relative to membership as a non-gambler was 9.7 times higher for heavy drinkers relative to non-drinkers, for women it was 85.6 times higher.

2.4 Discussion

The current study used LCA to identify types of college-student gamblers using data from a large northeastern university. It also used CLCA to predict gambling latent class membership from a variety of demographic characteristics and substance use behaviors previously shown to be related to gambling.

2.4.1 Types of College-student Gamblers

The data suggested that four types of gamblers can be identified. For both men and women, types of gamblers included: (1) non-gamblers, (2) cards and lotto players, (3) cards and games of skill players, and (4) multi-game players. Although the classes were characterized by similar item-response probabilities, there was some variation between men and women in their probabilities of playing some activities given latent class membership. This is indicative of a substantive difference in the gambling behavior of men and women. In addition, it appeared that there were substantial gender differences in the probabilities of membership in the latent classes. Men were more likely to be gamblers than women, which was reflected in men being more likely to be cards and games of skill players and multi-game players. Women who gambled were most likely to play cards and the lotto, and the proportion of women in this latent class was approximately equal to the proportion of men.

Gender differences in the item-response probabilities were seen to some degree in the cards and lotto, cards and games of skill, and multi-game latent classes. The largest gender differences were seen in the cards and games of skill latent class.

This latent class was defined by high probabilities of playing cards and games of skill, and lower and variable probabilities of playing the other activities. Men in this latent class had a high probability of also betting on sports whereas women in this latent class had a high probability of also playing the lotto. In addition, men in this class were more likely to gamble on casino table games and on the internet than were women, but women were more likely to play bingo and slots.

These findings suggest that although there may be a group of individuals whose gambling behavior is primarily defined by playing cards and games of skill, these individuals do play some other activities, unlike men and women who are cards and lotto players. However, these individuals are different from multi-game players because they are unlikely to gamble at a casino, play slots, and play other activities. Additional research is needed to determine whether games of skill may act as "gateway games" to new types of gambling activities for some individuals, and how this may differ by gender.

An interesting finding is that gambling on the internet does not appear to be common among participants engaging in only one or two types of gambling activities. Both men and women who gamble on the internet appear to also be engaging in a variety of other gambling activities — the only latent class for which internet gambling was endorsed was the multi-game players. Additional research on internet gambling is needed to make conclusions about characteristics of the behavior itself, but from these results it does not appear that internet gambling frequently occurs by itself or only with very common gambling behaviors like card playing and the lottery.

Examining the latent class membership probabilities in combination with the pattern of item-response probabilities, an order of severity of the latent classes is suggested. It is reasonable to believe that non-gamblers are at very low risk of developing problems with gambling as these participants were unlikely to engage in any type of gambling in the past year. It may also be reasonable to believe that multi-game players are at high risk of developing gambling problems as these participants were likely to engage in six or more types of gambling activities in the past year. However, it is important to note that this discussion does not take into account frequency or intensity of gambling behavior, which may be important factors in identifying those at highest risk.

Comparing cards and lotto players to cards and games of skill players to determine who is at higher risk of developing problems with gambling is particularly interesting. The cards and lotto latent class was defined by high probabilities of cards and lotto playing and low probabilities of playing all of the other activities. This suggests that although these participants do gamble and may do so frequently, they were restricting their gambling to the two most common and, arguably, the least problematic activities. In contrast, the cards and games of skill latent class had high probabilities of playing a few additional games, as discussed above, and variable probabilities of playing the remaining games. This suggests that cards and games of skill players had engaged in more variable gambling activity in the past year as compared to cards and lotto players. It appears, then, that the order of severity of lowest risk to highest risk for developing gambling problems is: non-gamblers, cards and lotto, cards and games of skill, and multi-game players.

If the order of severity just discussed is reasonable, it is important to consider whether different prevention strategies should be recommended for different latent classes, and specifically for cards and games of skill players compared to multi-game players. The results of the current study do not provide a way to make recommendations about whether prevention programs should contain different kinds of content for cards and games of skill players and multi-game players. It does, however, suggest that program providers should consider using a variety of strategies for recruiting program participants so that all types of at-risk gamblers may be targeted. Multi-game players may be comparatively easier to target than cards and games of skill players because multi-game players are more likely to gamble in public venues like casinos, bingo halls, and race tracks or off-track betting parlors. Cards and games of skill players may prefer private or covert venues that involve betting among individuals, at least at their current level of gambling.

If it is desirable to implement prevention programs to help those at high-risk prior to an escalation in behavior, cards and games of skill players may be the group to target but these individuals may be difficult to find. Additional research on the significance of gambling in public versus private venues, and on the importance of preferences for games available in each venue is needed to make conclusions about how different prevention strategies may be needed for different latent classes.

2.4.2 Predicting Gambling Latent Class Membership

Gambling latent class membership was predicted by several demographic characteristics and substance use behaviors including gender, school year, current-semester living arrangement, Greek membership, and past-year alcohol use. In the current study, the strongest effect of off-campus housing appears to be on the odds of membership as a multi-game player relative to membership as a non-gambler. It is important to note that it is possible that this finding was confounded with age; younger participants were both less likely to gamble and were more likely to live in a dorm. In future research, school year should be included in this analysis to control for the effect of age.

The strongest effects on latent class membership were arguably seen for Greek membership and moderate and heavy past-year alcohol use. The effects, however, were considerably different between men and women. LaBrie et al. (2003) found a variety of college-related characteristics to be predictive of gambling participation, including living arrangements and sorority or fraternity membership. In the current study, it appears that Greek membership was predictive of type of gambling, but that the predictive effect was variable. For men, Greek membership was related to increased odds of lighter types of gambling (cards and lotto, cards and games of skill) relative to no gambling, but decreased odds of multi-game gambling. Comparatively, for women Greek membership was related to decreased odds of all types of gambling relative to no gambling.

It appears, then, that Greek membership may be a risk factor for lighter types of gambling for men, but a protective factor for women and for heavier types of gambling for men. It may be that men play cards or bet on games of personal skill (like basketball or drinking games) as recreational activities within their fraternities, but that women have so many other social activities going on within their sororities they do not have the time or need to gamble. In addition, it may be that sororities at the university from which data was collected have stricter rules regarding drinking in the sorority houses than do the fraternities, and the sororities tend to abide by the rules better than the fraternities. It may be, then, that drinking and gambling go hand-in-hand as recreational activities in the fraternity houses. And, anecdotally, women do not tend to play cards in the sorority houses. The way in which Greek membership is related to gambling, and how this may differ between men and women, needs to be more closely examined in future research.

The predictive effects of moderate and heavy past-year alcohol use are especially interesting. For men, moderate and heavy past-year alcohol use were related to slightly increased odds of lighter types of gambling (cards and lotto, cards and games of skill) relative to no gambling, and both were related to highly increased odds of multi-game gambling. For women the predictive effect was somewhat different. For women, moderate and heavy past-year alcohol use were related to slightly increased odds of cards and lotto playing relative to no gambling, and both were related to decreased odds of cards and games of skill playing. In addition, moderate and heavy past-year alcohol use were related to highly increased odds of multi-game gambling. It appears that past-year alcohol use may be an important risk factor for heavier types of gambling where an individual is engaging in many types of gambling activities, especially for women. The way in which specific types of gambling behaviors are related to specific types of drinking behaviors, and how this relation differs between men and women, needs to be more closely examined in future research.

The strong effects of moderate and heavy alcohol use on membership as a multi-game player have important implications for prevention and treatment research in general. Individually, problems with gambling and drinking result in a variety of intra- and interpersonal problems, and financial difficulties. The combination of gambling and drinking, however, has the potential to exponentially increase the likelihood of severe, long-term negative consequences. It is easily imagined that individuals who drink heavily when gambling may have irresponsible gambling behavior, which leads to intra- or interpersonal problems and/or financial stress. This stress may then lead to heavier drinking to escape depressive feelings, or it may lead to heavier gambling to recover financial losses, or both. It may be that prevention and treatment programs need to target both behaviors in order to be effective for individuals at highest risk of developing problems with gambling. The ways in which gambling and drinking behaviors may interact to increase the likelihood of negative outcomes in a variety of domains needs to be explored in order to develop prevention and treatment programs that are effective at reducing both behaviors.

2.4.3 Implications for Prevention

Overall, the main implication for prevention offered by the results discussed here is that identifying types of gamblers based on the activities they play may be a good strategy to employ when trying to identifying individuals to target with prevention programs. It is likely that individuals who play a variety of types of gambling activities, like members of the cards and games of skill and multi-game latent classes, may be at increased risk for developing problems with gambling, regardless of the number of diagnostic criteria they meet. Members of these two latent classes play gambling activities that are: (1) arguably more problematic than very common gambling activities like playing cards and lotto, and (2) of concern to researchers and policy makers like internet gambling. In addition, membership in these latent classes is predicted by behaviors like heavy alcohol use that have been shown to be correlated with problem and pathological gambling.

The goal of prevention is to intervene in the early stages of potentially problematic behavior, which in this case may be identified by eclectic gambling behavior. The implication is that it may be important for prevention researchers move beyond the traditional method of identifying potential program participants with only the diagnostic criteria, as discussed in the introduction. Identifying potential participants by taking into account the variety of activities an individual plays may make it possible to intervene earlier, at least prior to an individual meeting diagnostic criteria. This strategy may be improved and become even more relevant for determining high-risk individuals when the frequency and intensity of gambling behavior are included.

2.4.4 Including Gender as a Grouping Variable

It is important to briefly discuss the differences between including gender as a grouping variable and including gender as a covariate in the latent class and conditional latent class models. Including gender as a grouping variable allows the latent class membership probabilities, item-response probabilities, and predictive effects of the other predictors to vary between men and women. When gender is included as a covariate, the latent class membership probabilities and item-response probabilities are not allowed to vary between men and women, and the predictive effect of gender on latent class membership is estimated and tested for significance. In this case, the effect of gender is interpreted as the increase (or decrease) in odds of membership in a latent class relative to membership in the reference latent class for men relative to women. This method is not appropriate when the latent class structure varies across the levels of the predictor.

The previous research discussed in the introduction about gender differences in gambling behavior suggested that there was reason to believe that the latent class structure of gambling was not the same for men and women. In addition, preliminary analyses conducted separately for men and women suggested that indeed the class structure was somewhat different by gender. Thus, it was deemed necessary to include gender as a grouping variable to allow parameter estimates to vary between men and women as needed, and to allow differences in the effects of the predictors between men and women to be explored.

Greek membership and past-year alcohol use appear to be stronger predictors of latent class membership for women than for men. The differences in effects of Greek membership and past-year alcohol use between men and women illustrate an advantage of using gender as a grouping variable instead of including it as a covariate. By including gender as a grouping variable, it was possible to see differences in the predictive effects of the other demographic characteristics and substance use behaviors. Had gender not been included as a grouping variable, the extremely large effects of moderate and heavy past-year alcohol use for women may have been obscured unless a gender by alcohol use interaction was included in the models. These results suggest that it may be particularly important for prevention and treatment researchers to understand the link between gambling and drinking behavior for women.

2.4.5 Limitations

There are four main limitations to the current study. The first is that the sample used in the current study was relatively small. It is likely that the problems with model identification mentioned earlier were due to a small number of participants in comparison to the number of cells in the contingency table ($2^{13} = 8,192$). Although the selected model made sense when considering the observed responses and theory, it was somewhat unstable as evidenced by multiple solutions being produced when different sets of random starting values were used.

A related issue is the relatively small proportions of participants in some of the latent classes (notably multi-game players for men and women, and cards and games of skill players for women). Combined with relatively small proportions of participants endorsing some levels of the predictors, odds ratios provide limited information due to the small sample sizes on which the odds ratios are based. In general, the results presented here should be interpreted cautiously.

The second limitation is one of generalizability. Due to the recruitment strategy of the original study, it is unclear the degree to which the results from the current study are generalizable to the entire undergraduate student population of a large northeastern university or to college students in general. Generalizing the results presented here to larger populations of college students should be done cautiously. In addition, a common critique of current college-student gambling research that applies to the current study is that studies do not often compare results with those of individuals that are the same age but not in college. This critique also pertains to many studies on college-student binge drinking.

The third limitation is that only one predictor was examined at a time in the conditional latent class models. The estimated effect of each predictor did not control for effects of other predictors. It was not possible to include all significant predictors in the same model due to small sample sizes (men) and small latent class membership probabilities (women). The small sample sizes and small latent class membership probabilities led to sparseness issues and it was not possible to reliability estimate the effects when many predictors were included in a single model. The models discussed in the current study are for exploratory purposes and further investigation is needed to estimate the effect of a predictor controlling for other predictors.

The fourth limitation is that the frequency and intensity of gambling behavior were not included in the identification of gambling types. For example, one participant identified as a cards and lotto player may have played cards once and bought one lottery ticket in the past year whereas another participant identified as a cards and lotto player may have played cards and bought lottery tickets several times a week in the past year. Alternatively, one participant identified as a multi-game player may have lost only a few dollars per activity in the past year whereas another participant identified as a multi-game player may have lost hundreds of dollars per month. Frequency and intensity of gambling behavior may be relevant for determining which individuals are at highest risk for developing gambling problems. When using the results discussed above to draw implications for prevention it is important to remember that latent class members may have different levels of frequency and intensity of activity participation.

2.4.6 Conclusions

The current study provides support for the idea that types of gamblers can be identified on the basis of the gambling activities in which they engage. This is particularly important for prevention and treatment researchers who currently need additional ways of identifying individuals at potential risk for developing problems with gambling. In addition, it illustrates the application and importance of CLCA to identifying predictors of types of gambling behavior. The current study suggests that future research needs to carefully examine how specific demographic characteristics and substance use behaviors are related to specific patterns of gambling behavior, and that it may be necessary to combine type and frequency/intensity to really understand how gambling problems develop.

CHAPTER 3

Empirical Study #2: Examining the Relation Between Gambling and Drinking Trajectories Among Adolescents and Young Adults

3.1 Introduction

As discussed in the previous chapter, gambling is becoming a particularly important public health concern for today's adults and youth. Research has begun to show that the initiation of gambling may take place during late childhood and early adolescence for a majority of individuals (Jacobs, 2000; Ladouceur et al., 1994). In addition, one of the most important correlates of gambling and problems with gambling is the use of alcohol and other substances (Stinchfield & Winters, 1998;

National Research Council, 1999; Wallisch, 1993; Kaminer & Haberek, 2004; Petry & Tawfik, 2001; Winters et al., 2002), behaviors also common during adolescence. It may be particularly important, then, to understand the nature of the relation between gambling and alcohol use long before individuals enter college.

In general, research on the relation between gambling and drinking during adolescence has shown that: (1) drinking tends to co-occur with gambling (Stinchfield & Winters, 1998; National Research Council, 1999; Wallisch, 1993; Petry & Tawfik, 2001; Winters et al., 2002); (2) problems with drinking tend to co-occur with problems with gambling (Kaminer & Haberek, 2004; Hardoon & Derevensky, 2002); and (3) gamblers drink at higher rates and more regularly than non-gamblers (Kassinove et al., 2000; Griffiths & Sutherland, 1998). A deeper understanding of how specific gambling behaviors are linked to specific alcohol use behaviors has not yet been reached. For example, it is not known how participation in certain types of gambling is linked to longitudinal patterns of frequent or heavy drinking.

There has been virtually no research exploring how gambling may be linked to longitudinal patterns of drinking behavior. There are, however, three prospective longitudinal studies that have looked at the development of problem gambling over time whose results are particularly relevant to the current study. The first is the work of Barnes et al. (1999, 2002, 2005). Their work has shown that gambling and drinking are both prevalent among high school students and that they share common demographic, psychological, social, and problem behavior predictors. This work suggested that trajectories of gambling behavior over time included flat-low, increasing, flat-medium, flat-high, and decreasing trajectories. Misuse of alcohol among males predicted increasing gambling and stability of high rates of gambling over time, and lower levels of alcohol misuse predicted decreasing gambling. This work, however, did not examine how longitudinal patterns of alcohol use were related to specific types of gambling behavior.

The second is the work of Vitaro et al. (1998, 2001, 2004). Their work has suggested that problem gambling and problem substance use may develop simultaneously during adolescence due to an impulse-control deficit origin. In addition, this work has shown that gambling does not explain increases in substance use, but that alcohol and other drug use predicts increases in gambling-related problems. Finally, Vitaro et al. have used PROC TRAJ to identify trajectories of gambling behavior among males. Three trajectories were identified: a low gambling trajectory followed by most participants, a chronic high trajectory, and a late-onset trajectory. Alcohol and other substance use were not discussed as predictors of trajectory membership. In general, this work has focused primarily on males and, similar to the first study discussed, did not examine how longitudinal patterns of alcohol use were related to specific gambling behaviors.

The third is the work of Winters and Anderson (2000) and Winters et al. (2002, 2005). These studies explored: (1) changes in the rates of problem and pathological gambling among a sample of adolescents who were followed into young adulthood, and (2) the patterns of onset and desistance of problems with gambling. Results from these studies have highlighted the need to examine individual trajectories of problems with gambling over time. These studies, however, did not look at normative gambling behaviors nor did they examine how alcohol and other substance use were related to patterns of problem and pathological gambling. MVLCM may provide a new way to look at comorbidity by linking specific types of gambling to specific patterns of drinking over time.

3.1.1 The Current Study: The Relation Between Adolescent and Young Adult Gambling and Drinking

In the first empirical study, types of college-student gamblers were identified using indicators of participation in a variety of gambling activities. In addition, a variety of demographic and substance use behaviors were used to predict gambling latent class membership using CLCA. This second empirical study takes a slightly different modeling approach and specifically examines the relation between gambling and longitudinal patterns of drinking.

In this study, data from the National Longitudinal Study of Adolescent Health (Add Health; Udry, 2003) are used to explore the extent to which gambling and drinking are related. The current study addresses the question, "How does gambling latent class membership relate to patterns of drinking behavior over time?" This question is addressed by: (1) identifying types of gamblers using LCA; (2) identifying longitudinal patterns of drinking behavior over time using RMLCA; and (3) describing the relation between gambling and drinking using MVLCM.

3.2 Methods

3.2.1 Participants

Data are provided by Add Health (Udry, 2003). Add Health is a nationally representative study that was designed to collect data about the causes of health-related behaviors of adolescents in middle school and high school. Data were collected on a sample of 7^{th}- through 12^{th}-graders in 1994, 1996, and 2001/2002. The public-use core data of Add Health were used for the current study. These data are publicly available and the sampling frame for the core data collection was designed to be nationally representative of the target population in 1994. The sample for the current study was limited to participants present at the third wave of data collection in 2001/2002 who responded to at least one question about their gambling behavior ($N = 4,536$; 46.4% male, 53.6% female). Data on drinking collected at all three waves were used; data on gambling were available only at the third wave.

A characteristic of the sample used in the current study that may have an important impact on the results is the distribution of age. The distribution of age at wave 3 is shown in Table 3.1. As Table 3.1 shows, participants ranged in age from 18 to 28 at the end of the study, meaning participants ranged in age from 10 to 20 at the beginning of the study. It was not expected that individuals at the beginning of middle school and individuals at the end of high school had similar patterns of gambling or drinking. In addition, there was no particularly good way of dealing with the five-year gap in data collection between the second and third waves during which participants became of legal age to drink and gamble at different points in time. It is likely, then, that age is a potentially serious confounder.

Exploring age differences is not the primary purpose of the current study. The primary purpose is to explore how gambling and drinking are related using MVLCM. It is unclear how large a sample is needed to fit the types of models used here, but overly sparse contingency tables should be avoided as much as possible. Therefore, it was desirable to use as large a sample as possible, regardless of possible age differences in the composition of the latent classes. However, the issue of possible age differences in the structure of gambling and drinking, or in the relation between gambling and drinking should not be ignored.

Table 3.1. Distribution of Age

	Age	Percent
Younger	18	0.75
	19	11.04
	20	15.83
Older	21	16.31
	22	17.48
	23	17.02
	24	15.59
	25	4.85
	26	0.84
	27	0.26
	28	0.02

As a preliminary way of accounting for age differences as much as possible, and to maintain as large a sample size as possible for the majority of the analyses, all analyses were conducted for two sets of participants. All analyses were conducted once with the total sample ($N = 4,536$), and then they were conducted again for "young" participants ($N = 1,253$), defined as participants under the age of 21 at the third wave of data collection. Conducting the analyses in this way allowed for limited comparison to see if the results from younger participants, for whom both behaviors were illegal during the entire length of the study, noticeably differed from those from the total sample.

3.2.2 Measures

3.2.2.1 Gambling

Gambling activity participation was measured at wave 3 using three binary indicators corresponding to three types of gambling activities. Participants were considered to be "players" of an activity if they reported having engaged in the activity at least once during their lifetimes; participants were considered to be "non-players" of an activity if they reported not having engaged in the activity during their lifetimes.

The first gambling activity was playing the lottery, which included lottery tickets of any type (daily, scratch-off, lotto). The second gambling activity was gambling at a casino, which included any type of casino tables or games (craps, blackjack, roulette, slot machines, video poker). The third gambling activity was other gambling, which included any other gambling activities (cards, BINGO, horse racing, sporting events, other). The proportions of participants engaging in each gambling activity are shown in Table 3.2. As Table 3.2 shows, playing the lottery was the most common gambling activity in the total sample and for young participants. Other gambling activity was the least common gambling activity in the total sample, but gambling at a casino was the least common for young participants.

Table 3.2. Lifetime Gambling Activities

	Percent Playing	
Activity	Total Sample	Young Participants
Lottery tickets: Daily, Scratch-offs, or Lotto	60.0	49.8
Casino tables/games: Craps, blackjack, roulette, slot machines, video poker	44.3	26.2
Other games: Cards, BINGO, horse racing, sporting events, other	40.7	36.1
N	4536	1253

3.2.2.2 Drinking

Three kinds of drinking behavior were of interest: drinking frequency, drinking quantity, and a combination of frequency and quantity, labeled “intensity” for the purposes of the current study. Drinking behavior was examined in these different ways because it is unknown which dimensions of drinking behavior are most important to gambling behavior. Although it is beyond the scope of the current study to develop a rich, multidimensional model of drinking behavior over time, this is an attempt to recognize the complex nature of drinking behavior and explore the ways in which different dimensions of drinking are related to gambling.[1] All three kinds of drinking were measured at waves 1, 2, and 3.

3.2.2.2.1 Drinking Frequency Frequent drinking was defined as drinking more than once per month in the past 12 months. A total of three binary indicators of frequent drinking, one from each wave of data, were used to measure drinking frequency. Endorsement of each indicator indicated frequent drinking in the past year (regardless of amount drank); non-endorsement indicated infrequent (or no) drinking in the past year. The proportions of participants endorsing each indicator of drinking frequency at waves 1, 2, and 3 are shown in Table 3.3.

3.2.2.2.2 Drinking Quantity Heavy drinking was defined as drinking three or more drinks per drinking day on average in the past 12 months. A total of three binary indicators of heavy drinking, one from each wave of data, were used to measure drinking quantity. Endorsement of each indicator indicated heavy drinking in the past year (regardless of how often drinking occurred); non-endorsement indicated light (or no) drinking in the past year. The proportions of participants endorsing each indicator of drinking quantity at waves 1, 2, and 3 are shown in Table 3.3.

[1] Developing a rich, multidimensional model of drinking behavior over time was beyond the scope of the current study because preliminary analyses that combined drinking frequency and quantity into the same model suggested that a very large number of latent classes would be needed to describe patterns of behavior over time. To identify models that were more easily interpreted and that could then be related to gambling, it was decided that drinking frequency and quantity should be examined separately. Models that include both drinking frequency and quantity simultaneously, however, should be examined in future research.

Table 3.3. Drinking Frequency, Quantity, and Intensity Indicators

		Percent Endorsing	
Activity		Total Sample	Young Participants
Frequency:			
Drank more than once a month	Wave 1	17.4	7.1
in the past 12 months	Wave 2	18.5	9.3
	Wave 3	45.9	38.4
Quantity:			
Drank three or more drinks on	Wave 1	28.6	10.7
average when drinking in the	Wave 2	30.9	19.3
past 12 months	Wave 3	50.7	52.9
Intensity:			
Drank more than once a month	Wave 1	13.9	4.5
and drank three or more drinks	Wave 2	15.5	7.3
on average when drinking in	Wave 3	36.1	34.2
the past 12 months			
N		4536	1253

3.2.2.2.3 Drinking Intensity High intensity drinking was defined as drinking three or more drinks on average more than once per month in the past 12 months. A total of three binary indicators of high intensity drinking, one from each wave of data, were used to measure drinking intensity. Endorsement of each indicator indicated frequent heavy drinking in the past year; non-endorsement indicated more infrequent or lighter (or no) drinking in the past year. The proportions of participants endorsing each indicator of drinking intensity at waves 1, 2, and 3 are shown in Table 3.3.

3.2.3 Overview of Models Fit in the Current Study

Before the primary objective of the current study was addressed with MVLCM, individual models of gambling and patterns of drinking over time were developed. All models were developed for the total sample and for young participants separately.[2]

First, LCA was used to identify types of adolescent and young adult gamblers using the three indicators of gambling activity participation. Second, RMLCA was used to identify latent classes describing longitudinal patterns of drinking frequency, quantity, and intensity. In RMLCA, the three binary indicators of each kind of drinking separately were used to identify latent classes that were interpreted as drinking trajectories over time. For example, the three binary indicators of frequent drinking were used to identify latent classes of trajectories of drinking frequency. An example latent class in this model may be defined by high probabilities of endorsing frequent drinking at all three waves; this class may be labeled "consistent frequent drinkers" because participants who were members of that class were consistently drinking frequently during the course of the study. Similar interpretations stem from the models of drinking quantity and intensity trajectories.

Absolute model fit assessment for the latent class and latent class models for repeated measures were based on the G^2 fit statistic and the interpretations of the

[2] It should be noted that in the first empirical study, large gender differences were seen in the relation between gambling and drinking. A limitation of the models presented here is that they do not account for gender; future work with these data must account for both gender and age differences. This limitation is discussed in more detail in the discussion section.

models. Model selection was based on the AIC and BIC.

Third, to address the primary objective of the current study, the relation between gambling and patterns of drinking over time was explored using MVLCM. In any one multivariable latent class model, gambling and drinking may be independent or they may be dependent. If gambling and drinking are independent, they are unrelated; if gambling and drinking are dependent, they are related. The loglinear notation for the structural part of the independence model is (D, G), where "D" represents drinking frequency, quantity, or intensity trajectories and "G" represents gambling. The loglinear notation for the structural part of the dependence model is (DG).

Both the independence and dependence models were estimated for the three combinations of gambling and drinking: gambling and drinking frequency trajectories, gambling and drinking quantity trajectories, and gambling and drinking intensity trajectories. The independence and dependence models were compared to determine whether gambling and drinking were related. Results from the dependence model were used to describe the relation between gambling and drinking. In addition, results from the dependence model were compared with the results from the independence model to see how the results from the dependence model differed from what would be expected if gambling and drinking were independent.

All latent class and latent class models for repeated measures were estimated using PROC LCA. The multivariable latent class models that related gambling and drinking frequency, quantity, and intensity trajectories were estimated using M*plus*. M*plus* is a statistical modeling program that allows the estimation of a variety of models that include single-level and multilevel data, combinations of manifest and latent variables, and combinations of continuous and categorical variables. M*plus* is available by paid order from the M*plus* website at `http://www.statmodel.com/`. Examples of the SAS and M*plus* programming code used to fit the models discussed above may be found in Appendix B.

3.3 Results

3.3.1 Gambling Latent Class Models

First, LCA was used to identify types of adolescent and young adult gamblers for the total sample and separately for young participants. Models with 2 to 8 latent classes were investigated for each set of participants. All models were restricted to estimate two measurement parameters; one parameter was used to estimate a "high" probability of endorsing a gambling activity conditional on latent class membership, the other was used to estimate a "low" probability. Even with these restrictions on the measurement, models with more than 6 latent classes were not fit because they required the estimation of more parameters than the number of available *df*.

Starting with the 6-class model and moving to models with fewer latent classes, the G^2 fit statistic, AIC, BIC, and substantive interpretations of the models were investigated to select a model identifying types of adolescent and young adult gamblers. When the AIC and BIC did not point to the same model, the model with the lowest BIC was selected if it also corresponded to the greatest drop in the AIC. Fit statistics for the gambling latent class models, for the total sample and for young participants, are shown in Table 3.4.

For the total sample, the 5-class model was selected ($G^2 = 3.14$, $df = 1$, AIC = 15.14, BIC = 53.66). This model is characterized by classes of: (1) "non-gamblers," (2) "lotto" players, (3) "lotto and casino" players, (4) "lotto and other" players, and (5) "all" players who played all three activities. For young participants, the 4-class model was selected ($G^2 = 7.17$, $df = 2$, AIC = 17.17, BIC = 42.83). This model is characterized by classes of "non-gamblers," "lotto" players, "lotto and other" players, and "all" players.

The latent class membership probabilities for the 5- and 4-class models for the total sample and for young participants, respectively, are shown in Table 3.5. The item-response probabilities of playing an activity given latent class membership for the selected models are shown in Table 3.6. Item-response probabilities restricted to be equal during estimation are denoted with the same letter in the table. The classes were named by examining for which of the three gambling activities there was a high item response probability of endorsing a gambling activity within the

Table 3.4. Fit Statistics for Models of Gambling

Total Sample	Model	G^2	*df*	AIC	BIC
	4-class	58.36	2	68.36	100.46
	5-class	**3.14**	**1**	**15.14**	**53.66**
	6-class	0.00	0	14.00	58.94
Young Participants	Model	G^2	p	AIC	BIC
	3-class	83.12	3	91.12	111.65
	4-class	**7.17**	**2**	**17.17**	**42.83**
	5-class	4.15	1	16.15	46.95

Note: Boldface indicates the selected model.

Table 3.5. Latent Class Membership Probabilities for Gambling

	Latent Class Membership				
	None	Lotto	Lotto & Casino[a]	Lotto & Other	All
Total Sample	0.336	0.148	0.100	0.058	0.358
Young Participants	0.434	0.179	—	0.140	0.247

Note: Probabilities sum to 1.0 across rows.
[a] Latent class not applicable for young participants.

Table 3.6. Item Response Probabilities for Gambling: Probability of Playing an Activity*

Total Sample	Latent Class Membership				
Gambling Indicator	None	Lotto	Lotto & Casino	Lotto & Other	All
Lotto	0.090^{a}	0.857^{b}	0.857^{b}	0.857^{b}	0.857^{b}
Casino	0.090^{a}	0.090^{a}	0.857^{b}	0.090^{a}	0.857^{b}
Other	0.090^{a}	0.090^{a}	0.090^{a}	0.857^{b}	0.857^{b}

Young Participants	Latent Class Membership			
Gambling Indicator	None	Lotto	Lotto & Other	All
Lotto	0.085^{c}	0.806^{d}	0.806^{d}	0.806^{d}
Casino	0.085^{c}	0.085^{c}	0.085^{c}	0.806^{d}
Other	0.085^{c}	0.085^{c}	0.806^{d}	0.806^{d}

Note: Probabilities superscripted with the same letter were restricted to be equal during estimation.
* Probability of playing an activity given latent class membership.

latent class. For example, lotto and casino players had a high probability of endorsing the indicators of lottery and casino gambling, but a low probability of endorsing other gambling.

As shown in Table 3.5, in the total sample, participants were approximately equally likely to be non-gamblers and gamblers who engaged in all types of gambling. Compared to the total sample, young participants were more likely to be non-gamblers and to be lotto and other players, and less likely to engage in all types of gambling.

3.3.2 Drinking Latent Class Models for Repeated Measures

Second, RMLCA was used to identify adolescent and young adult trajectories of drinking frequency, quantity, and intensity, for the total sample and separately for young participants. Models with 2 to 8 latent classes were investigated for drinking frequency, quantity, and intensity, for each set of participants. All models were restricted to estimate two measurement parameters; one parameter was used to estimate a high probability of endorsing a drinking frequency, quantity, or intensity indicator conditional on latent trajectory class membership, the other was used to estimate a low probability. Even with these restrictions on the measurement, models with more than 6 latent classes were not fit because they required the estimation of more parameters than the number of available *df*. The same approach to model selection as the one described above for the gambling models was taken for all models of drinking trajectories. Fit statistics for the drinking frequency, quantity, and intensity latent trajectory class models, for the total sample and for young participants, are shown in Table 3.7.

For the total sample, the 4-class model was selected for drinking frequency (G^2 = 3.89, *df* = 2, AIC = 13.89, BIC = 45.99). This model is characterized by trajectory classes of: (1) "consistent infrequent" drinkers who did not drink frequently at any of the three waves, (2) "time 3 starters" who started drinking frequently at the third wave but did not do so previously, (3) "time 2 starters" who started drinking frequently at the second wave and continued to do so at the third, and (4) "consistent frequent" drinkers who drank frequently at all three waves.

Table 3.7. Fit Statistics for Models of Drinking Frequency, Quantity, and Intensity Trajectories

Drinking Frequency					
Total Sample	Model	G^2	*df*	AIC	BIC
	3-class	22.09	3	30.09	55.76
	4-class	**3.89**	**2**	**13.89**	**45.99**
	5-class[a]	3.89	1	15.89	54.41
Young Participants	Model	G^2	p	AIC	BIC
	3-class	19.52	3	27.52	48.05
	4-class	**4.70**	**2**	**14.70**	**40.37**
	5-class[b]	19.52	1	31.52	62.32

Drinking Quantity					
Total Sample	Model	G^2	*df*	AIC	BIC
	4-class	9.55	2	19.55	51.65
	5-class	**0.03**	**1**	**12.03**	**50.55**
	6-class	0.00	0	14.00	58.94
Young Participants	Model	G^2	p	AIC	BIC
	3-class	59.40	3	67.40	87.93
	4-class	**0.48**	**2**	**10.48**	**36.14**
	5-class	0.04	1	12.04	42.84

Drinking Intensity					
Total Sample	Model	G^2	*df*	AIC	BIC
	3-class	22.81	3	30.81	56.49
	4-class	**9.64**	**2**	**19.64**	**51.73**
	5-class	5.64	1	17.64	56.15
Young Participants	Model	G^2	p	AIC	BIC
	3-class	18.04	3	26.04	46.57
	4-class	**12.14**	**2**	**22.14**	**47.81**
	5-class[b]	18.04	1	30.04	60.84

Note: Boldface indicates the selected model.
[a] One class was estimated to be empty.
[b] Two classes were estimated to be empty.

For young participants, the 4-class model was selected for drinking frequency (G^2 = 4.70, *df* = 2, AIC = 14.70, BIC = 40.37). This model is characterized by the same latent trajectory classes as those for the total sample.

For drinking quantity, the 5-class model was selected for the total sample (G^2 = 0.03, *df* = 1, AIC = 12.03, BIC = 50.55). This model is characterized by trajectory classes of: (1) "consistent light" drinkers who did not drink heavily at any of the three waves, (2) "time 3 starters" who started drinking heavily at the third wave but did not do so previously, (3) "time 2 starters" who started drinking heavily at the second wave and continued to do so at the third, (4) "desisters" who drank heavily at the first two waves but not at the third, and (5) "consistent heavy" drinkers who drank heavily at all three waves. For young participants, the 4-class model was selected for drinking quantity (G^2 = 0.48, *df* = 2, AIC = 10.48, BIC = 36.14). This model is characterized by the same trajectory classes as those for the total sample with the exception of the "desisters," which was not identified as a latent trajectory class for young participants.

For drinking intensity, the 4-class model was selected for the total sample (G^2 = 9.64, *df* = 2, AIC = 19.64, BIC = 51.73). This model is characterized by trajectory classes of: (1) "consistent light or infrequent drinkers," who did not drink intensely at any of the three waves, (2) "time 3 starters" who started drinking heavily and frequently at the third wave but did not do so previously, (3) "time 2 starters" who started drinking heavily and frequently at the second wave and continued to do so at the third, and (4) "consistent frequent heavy" drinkers who drank heavily and frequently at all three measurement occasions. For young participants, the 4-class model was selected for drinking intensity (G^2 = 12.14, *df* = 2, AIC = 22.14, BIC = 47.81). This model is characterized by trajectory classes of: (1) "consistent light or infrequent drinkers," (2) "time 3 starters," (3) "time 2 starters," and (4) "desisters."

The latent trajectory class membership probabilities for the models of drinking frequency, quantity, and intensity, for the total sample and for young participants, are shown in Table 3.8. The item-response probabilities of endorsing an indicator given latent trajectory class membership for the selected models are shown in Tables 3.9 and 3.10. Item-response probabilities restricted to be equal during estimation are denoted with the same letter in the tables. The classes in all three

Table 3.8. Latent Class Membership Probabilities for Drinking Frequency, Quantity, and Intensity Trajectories

Frequency	Latent Class Membership				
	Consistent Infrequent	Time 3 Starters	Time 2 Starters	Consistent Frequent	
Total Sample	0.380	0.396	0.038	0.186	
Young Participants	0.412	0.464	0.052	0.072	
Quantity	**Latent Class Membership**				
	Consistent Light	Time 3 Starters	Time 2 Starters	Desisters[a]	Consistent Heavy
Total Sample	0.298	0.300	0.056	0.044	0.302
Young Participants	0.221	0.508	0.131	—	0.141
Intensity	**Latent Class Membership**				
	Consistent Infreq. or Light	Time 3 Starters	Time 2 Starters	Desisters[b]	Consistent Freq. & Heavy[a]
Total Sample	0.455	0.340	0.033	—	0.172
Young Participants	0.644	0.313	0.028	0.016	—

Note: Probabilities sum to 1.0 across rows.
[a] Latent class not applicable for young participants.
[b] Latent class not applicable for the total sample.

Table 3.9. Item Response Probabilities for Drinking in the Total Sample: Probability of Endorsing an Indicator*

	Latent Class Membership				
Frequency Indicator	Consistent Infrequent	Time 3 Starters	Time 2 Starters	Consistent Frequent	
Wave 1	0.050[a]	0.050[a]	0.050[a]	0.710[b]	
Wave 2	0.050[a]	0.050[a]	0.710[b]	0.710[b]	
Wave 3	0.050[a]	0.710[b]	0.710[b]	0.710[b]	
	Latent Class Membership				
Quantity Indicator	Consistent Light	Time 3 Starters	Time 2 Starters	Desisters	Consistent Heavy
Wave 1	0.041[c]	0.041[c]	0.041[c]	0.748[d]	0.748[d]
Wave 2	0.041[c]	0.041[c]	0.748[d]	0.748[d]	0.748[d]
Wave 3	0.041[c]	0.748[d]	0.748[d]	0.041[c]	0.748[d]
	Latent Class Membership				
Intensity Indicator	Consistent Infreq. or Light	Time 3 Starters	Time 2 Starters	Consistent Freq. Heavy	
Wave 1	0.036[e]	0.036[e]	0.036[e]	0.639[f]	
Wave 2	0.036[e]	0.036[e]	0.639[f]	0.639[f]	
Wave 3	0.036[e]	0.639[f]	0.639[f]	0.639[f]	

Note: Probabilities superscripted with the same letter were restricted to be equal during estimation.
* Probability of endorsing an indicator given latent class membership.

Table 3.10. Item Response Probabilities for Drinking in the Young Participants Sample: Probability of Endorsing any Indicator*

	Latent Class Membership			
Frequency Indicator	Consistent Infrequent	Time 3 Starters	Time 2 Starters	Consistent Frequent
Wave 1	0.025^a	0.025^a	0.025^a	0.643^b
Wave 2	0.025^a	0.025^a	0.643^b	0.643^b
Wave 3	0.025^a	0.643^b	0.643^b	0.643^b
	Latent Class Membership			
Quantity Indicator	Consistent Light	Time 3 Starters	Time 2 Starters	Consistent Heavy
Wave 1	0.015^c	0.015^c	0.015^c	0.674^d
Wave 2	0.015^c	0.015^c	0.674^d	0.674^d
Wave 3	0.015^c	0.674^d	0.674^d	0.674^d
	Latent Class Membership			
Intensity Indicator	Consistent Infreq. or Light	Time 3 Starters	Time 2 Starters	Desisters
Wave 1	0.033^e	0.033^e	0.033^e	0.937^f
Wave 2	0.033^e	0.033^e	0.937^f	0.937^f
Wave 3	0.033^e	0.937^f	0.937^f	0.033^e

Note: Probabilities superscripted with the same letter were restricted to be equal during estimation.
* Probability of endorsing an indicator given latent class membership.

drinking models were named by examining for which of the three times there was a high item response probability of endorsing drinking within the latent class; this approach was similar to the one taken in the gambling models.

As shown in Table 3.8, for drinking frequency it appears that in the total sample and in the young participants sample, participants were most likely to start drinking frequently at the third wave. This was also the case for heavy drinking. For drinking intensity, it appears that most participants were consistently not frequent heavy drinkers, but if participants were going to become frequent heavy drinkers they were most likely to do so at the third wave.

It is interesting that the latent class "desisters" of intense drinking was not present for the total sample, but was present for young participants. It is hypothesized that this is because this latent class was so small in the total sample analysis (only 1.6% of young participants were members), it was not required for a well-fitting, parsimonious model of patterns of drinking for the total sample. It makes sense that the latent class "consistent frequent and heavy" drinking was not present for young participants. Most of the young participants were not drinking intensely at wave 1 (64.4%) and additional third of them did not report intense drinking until wave 3 (31.1%), making it likely that a class for consistent frequent and heavy drinking at all three times was unnecessary.

3.3.3 Multivariable Latent Class Models

Third, to address the primary objective of the current study, multivariable latent class models specifying independence and dependence between gambling latent class membership and drinking frequency, quantity, and intensity trajectories were used to examine how patterns of drinking behavior over time relate to gambling latent class membership, for the total sample and separately for young participants. Multivariable latent class models were estimated for: (1) gambling and drinking frequency trajectories; (2) gambling and drinking quantity trajectories; and (3) gambling and drinking intensity trajectories. One multivariable latent class model specifying independence, and one specifying dependence between the two variables were estimated for each of the three variable combinations, for the total sample and separately for young participants.

Fit statistics for the multivariable latent class models specifying independence and dependence for each of the three variable combinations, for the total sample and for young participants, are shown in Table 3.11. Comparing the fit of the independence and dependence models for each of the three variable combinations, gambling appears to be highly related to patterns of drinking over time, regardless of type of drinking and regardless of age — the dependence model was always selected over the independence model.

To examine how patterns of drinking behavior over time were related to gambling latent class membership, the probabilities of gambling latent class membership conditional on drinking latent trajectory class membership were interpreted. However, it would be possible to instead interpret the probabilities of membership in a drinking latent trajectory class conditional on membership in the gambling latent classes if so desired. The probabilities of gambling latent class membership conditional on drinking latent trajectory class membership for drinking frequency, quantity, and intensity, for the total sample and for young participants, are shown in Tables 3.12, 3.13, and 3.14.

The entries of particular importance to note in Tables 3.12, 3.13, and 3.14 are the probabilities that are substantially different from what would be expected if gambling and drinking trajectories were independent. The expected probabilities are shown in the last row of each table. To facilitate interpretation, entries appearing in boldface type are more than 0.100 different from the expected probability.

Finally, it should be noted that several entries were fixed to 0.000 during estimation due to sparseness; fixed entries are labeled in the tables. An entry was fixed to 0.000 when the probability was so small that it was too difficult to estimate. For example, the probability of membership in the lotto latent class, conditional on membership in the consistent frequent drinking latent trajectory class for young participants was fixed to 0.000. This entry being fixed to 0.000 forces all young consistent frequent drinkers to be members of gambling latent classes other than the lotto latent class. This may slightly increase the other estimates in that row of the conditional probability matrix in Table 3.12. Similar interpretations may be made about the fixed entries in Tables 3.13 and 3.14.

As shown in Tables 3.12, 3.13, and 3.14, participants who were consistent infrequent, light, or not intense drinkers were most likely to be non-gamblers. This

Table 3.11. Fit Statistics for Independence and Dependence Models of Gambling and Drinking Frequency, Quantity, and Intensity Trajectories

Total Sample	Model	LL[a]	χ^2	p	AIC	BIC
Frequency	Ind.	-14953.916	530.49	11	29929.823	30000.450
	Dep.	**-14703.198**	**70.288**	**23**	**29452.396**	**29600.052**
Quantity	Ind.	-15861.097	495.361	12	31746.195	31823.232
	Dep.	**-15607.975**	**39.600**	**27**	**31269.949**	**31443.284**
Intensity	Ind.	-14339.551	493.855	11	28701.103	28771.721
	Dep.	**-14124.699**	**67.713**	**23**	**28295.397**	**28443.053**
Young Participants	**Model**	**LL[a]**	**χ^2**	**p**	**AIC**	**BIC**
Frequency	Ind.	-3702.740	120.923	10	7425.476	7476.812
	Dep.	**-3664.726**	**34.436**	**18**	**7365.452**	**7457.851**
Quantity	Ind.	-3999.547	194.598	10	8019.093	8070.426
	Dep.	**-3954.187**	**97.718**	**18**	**7944.375**	**8036.774**
Intensity	Ind.	-3519.541	178.385	10	7059.081	7110.414
	Dep.	**-3483.248**	**66.495**	**17**	**7000.495**	**7087.761**

Note: Boldface indicates the selected model.
Note: The independence model is labeled Ind.; the dependence model is labeled Dep.
[a] Loglikelihood value.

Table 3.12. Conditional Latent Class Membership Probabilities: Probability of Membership in a Gambling Latent Class Given Drinking Frequency Latent Trajectory Class

	Gambling Latent Class Membership				
Drinking Frequency Latent Class Membership	None	Lotto	Lotto & Casino[a]	Lotto & Other	All
Total Sample					
Consistent Infrequent	**0.583**	0.183	0.089	0.088	**0.057**
Time 3 Start	**0.191**	0.132	0.133	0.051	**0.493**
Time 2 Start	**0.114**	**0.283**	0.100	0.034	**0.468**
Consistent Frequent	**0.185**	0.084	0.051	0.016	**0.664**
Expected	0.336	0.148	0.100	0.058	0.358
Young Participants					
Consistent Infrequent	**0.621**	0.198	—	0.140	**0.041**
Time 3 Start	**0.288**	0.184	—	0.161	**0.366**
Time 2 Start	**0.231**	0.234	—	**0.012**	**0.522**
Consistent Frequent	0.441	0.000[b]	—	0.098	**0.461**
Expected	0.434	0.179	—	0.140	0.247

Note: Boldface entries represent probabilities that are more than 0.100 above or 0.100 below the expected probability if drinking and gambling were independent.

[a] Latent class not applicable for young participants.

[b] Entry fixed due to sparseness.

Table 3.13. Conditional Latent Class Membership Probabilities: Probability of Membership in a Gambling Latent Class Given Drinking Quantity Latent Trajectory Class

	Gambling Latent Class Membership				
Drinking Quantity Latent Class Membership	None	Lotto	Lotto & Casino[a]	Lotto & Other	All
Total Sample					
Consistent Light	**0.612**	0.148	0.087	0.041	**0.112**
Time 3 Start	0.248	0.197	0.120	0.093	0.341
Time 2 Start	0.284	**0.039**	0.102	0.150	0.424
Desist	**0.584**	0.215	0.073	0.128	0.000[b]
Consistent Heavy	**0.130**	0.111	0.096	0.011	**0.651**
Expected	0.336	0.148	0.100	0.058	0.358
Young Participants					
Consistent Light	**0.797**	0.135	—	0.069	0.000[b]
Time 3 Start	**0.312**	0.243	—	0.158	0.287
Time 2 Start	**0.329**	**0.047**	—	**0.270**	**0.354**
Consistent Heavy	**0.297**	0.133	—	0.083	**0.488**
Expected	0.434	0.179	—	0.140	0.247

Note: Boldface entries represent probabilities that are more than 0.100 above or 0.100 below the expected probability if drinking and gambling were independent.

[a] Latent class not applicable for young participants.

[b] Entry fixed due to sparseness.

Table 3.14. Conditional Latent Class Membership Probabilities: Probability of Membership in a Gambling Latent Class Given Drinking Intensity Latent Trajectory Class

	Gambling Latent Class Membership				
Drinking Intensity Latent Class Membership	None	Lotto	Lotto & Casino[a]	Lotto & Other	All
Total Sample					
Consistent Infreq. Light	**0.545**	0.178	0.101	0.062	**0.114**
Time 3 Start	**0.176**	0.138	0.116	0.070	**0.501**
Time 2 Start	**0.085**	0.221	0.111	0.064	**0.519**
Consistent Freq. Heavy	**0.173**	0.079	0.059	0.023	**0.667**
Expected	0.336	0.148	0.100	0.058	0.358
Young Participants					
Consistent Infreq. Light	0.529	0.194	—	0.130	0.147
Time 3 Start	**0.286**	0.177	—	0.162	**0.375**
Time 2 Start	**0.246**	**0.051**	—	0.145	**0.558**
Consistent Freq. Heavy	**0.559**	0.000[b]	—	0.000[b]	**0.441**
Expected	0.434	0.179	—	0.140	0.247

Note: Boldface entries represent probabilities that are more than 0.100 above or 0.100 below the expected probability if drinking and gambling were independent.
[a] Latent class not applicable for young participants.
[b] Entry fixed due to sparseness.

appears to be true for the total sample and for young participants. Conversely, participants who were frequent, heavy, or intense drinkers at any time were most likely to be members of the all types of gambling latent class. This, too, appears to be true for the total sample and for young participants. In the total sample, consistent frequent, consistent heavy, and consistent intense drinkers had the highest probabilities of any drinking latent trajectory class of being members of the all types of gambling latent class. For young participants, however, time 2 starters of frequent, heavy, or intense drinking may be at higher risk of belonging to the all types of gambling latent class than the consistent frequent, heavy, or intense drinkers.

In general, the greatest differences between the conditional gambling latent class membership probabilities and the expected probabilities if gambling and drinking were independent were seen for non-gamblers and gamblers who participated in all three types of gambling activities, for all kinds of drinking. That is, it appears that, in general, a greater proportion than expected of consistent infrequent, light, and not intense drinkers were non-gamblers, and a greater proportion than expected of consistent frequent, heavy, and intense drinkers were gamblers who engaged in all three types of gambling activities.

Finally, it appears that drinking frequency was more predictive of gambling latent class membership than was drinking quantity for the total sample. Comparatively, for young participants it appears that they were about equally predictive. Evidence for this can be found by comparing the number of conditional gambling latent class membership probabilities in the drinking frequency and quantity models that were substantially different from what would be expected if gambling and drinking were independent.

3.4 Discussion

The current study used LCA to identify types of adolescent and young adult gamblers using data from Add Health. It also used RMLCA to identify trajectory classes of drinking frequency, quantity, and intensity over time. To address the primary objective, the current study modeled the relation between gambling latent class membership and drinking latent trajectory class membership using MVLCM.

The goal of the current study was to examine how gambling latent class membership relates to patterns of drinking behavior over time.

3.4.1 Relation Between Gambling and Longitudinal Patterns of Drinking

As discussed above, drinking frequency appeared to be more predictive of gambling latent class membership than was drinking quantity. The overall pattern of results for the relation between gambling latent class membership and drinking frequency latent trajectory class membership was particularly interesting. Participants who were consistent infrequent drinkers had a 58.3% chance of not yet having engaging in any gambling behavior; this chance was 62.1% for young participants. These percentages were substantially higher than what would be expected if gambling and drinking were independent. In addition, participants and young participants who drank frequently at any point during the course of the study were substantially more likely to engage in all types of gambling, and were substantially less likely to have not gambled at all when compared to what would be expected if gambling and drinking were independent.

Interestingly, for the total sample there was an increasing trend in the chance of engaging in all types of gambling given drinking trajectory — the longer participants had been frequent drinkers, the greater their chance of engaging in all types of gambling. (This increasing trend did not strictly hold for young participants, probably due to issues of sparseness.) There was not, however, a stronger than expected relation between drinking frequency latent trajectory class membership and membership in gambling latent classes characterized by participation in one or two types of gambling activities (lotto, lotto and casino, lotto and other).

The overall pattern of results for drinking quantity trajectories was similar to that of drinking frequency trajectories for the total sample and for young participants, but the relation between gambling and drinking was less in evidence. Consistent light drinkers were more likely than expected to be non-gamblers and less likely than expected to be engaging in all types of gambling; consistent heavy drinkers were more likely than expected to be engaging in all types of gambling and less likely than expected to be non-gamblers.

The overall pattern of results for drinking intensity trajectories for the total sample was similar to that of drinking frequency trajectories. The increasing trend in the chance of engaging in all types of gambling given a drinking intensity trajectory with a longer amount of time as an intense drinker was again evident. For young participants, results were in the expected direction but sparseness appears to be a problem in this model, especially for consistent frequent heavy drinkers.

In general, it appears that the majority of participants who drank frequently, heavily, or intensely at some point during the course of the study had already participated in some type of gambling. This was not the case for participants who consistently drank infrequently, lightly, or not intensely (or not at all) — the majority of these participants reported that they had not yet gambled. In addition, as was specifically noted with drinking frequency trajectories, it appears that there was no particularly unique or strong association between specific drinking trajectories of any kind and one or two specific gambling behaviors.

3.4.2 Implications for Prevention

These findings suggest that because drinking is linked to participation in multiple types of gambling, not just playing one or two specific games, prevention and treatment programs may have to target gambling and drinking behaviors simultaneously to be effective for individuals at highest risk of developing problem or pathological gambling. If all types of drinking are linked to participation in multiple types of gambling, as the results of the current study suggest, then any kind of drinking may be an important risk factor for problems with gambling. It is also interesting that drinking frequency, regardless of quantity, appears to be most strongly related to gambling behavior in general. This suggests that prevention and treatment providers should be concerned about any amount of regular drinking behavior, even if it is low in quantity, when targeting gambling.

3.4.3 Limitations

There are several limitations to the current study that should be noted. The first is that whereas the questions asked as part of Add Health do provide a way to begin exploring how gambling behavior is related to patterns of drinking behavior over time, they do so in a limited way. It is difficult to provide a clear look at the relation between specific gambling behaviors and specific drinking behaviors within the same time frame given that questions about gambling were asked only at the third wave, and that the questions concerned behavior over a different period of time than the questions about drinking (lifetime vs. past-year). Another challenge of the gambling questions is that many gambling behaviors common to youth, like games of skill and card playing with friends and family, are all part of the "other" gambling indicator, which does not provide an opportunity for a nuanced look at gambling activity participation. In addition, the length of time between the second and third waves prevents researchers from modeling drinking development over a long period of time during which drinking behavior is highly variable.

The second limitation is that although the core sample of Add Health was originally designed to be nationally representative, and therefore highly generalizable to the population of 7^{th} to 12^{th} graders (in 1994), it remains to be seen whether or not the sample selected for the current study is nationally representative in the same way. So that gambling could be modeled, only those participants who responded to at least one gambling question at the third wave were included. At this time, it is unknown how well the results of the current study generalize.

The third limitation is that despite a fairly large sample size for the total sample and young participants analyses, there were still issues of sparseness in the multivariable latent class models. This challenge is particularly evident in the young participants models, as every multivariable latent class model for the young participants had at least one parameter fixed to 0.000 to avoid estimation problems. The sparseness issues likely stem from the drinking models because some of the latent classes in those models are particularly small. Taking a different approach to modeling drinking behavior over time may overcome some of the issues of sparseness. Conclusions that may be made about the relation between gambling and drinking intensity trajectories for young participants were especially limited.

The fourth limitation is that gender differences in the relation between gambling and drinking were not explored. The results of the first empirical study suggest that these analyses should control for gender or should be run for men and women separately. It is planned that these analyses will be conducted again with gender included as a grouping variable, similar to the approach taken in the first empirical study. This was beyond the scope of the current study due to limitations of properly including a grouping variable when programming the multivariable latent class models.

3.4.4 Conclusions

Like the first empirical study, the current study provides support for the idea that types of gamblers can be identified on the basis of the gambling activities in which they engage. Again, this is particularly important for prevention and treatment researchers who currently need additional ways of identifying individuals at potential risk for developing problems with gambling. In addition, the current study illustrates the usefulness of multivariable latent class models to developmental researchers when modeling behavior over time is the goal. Finally, although the current study does not provide a particularly nuanced look at the two behaviors, it does provide strong support for the relation between multiple dimensions of drinking behavior and engaging in multiple types of gambling behavior. It suggests that future research needs to carefully examine how specific drinking behaviors are related to specific gambling behaviors over time in order to disentangle more details about the nature of the relation between the two processes.

CHAPTER 4

Empirical Study #3: Examining Developmental Relations Between Smoking and Drinking Among Adolescents

4.1 Introduction

The first two empirical studies addressed research questions about gambling and substance use. A nice follow-up to these would be to use MPM to examine the nature of the relation between gambling development and drinking or other substance use development. Unfortunately, good longitudinal data about gambling behaviors are difficult to find because researchers are just beginning to collect such data. Instead, to illustrate the use of MPM, the developmental relation between two related problem behaviors is examined. That is, because it was not possible to acquire appropriate longitudinal data on gambling, the relation between drinking and smoking development is examined to illustrate the application of MPM. As discussed in previous chapters, gambling, drinking, and smoking are all considered to be problem behaviors and previous research has shown strong associations among them. When longitudinal data about gambling are available it is planned that this method will be used to model the relation between gambling and drinking, and between gambling and smoking.

4.1.1 Relation Between Smoking and Drinking

Research has provided consistent evidence that there is a strong association between alcohol and cigarette use for a variety of types of people, and for a variety of types of drinkers (Jackson et al., 2000a; Bien & Burge, 1990; Rohde et al., 1995; Istvan & Matarazzo, 1984). Much of this evidence, however, has stemmed from research using variable-centered methodologies to look at the comorbidity of problem alcohol use and tobacco dependence (Jackson et al., 2000b; Sher et al., 1996). In general, these methods look at average patterns of inter-individual change instead of person-specific change (Schulenberg, O'Malley, et al., 1996).

As discussed by Jackson et al. (2000a, 2000b), it is important to understand how specific alcohol use behaviors are linked to specific cigarette use behaviors. This is because, as they suggest, different relations among behaviors may provide varying implications for prevention, and may suggest a variety etiological processes at work. To do this, it is necessary to take a person-centered approach to looking at the relation between smoking and drinking. That is, it is important to look at person-specific patterns of change over time (Schulenberg, O'Malley, et al., 1996). For discussions that contrast variable-centered and person-centered approaches to substance use research, and how one documents patterns of stability and change over time in substance use, see Schulenberg, O'Malley, et al. (1996), Schulenberg, Wadsworth, et al. (1996), and Horn (2000).

Currently, there is a limited amount of work that looks at the relation between smoking and drinking development prospectively over time. However, there are at least three exceptions that have provided some information about the relation between smoking and drinking development. The first is the work of John et al. (2003). John et al. (2003) showed that smoking 30 or more cigarettes per day, smoking onset prior to age 18, and nicotine dependence were all related to alcohol dependence. However, logistic regression was used in this study to find associations between variables; it did not take a person-centered approach nor did it link smoking behaviors to changes in alcohol dependence.

The second is the work of Jackson et al. (2000a, 2000b). Jackson et al. (2000a, 2000b) examined changes in dual substance use disorder diagnoses prospectively over time. In these studies, latent class analysis and state-trait models were used to examine longitudinal patterns of alcohol and tobacco disorder diagnoses. However, this work did not look at the relation between more normative smoking and drinking behaviors. In addition, this work did not link tobacco disorder diagnoses to changes in alcohol disorder diagnoses.

The third is additional work by Jackson et al. (2005) exploring the concurrent courses of smoking and drinking using growth mixture modeling. In this study, seven classes characterized by different combinations of smoking and drinking behavior over time were identified. These classes included: non-heavy drinking and non-smoking, chronic heavy drinking and chronic heavy smoking, low drinking and heavy smoking, heavy drinking and low smoking, moderate drinking and late-onset heavy smoking, moderate drinking and developmentally-limited heavy smoking, and moderate drinking and moderate smoking. Although a person-centered approach to looking at concurrent development of smoking and drinking was taken, this study did not look at how specific developmental transitions in one behavior were linked to the other behavior. MPM may provide a new way to look at comorbidity by modeling the courses of smoking and drinking simultaneously.

4.1.2 The Current Study: The Relation Between Smoking and Drinking Development

This third empirical study examines the relation between the development of smoking and the development of drinking. In this study, data from the National Longitudinal Survey of Youth, 1997 (NLSY97; Bureau of Labor Statistics, U.S. Department of Labor) are used to explore the extent to which developmental transitions in drinking are related to smoking. The current study addresses the question, "How do developmental transitions in drinking behavior vary by type of smoking behavior?" This question is addressed by: (1) modeling the development of smoking over time using LTA; (2) modeling the development of drinking over time using LTA; and (3) describing the relation between the development of drinking and the development of smoking using MPM. Eight sets of relations between smoking and drinking are hypothesized, which correspond to eight multiprocess models that comprise the set of relevant models for consideration. The hypothesized relations range in complexity from no relation between smoking and drinking, to complex interactions of earlier smoking and drinking being related to later drinking. Each of the corresponding models, labeled Model 1 through Model 8, are discussed in detail below.

4.2 Methods

4.2.1 Participants

The NLSY97 is sponsored and directed by the U.S. Bureau of Labor Statistics and conducted by the National Opinion Research Center at the University of Chicago, with assistance from the Center for Human Resource Research at The Ohio State University. The study was designed to collect data about characteristics affecting the transition from school to the labor market. Data have been collected annually since 1997 on approximately 9,000 participants (currently, there are seven waves of data available). The original NLSY97 sample consisted of a nationally representative (in 1997) sample of 6,748 adolescents born from 1980 to 1984 (the cross-sectional sample), and a supplemental sample of 2,236 adolescents that oversamples Hispanic or Latino and black people born from 1980 to 1984 (the oversample).

Data for the current study come from all participants aged 15-16 years in 1998 who are members of the cross-sectional sample, and who responded to at least one question about their drinking or smoking behavior ($N = 2,563$; 52.1% male, 47.9% female; 49.1% age 15, 50.9% age 16). The NLSY97 has participants aged anywhere from 12 to 20 during the times used in the current study; a sample of younger participants was not selected because it was desirable to have a reasonable proportion of participants smoking and drinking; a sample of older participants was not selected because it was desirable to avoid the possibility of participants graduating high school between times 1 and 2. Data from 1998 and 1999 were used to model the drinking and smoking behavior of these participants.

4.2.2 Measures

4.2.2.1 Smoking

Two questions about cigarette use were used as indicators of smoking behavior. The first question asked about the number of days, in the past 30 days, on which a cigarette was smoked. This question was used to create a trichotomous indicator of monthly smoking frequency, with the responses corresponding to: (1) no use, (2) one to fourteen smoking days, and (3) fifteen or more smoking days. The second question asked about the number of cigarettes smoked per smoking day during the past 30 days. This question was used to create a trichotomous indicator of monthly smoking quantity, with the responses corresponding to: (1) zero cigarettes, (2) one or two cigarettes per smoking day, and (3) three or more cigarettes per smoking day. The categories for each indicator were selected based on the frequency distribution in order to create groups large enough to compare while capturing most of the variation, and on preliminary analyses used to investigate which cut-points produced the most distinct classes. Table 4.1 shows the distribution of the smoking indicator responses at times 1 and 2.

4.2.2.2 Drinking

Three questions about alcohol use were used as indicators of drinking behavior. The first question asked about the number of days, in the past 30 days, on which one or more drinks were consumed. This question was used to create a trichotomous indicator of monthly drinking frequency, with the responses corresponding to: (1) no use, (2) one or two drinking days, and (3) three or more drinking days. The second question asked about the number of drinks consumed per drinking day during the past 30 days. This question was used to create a trichotomous indicator of monthly drinking quantity, with the responses corresponding to: (1) zero drinks, (2) one or two drinks per drinking day, and (3) three or more drinks per drinking day. The third question asked about the number of days, in the past 30 days, on which five or more drinks were consumed on the same occasion (binge

Table 4.1. Smoking and Drinking Indicators

Activity	Time	Percent "Endorsing"		
Cigarette Frequency:		0 Days	1-14 Days	15+ Days
Number of days, during the past 30 days, on which a cigarette was smoked	1	75.7	11.2	12.4
	2	68.6	11.4	15.5
Cigarette Quantity:		0 Cigs	1-2 Cigs	3+ Cigs
Number of cigarettes per smoking day during the past 30 days	1	77.7	9.6	12.0
	2	69.9	9.6	16.0
Alcohol Frequency:		0 Days	1-2 Days	3+ Days
Number of days, during the past 30 days, on which one or more drinks were consumed	1	69.5	15.5	14.5
	2	59.4	17.2	18.6
Alcohol Quantity:		0 Drnks	1-2 Drnks	3+ Drnks
Number of drinks per drinking day during the past 30 days	1	70.8	13.1	15.5
	2	60.0	14.5	20.5
Binge Drinking:		No	Yes	
Consumed five or more drinks on same occasion during the past 30 days	1	85.4	14.1	
	2	75.9	19.0	

Note: Percentages may not sum to 100% due to missing data.

drinking). This question was used to create a dichotomous indicator of monthly binge drinking ("no" or "yes"). The categories for each indicator were selected based on the frequency distribution in order to create groups large enough to compare while capturing most of the variation, and on preliminary analyses used to investigate which cut-points produced the most distinct classes. Table 4.1 shows the distribution of the drinking indicator responses at times 1 and 2.

4.2.3 Overview of Models Fit in the Current Study

Before the primary objective of the current study was addressed with MPM, individual models of smoking and drinking development were developed. First, LTA was used to identify types of smokers and describe change over time in smoking. Second, LTA was used to identify types of drinkers and describe change over time in drinking. Latent transition models for smoking and drinking were assessed and selected based on the G^2 fit statistic, AIC, BIC, and substantive interpretations of the models. The validity of imposing parameter restrictions specifying measurement invariance over time and response similarity across indicators were also investigated for each model. The result was two highly restricted, individual latent transition models that described the development of adolescent smoking and drinking behavior.

Third, to address the primary objective of the current study, a variety of multiprocess models were explored. Eight possible sets of relations between smoking and drinking were hypothesized, each of which corresponds to a multiprocess model. The eight models were compared using the AIC and BIC to determine which was best at balancing fit and parsimony in its description of the relation between smoking and drinking development. A discussion of each of the eight hypothesized sets of relations and its corresponding multiprocess model is presented below. The models are labeled Model 1 through Model 8.

Graphical representations of Models 1 though 8 are shown in Figures 4.1 through 4.8, and are provided as a way to visualize the differences between the models. These graphical representations are based on the graphical representations of the M*plus* Modeling Framework (Muthén & Muthén, 1998-2006).

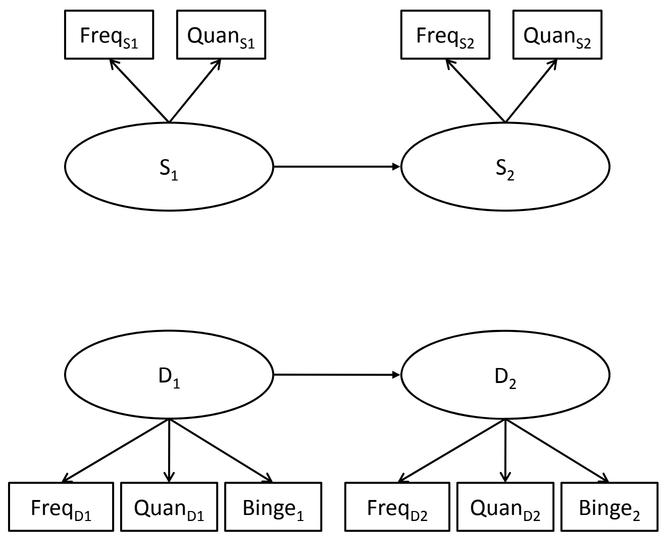
In the graphical representations, an ellipse represents a discrete latent variable (smoking or drinking behavior at a particular time); a rectangle represents a manifest indicator of smoking or drinking behavior at a particular time. An arrow represents a conditional dependency among variables and the corresponding interaction effect included in the loglinear model; the direction of the arrow specifies the direction in which the predictive effect will be interpreted. A solid arrow represents conditional dependence between two variables; a dotted arrow represents conditional dependence between three variables (the variable from which the dotted arrow originates and the two variables that are related via the solid arrow at which the dotted arrow terminates).[1] A dotted arrow specifies that the conditional relation between any two variables given the third varies across levels of the third.

In the figures, smoking at times 1 and 2 are labeled "S_1" and "S_2;" drinking at times 1 and 2 are labeled "D_1" and "D_2;" indicators of smoking frequency at times 1 and 2 are labeled "Freq_{S1}" and "Freq_{S2};" indicators of smoking quantity at times 1 and 2 are labeled "Quan_{S1}" and "Quan_{S2};" indicators of drinking frequency at times 1 and 2 are labeled "Freq_{D1}" and "Freq_{D2};" indicators of drinking quantity at times 1 and 2 are labeled "Quan_{D1}" and "Quan_{D2};" indicators of binge drinking at times 1 and 2 are labeled "Binge_1" and "Binge_2." In addition, each model is labeled with its loglinear notation for the structural part of the model for clarity.

Model 1 specifies that the development of smoking and the development of drinking are unrelated. In Model 1: (1) time 2 smoking depends on time 1 smoking; (2) time 2 drinking depends on time 1 drinking; and (3) smoking and drinking are independent. The loglinear notation for the structural part of Model 1 is (S_1S_2, D_1D_2).

Model 2 specifies that allowing a cross-sectional relation between smoking and drinking at time 1 (baseline relation) is enough to account for any observed relation between the two processes. In Model 2: (1) time 2 smoking depends on time 1 smoking; (2) time 2 drinking depends on time 1 drinking; and (3)

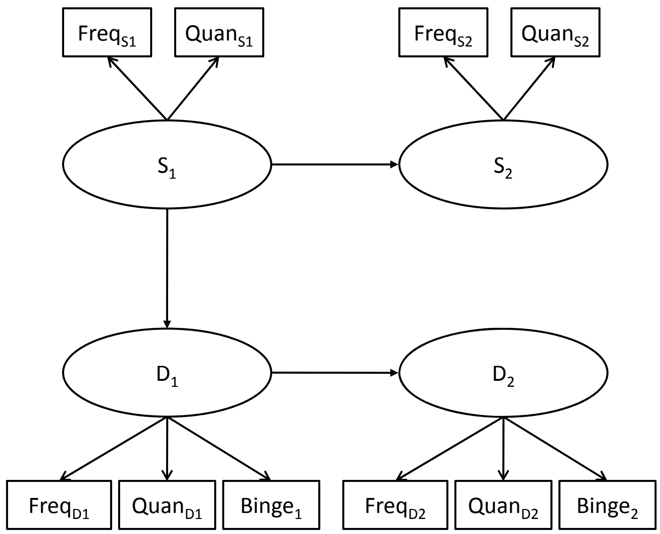
[1] Predictive relations between variables are described by odds ratios or by conditional probabilities, rather than correlations. When there is no relation between two variables, levels of the first variable are independent of levels of the second. For example, if time 1 smoking is independent of time 1 drinking, time 1 drinking latent status membership does not depend on time 1 smoking latent status membership. Statistically, this means: $P(D_1|S_1) = P(D_1)$. This is conceptually equivalent to a correlation coefficient equaling zero; arrows do not represent regression as they do in SEM diagrams.



S_1 and S_2: smoking at times 1 and 2; Freq$_{S1}$ and Freq$_{S2}$: indicators of smoking frequency at times 1 and 2; Quan$_{S1}$ and Quan$_{S2}$: indicators of smoking quantity at times 1 and 2

D_1 and D_2: drinking at times 1 and 2; Freq$_{D1}$ and Freq$_{D2}$: indicators of drinking frequency at times 1 and 2; Quan$_{D1}$ and Quan$_{D2}$: indicators of drinking quantity at times 1 and 2; Binge$_1$ and Binge$_2$: indicators of binge drinking at times 1 and 2

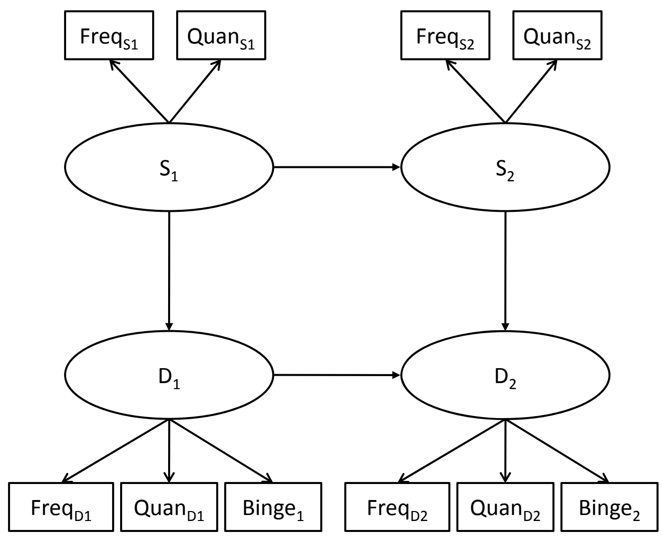
Figure 4.1. Graphical representation of Model 1. The loglinear notation for the structural part of Model 1 is (S_1S_2, D_1D_2). This model corresponds to the hypothesis that the development of smoking and the development of drinking are unrelated.



S$_1$ and S$_2$: smoking at times 1 and 2; Freq$_{S1}$ and Freq$_{S2}$: indicators of smoking frequency at times 1 and 2; Quan$_{S1}$ and Quan$_{S2}$: indicators of smoking quantity at times 1 and 2

D$_1$ and D$_2$: drinking at times 1 and 2; Freq$_{D1}$ and Freq$_{D2}$: indicators of drinking frequency at times 1 and 2; Quan$_{D1}$ and Quan$_{D2}$: indicators of drinking quantity at times 1 and 2; Binge$_1$ and Binge$_2$: indicators of binge drinking at times 1 and 2

Figure 4.2. Graphical representation of Model 2. The loglinear notation for the structural part of Model 2 is (S_1S_2, D_1D_2, S_1D_1). This model corresponds to the hypothesis that there is a cross-sectional relation between time 1 smoking and time 1 drinking (baseline relation).



S_1 and S_2: smoking at times 1 and 2; $Freq_{S1}$ and $Freq_{S2}$: indicators of smoking frequency at times 1 and 2; $Quan_{S1}$ and $Quan_{S2}$: indicators of smoking quantity at times 1 and 2

D_1 and D_2: drinking at times 1 and 2; $Freq_{D1}$ and $Freq_{D2}$: indicators of drinking frequency at times 1 and 2; $Quan_{D1}$ and $Quan_{D2}$: indicators of drinking quantity at times 1 and 2; $Binge_1$ and $Binge_2$: indicators of binge drinking at times 1 and 2

Figure 4.3. Graphical representation of Model 3. The loglinear notation for the structural part of Model 3 is $(S_1S_2, D_1D_2, S_1D_1, S_2D_2)$. This model corresponds to the hypothesis that there is a cross-sectional relation between time 1 smoking and time 1 drinking, and between time 2 smoking and time 2 drinking (concurrent relation).

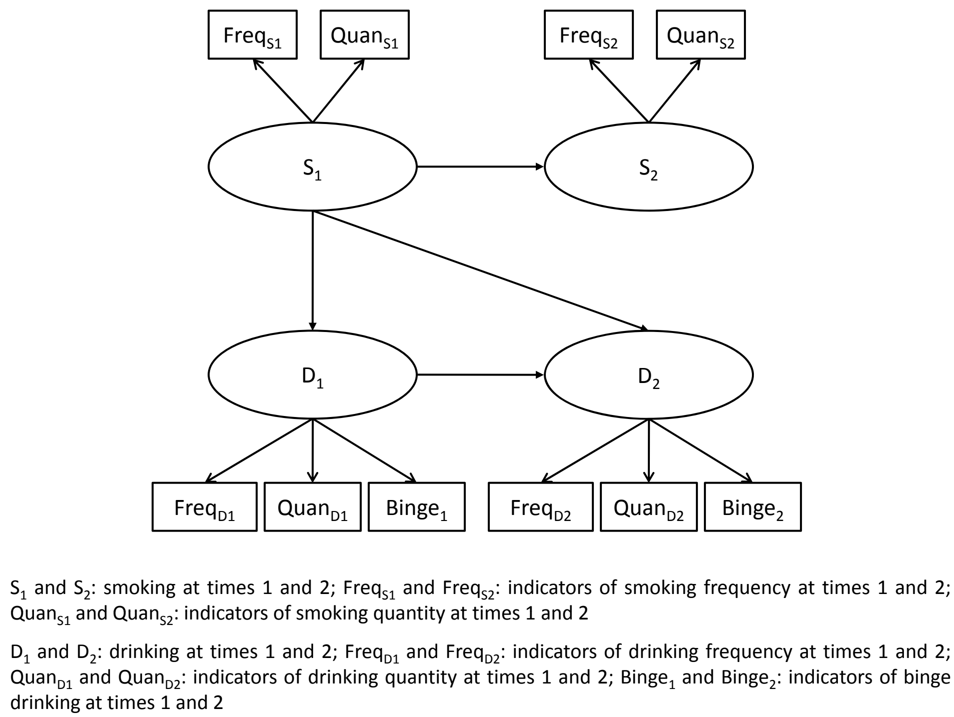
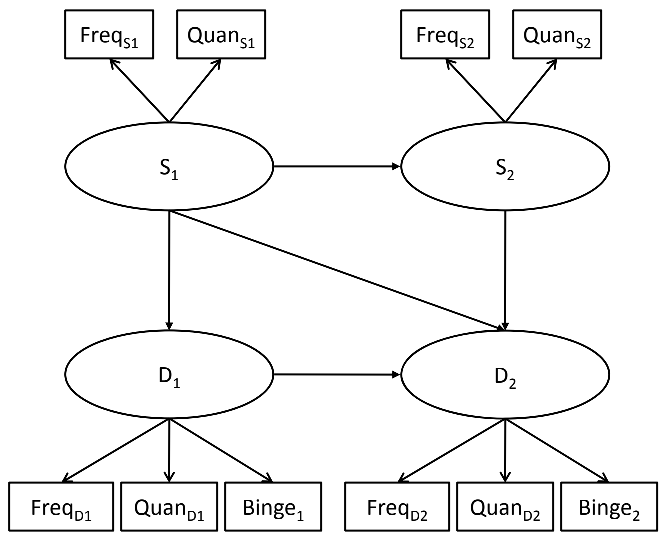


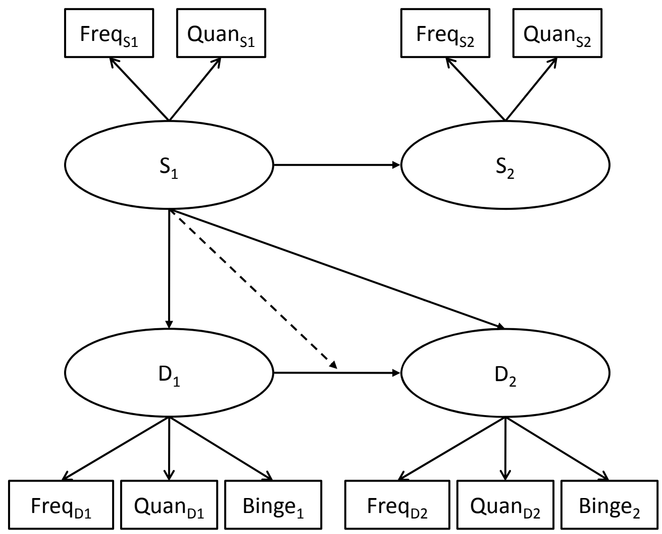
Figure 4.4. Graphical representation of Model 4. The loglinear notation for the structural part of Model 4 is $(S_1S_2, D_1D_2, S_1D_1, S_1D_2)$. This model corresponds to the hypothesis that there is a baseline and lagged relation between smoking and drinking.



S_1 and S_2: smoking at times 1 and 2; Freq$_{S1}$ and Freq$_{S2}$: indicators of smoking frequency at times 1 and 2; Quan$_{S1}$ and Quan$_{S2}$: indicators of smoking quantity at times 1 and 2

D_1 and D_2: drinking at times 1 and 2; Freq$_{D1}$ and Freq$_{D2}$: indicators of drinking frequency at times 1 and 2; Quan$_{D1}$ and Quan$_{D2}$: indicators of drinking quantity at times 1 and 2; Binge$_1$ and Binge$_2$: indicators of binge drinking at times 1 and 2

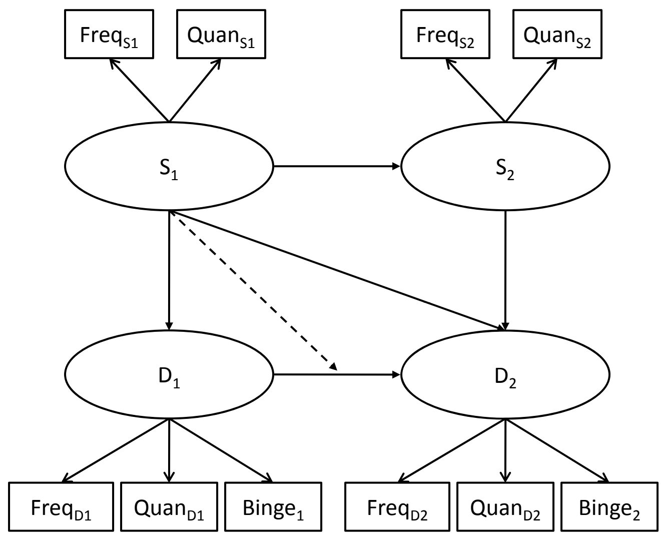
Figure 4.5. Graphical representation of Model 5. The loglinear notation for the structural part of Model 5 is $(S_1S_2, D_1D_2, S_1D_1, S_2D_2, S_1D_2)$. This model corresponds to the hypothesis that there is a concurrent and lagged relation between smoking and drinking.



S_1 and S_2: smoking at times 1 and 2; Freq$_{S1}$ and Freq$_{S2}$: indicators of smoking frequency at times 1 and 2; Quan$_{S1}$ and Quan$_{S2}$: indicators of smoking quantity at times 1 and 2

D_1 and D_2: drinking at times 1 and 2; Freq$_{D1}$ and Freq$_{D2}$: indicators of drinking frequency at times 1 and 2; Quan$_{D1}$ and Quan$_{D2}$: indicators of drinking quantity at times 1 and 2; Binge$_1$ and Binge$_2$: indicators of binge drinking at times 1 and 2

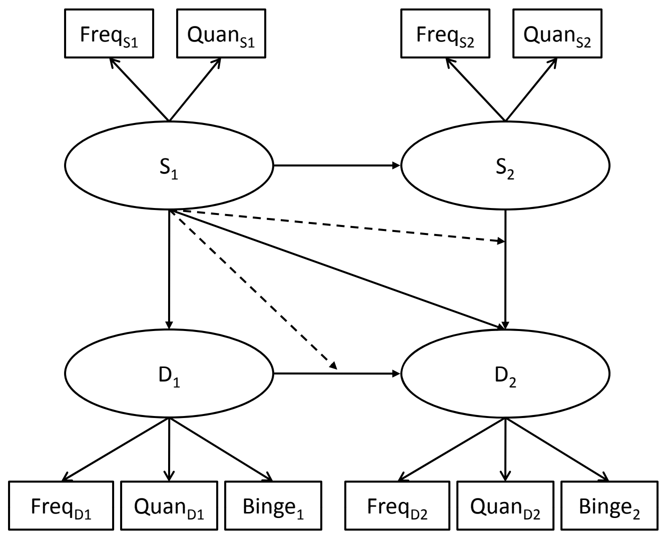
Figure 4.6. Graphical representation of Model 6. The loglinear notation for the structural part of Model 6 is $(S_1D_1D_2, S_1S_2)$. This model corresponds to the hypothesis that there is a baseline relation and an interaction between time 1 smoking and time 1 drinking.



S_1 and S_2: smoking at times 1 and 2; $Freq_{S1}$ and $Freq_{S2}$: indicators of smoking frequency at times 1 and 2; $Quan_{S1}$ and $Quan_{S2}$: indicators of smoking quantity at times 1 and 2

D_1 and D_2: drinking at times 1 and 2; $Freq_{D1}$ and $Freq_{D2}$: indicators of drinking frequency at times 1 and 2; $Quan_{D1}$ and $Quan_{D2}$: indicators of drinking quantity at times 1 and 2; $Binge_1$ and $Binge_2$: indicators of binge drinking at times 1 and 2

Figure 4.7. Graphical representation of Model 7. The loglinear notation for the structural part of Model 7 is $(S_1D_1D_2, S_1S_2, S_2D_2)$. This model corresponds to the hypothesis that there is a concurrent relation and an interaction between time 1 smoking and time 1 drinking.



S_1 and S_2: smoking at times 1 and 2; $Freq_{S1}$ and $Freq_{S2}$: indicators of smoking frequency at times 1 and 2; $Quan_{S1}$ and $Quan_{S2}$: indicators of smoking quantity at times 1 and 2

D_1 and D_2: drinking at times 1 and 2; $Freq_{D1}$ and $Freq_{D2}$: indicators of drinking frequency at times 1 and 2; $Quan_{D1}$ and $Quan_{D2}$: indicators of drinking quantity at times 1 and 2; $Binge_1$ and $Binge_2$: indicators of binge drinking at times 1 and 2

Figure 4.8. Graphical representation of Model 8. The loglinear notation for the structural part of Model 8 is $(S_1D_1D_2, S_1S_2D_2)$. This model corresponds to the hypothesis that there is a concurrent relation, an interaction between time 1 smoking and time 1 drinking, and an interaction between time 1 smoking and time 2 smoking. That is, it is hypothesized that time 2 drinking is explained by combinations of time 1 smoking and time 1 drinking, and transitions in smoking behavior between times 1 and 2.

time 1 drinking depends on time 1 smoking. The loglinear notation for the structural part of Model 2 is (S_1S_2, D_1D_2, S_1D_1).

Model 3 specifies a cross-sectional relation between smoking and drinking at times 1 and 2 (concurrent relation). In Model 3: (1) time 2 smoking depends on time 1 smoking; (2) time 2 drinking depends on time 1 drinking; (3) time 1 drinking depends on time 1 smoking; and (4) time 2 drinking depends on time 2 smoking. The loglinear notation for the structural part of Model 3 is $(S_1S_2, D_1D_2, S_1D_1, S_2D_2)$.

Model 4 specifies a baseline and lagged relation between smoking and drinking. In Model 4: (1) time 2 smoking depends on time 1 smoking; (2) time 2 drinking depends on time 1 drinking; (3) time 1 drinking depends on time 1 smoking; and (4) time 2 drinking depends on time 1 smoking. The loglinear notation for the structural part of Model 4 is $(S_1S_2, D_1D_2, S_1D_1, S_1D_2)$.

Model 5 specifies a concurrent and lagged relation between smoking and drinking. In Model 5: (1) time 2 smoking depends on time 1 smoking; (2) time 2 drinking depends on time 1 drinking; (3) time 1 drinking depends on time 1 smoking; (4) time 2 drinking depends on time 2 smoking; and (5) time 2 drinking depends on time 1 smoking. The loglinear notation for the structural part of Model 5 is $(S_1S_2, D_1D_2, S_1D_1, S_2D_2, S_1D_2)$.

Model 6 specifies that both an interaction between smoking and drinking at time 1 and a baseline relation are required to account for any observed relation between the two processes. In Model 6: (1) time 2 smoking depends on time 1 smoking; (2) time 2 drinking depends on time 1 drinking and time 1 smoking jointly; and (3) time 1 drinking depends on time 1 smoking. The loglinear notation for the structural part of Model 6 is $(S_1D_1D_2, S_1S_2)$.

Model 7 specifies both a concurrent and interactive relation between smoking and drinking. In Model 7: (1) time 2 smoking depends on time 1 smoking; (2) time 2 drinking depends on time 1 drinking and time 1 smoking jointly; (3) time 1 drinking depends on time 1 smoking; and (4) time 2 drinking depends on time 2 smoking. The loglinear notation for the structural part of Model 7 is $(S_1D_1D_2, S_1S_2, S_2D_2)$.

Model 8 specifies that an interaction between smoking and drinking at time 1, an interaction between smoking at time 1 and smoking at time 2, and a concurrent relation are all required to account for any observed relation between the two processes. That is, it is hypothesized that drinking at time 2 is explained by combinations of time 1 smoking and time 1 drinking, and transitions in smoking behavior between times 1 and 2. In Model 8, time 2 drinking depends on: (1) time 1 drinking and time 1 smoking jointly; and (2) time 1 and time 2 smoking jointly. The loglinear notation for the structural part of Model 8 is $(S_1D_1D_2, S_1S_2D_2)$.

All latent transition and multiprocess models were estimated using M*plus*. Examples of the M*plus* programming code used to fit the models discussed above may be found in Appendix C.

4.3 Results

4.3.1 Latent Transition Models

First, LTA was used to model the development of smoking and drinking individually. Latent transition models with 2 to 4 latent statuses were explored for both smoking development and drinking development. In each model, item-response probabilities were restricted to be equal across time to keep the interpretations of the latent statuses consistent but otherwise the parameters were freely estimated. Model selection was based on the G^2 fit statistic, AIC, BIC, and substantive interpretations of the models. A 3-class model was selected for smoking development, and a 3-class model was selected for drinking development. The fit statistics for the 2- to 4-class models for smoking and drinking are shown in Table 4.2.

Once the 3-class models were selected for smoking and drinking, the validity of restrictions on the measurement model was explored in preparation for fitting the multiprocess models. First, the validity of restrictions specifying measurement invariance across time in the selected 3-class smoking model and selected 3-class drinking model were tested. It was confirmed that measurement invariance across time was appropriate in both the smoking and drinking models. In the smoking model, measurement invariance was confirmed by a non-significant G^2-difference

Table 4.2. Fit Statistics for Latent Transition Models of Smoking and Drinking

Smoking	Model	LL[a]	χ^2	p^c	AIC	BIC
	2-class	-4743.761	817.491	11	9509.521	9573.855
	3-class	**-4242.674**	**79.607**	**20**	**8525.349**	**8642.320**
	4-class[b]	-4225.124	25.615	31	8512.247	8693.552
Drinking	**Model**	**LL[a]**	**χ^2**	**p^c**	**AIC**	**BIC**
	2-class	-6980.442	986.721	13	13986.884	14062.915
	3-class	**-6339.345**	**196.019**	**23**	**12724.690**	**12859.206**
	4-class[b]	-6321.127	174.501	35	12712.254	12916.953

Note: Boldface indicates the selected model.
[a] Loglikelihood value.
[b] Model may not be identified.
[c] Number of parameters estimated.

test ($0.05 < p < 0.10$). In the drinking model, the G^2-difference test was significant ($0.01 < p < 0.05$), but a comparison of the parameter estimates in the restricted and unrestricted models showed that the estimates did not change substantially and the interpretation of the model did not change when the restrictions were added. Thus, restrictions specifying measurement invariance in the drinking model were also deemed appropriate.

Second, a close examination of the measurement parameters for each smoking indicator within each latent status suggested that it was possible to constrain the item-response probabilities for the two smoking indicators to be equal within each latent status. In addition, a close examination of the measurement parameters for each drinking indicator within each latent status suggested that it was possible to constrain the item-response probabilities for two of the drinking indicators (drinking frequency and drinking quantity) to be equal within each latent status. The appropriateness of these additional restrictions was tested. Despite significant increases in the G^2 fit statistic when these parameter restrictions were imposed on each model, these restrictions were deemed appropriate in both models because they did not affect the interpretation of the latent classes in either model, and it was desirable to restrict the measurement model as much as possible before combining the individual models using MPM.

The smoking model was characterized by classes of: (1) non-smokers, (2) light smokers, and (3) heavy smokers. Similarly, the drinking model was characterized by classes of: (1) non-drinkers, (2) light drinkers, and (3) heavy drinkers. The latent status membership probabilities for the smoking and drinking models at times 1 and 2 are shown in Table 4.3. In addition, the item-response probabilities of "endorsing" an indicator given latent status membership for the smoking and drinking models are shown in Table 4.4. As discussed above, many of the item-response probabilities were restricted to be equal during estimation; these probabilities are denoted with the same letter in the table. In addition, item-response probabilities were restricted to be equal across time (measurement invariance) so there was only one set of item-response probabilities.

Table 4.3 shows that the majority of participants were non-smokers and non-drinkers at times 1 and 2, but the proportions of heavy smokers and heavy drinkers both increased between times 1 and 2; the proportions of light smokers and light

Table 4.3. Latent Status Membership Probabilities for Smoking and Drinking

Smoking	Latent Status Membership		
	Non-smokers	Light Smokers	Heavy Smokers
Time 1	0.744	0.119	0.137
Time 2	0.691	0.123	0.186

Drinking	Latent Status Membership		
	Non-drinkers	Light Drinkers	Heavy Drinkers
Time 1	0.686	0.161	0.153
Time 2	0.602	0.172	0.226

Table 4.4. Item Response Probabilities for Smoking and Drinking: Probability of Endorsing an Indicator*

Smoking		Latent Status Membership		
Indicator	Response	Non-smokers	Light Smokers	Heavy Smokers
Frequency	0 Days	0.998^a	0.070^d	0.000^g
	1-14 Days	0.002^b	0.826^e	0.060^h
	15+ Days	0.000^c	0.104^f	0.940^i
Quantity	0 Cigs	0.998^a	0.070^d	0.000^g
	1-2 Cigs	0.002^b	0.826^e	0.060^h
	3+ Cigs	0.000^c	0.104^f	0.940^i

Drinking		Latent Status Membership		
Indicator	Response	Non-drinkers	Light Drinkers	Heavy Drinkers
Frequency	0 Days	0.994^j	0.000^m	0.006^p
	1-2 Days	0.005^k	0.790^n	0.148^q
	3+ Days	0.002^l	0.210^o	0.846^r
Quantity	0 Drnks	0.994^j	0.000^m	0.006^p
	1-2 Drnks	0.005^k	0.790^n	0.148^q
	3+ Drnks	0.002^l	0.210^o	0.846^r
Binge	No	0.892	1.000	0.113
	Yes	0.108	0.000	0.887

Note: Probabilities superscripted with the same letter were restricted to be equal during estimation.
Note: Restrictions specifying measurement invariance across time were imposed in both the smoking and drinking models; item-response probabilities at time 1 were restricted to be equal to item-response probabilities at time 2.
* Probability of endorsing an indicator given latent status membership.

drinkers stayed roughly the same. The smoking and drinking latent statuses were labeled according to the indicator responses having the highest probabilities of endorsement for the status. For example, from Table 4.4, non-smokers had high probabilities of responding that they smoked 0 cigarettes on 0 days in the past month (0.998), and light drinkers had high probabilities of responding that they drank 1-2 drinks on 1-2 days in the past month (0.790).

Development in smoking and drinking individually were described with transition probabilities between times 1 and 2. The transition probabilities from time 1 to 2 for smoking and drinking are shown in Table 4.5. The individual table entries are the probabilities of time 2 smoking (or drinking) latent status membership conditional on time 1 smoking (or drinking) latent status membership. From Table 4.5, it appears that there was high stability of non-smoking, heavy smoking, non-drinking, and heavy drinking between times 1 and 2; participants who were non-smokers, heavy smokers, non-drinkers, or heavy drinkers at time 1 were likely to be members of the same latent status at time 2. Comparatively, the behavior of light smokers and light drinkers was more variable; light smokers and light drinkers at time 1 were likely to be members of any smoking and drinking status at time 2.

4.3.2 Multiprocess Models

Second, to address the primary objective of the current study, MPM was used to model smoking and drinking development simultaneously. The fit statistics for the eight models that comprise the set of relevant models for consideration are shown in Table 4.6. To aid in visual comparison of the models, Figure 4.9 plots the AIC and BIC values for each model. Model 3 has one of the lowest AIC values and the lowest BIC value, and was selected as the best at balancing fit and parsimony in its description of the relation between smoking and drinking development. The results of Model 3 were used to describe the relation between smoking and drinking.

The relation between smoking and drinking was explored using two sets of probabilities. The first set of probabilities described the relation between smoking and drinking cross-sectionally. The probabilities for the cross-sectional relation between smoking and drinking at times 1 and 2 are shown in Table 4.7. The

Table 4.5. Transition Probabilities for Smoking and Drinking

Smoking

Time 1 Latent Status Membership	Time 2 Latent Status Membership: Non-smokers	Light Smokers	Heavy Smokers
Non-smokers	0.841	0.107	0.053
Light Smokers	0.433	0.342	0.225
Heavy Smokers	0.102	0.027	0.872

Drinking

Time 1 Latent Status Membership	Time 2 Latent Status Membership: Non-drinkers	Light Drinkers	Heavy Drinkers
Non-drinkers	0.731	0.169	0.100
Light Drinkers	0.437	0.280	0.283
Heavy Drinkers	0.196	0.079	0.725

Note: Probabilities sum to 1.000 across each row.

Table 4.6. Fit Statistics for Multiprocess Models

Model	p^a	AIC	BIC
1 - (S_1S_2, D_1D_2)	31	21468.8	21650.1
2 - (S_1S_2, D_1D_2, S_1D_1)	35	20926.0	21130.7
3 - $(\mathbf{S_1S_2, D_1D_2, S_1D_1, S_2D_2})$	**39**	**20686.5**	**20914.6**
4 - $(S_1S_2, D_1D_2, S_1D_1, S_1D_2)$	39	20881.9	21110.1
5 - $(S_1S_2, D_1D_2, S_1D_1, S_2D_2, S_1D_2)$	43	20673.6	20925.1
6 - $(S_1D_1D_2, S_1S_2)$	47	20859.8	21134.7
7 - $(S_1D_1D_2, S_1S_2, S_2D_2)$	51	20662.5	20960.8
8 - $(S_1D_1D_2, S_1S_2D_2)$	59	20672.5	21017.6

Note: Boldface indicates the selected model.
^{a} Number of parameters estimated.

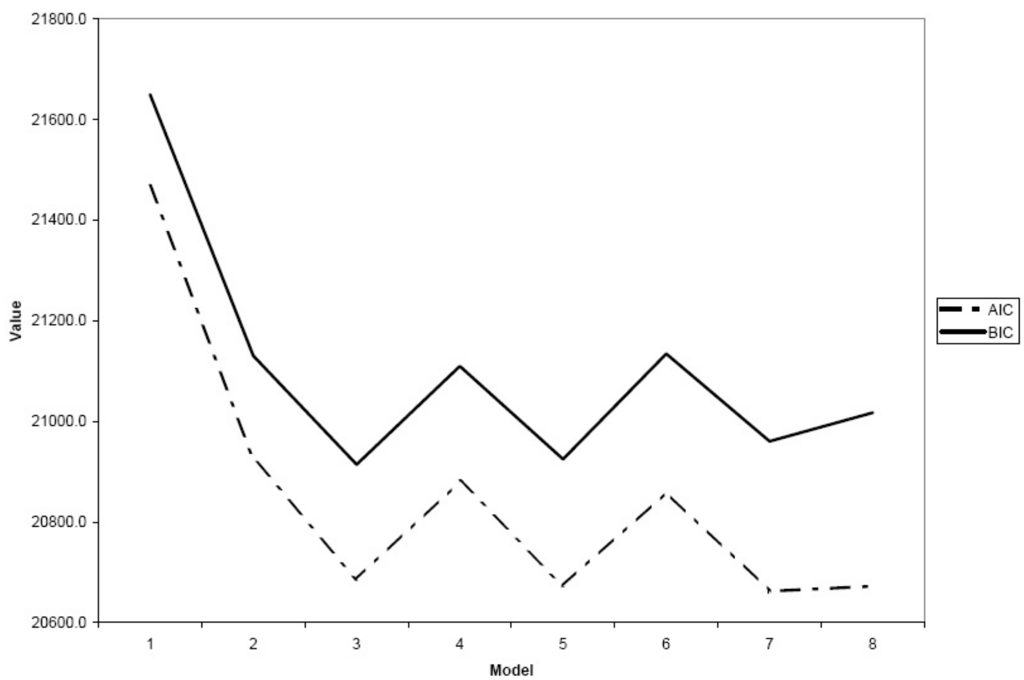


Figure 4.9. Fit statistics for the multiprocess models.

individual entries are the probabilities of time 1 drinking (or time 2 drinking) latent status membership conditional on time 1 smoking (or time 2 smoking) latent status membership.

From Table 4.7, the cross-sectional relation between smoking and drinking appears to be about the same at times 1 and 2. Non-smokers were most likely to also be non-drinkers, and heavy smokers were most likely to also be heavy drinkers. Interestingly, heavy smokers also had a fairly high probability of being non-drinkers, especially at time 1. Comparatively, the drinking behavior of light smokers appears to be quite variable. At time 1, light smokers were approximately equally likely to engage in any of the three types of drinking behaviors; at time 2 light smokers were slightly more likely to be non-drinkers, but were likely to engage in any of the three types of drinking behaviors.

The second set of probabilities described how developmental transitions in drinking varied by smoking behavior. The probabilities for developmental transitions in drinking from time 1 to 2, controlling for time 1 smoking are shown in Table 4.8. The individual entries are the probabilities of time 2 drinking latent status membership conditional on time 1 drinking and time 1 smoking latent status memberships jointly. For example, a participant who was a heavy drinker and a heavy smoker at time 1 had an 84.6% chance of being a heavy drinker at time 2.

From Table 4.8, transitioning from the non-drinking latent status to the heavy drinking latent status between times 1 and 2 was progressively more likely for more serious types of smoking: non-smokers had a low probability of making the transition (0.083), light smokers had a slightly higher probability (0.157), and heavy smokers had the highest probability of making the transition (0.282). Similar progressively increasing and decreasing patterns are seen for: (1) the probabilities of being a heavy drinker at time 2 given heavy drinking latent status membership at time 1, and (2) the probabilities transitioning from the heavy drinking latent status to the non-drinking latent status between times 1 and 2. Non-smokers were least likely to be heavy drinkers at both times (0.571), light smokers were more likely (0.696), and heavy smokers were most likely to be heavy drinkers at both times (0.846). Heavy smokers were least likely to transition from heavy drinking to non-drinking (0.099), light smokers were more lightly to make the transition (0.213), and non-smokers were most likely to make the transition (0.326).

Table 4.7. Relational Probabilities for the Cross-sectional Relation Between Smoking and Drinking: Drinking Conditional on Smoking

Time 1	Drinking Latent Status Membership		
Smoking Latent Status Membership	Non-drinkers	Light Drinkers	Heavy Drinkers
Non-smokers	0.804	0.134	0.062
Light Smokers	0.358	0.313	0.329
Heavy Smokers	0.313	0.191	0.496
Time 2	Drinking Latent Status Membership		
Smoking Latent Status Membership	Non-drinkers	Light Drinkers	Heavy Drinkers
Non-smokers	0.726	0.161	0.114
Light Smokers	0.406	0.274	0.320
Heavy Smokers	0.267	0.154	0.578

Note: Probabilities sum to 1.000 across each row.

Table 4.8. Relational Probabilities for the Developmental Relation Between Smoking and Drinking: Time 2 Drinking Conditional on Time 1 Smoking and Drinking

		T2 Drinking Latent Status Membership		
T1 Smoking Latent Status Membership	T1 Drinking Latent Status Membership	Non-drinkers	Light Drinkers	Heavy Drinkers
Non-smokers	Non-drinkers	0.754	0.164	0.083
	Light Drinkers	0.512	0.281	0.207
	Heavy Drinkers	0.326	0.104	0.571
Light Smokers	Non-drinkers	0.634	0.209	0.157
	Light Drinkers	0.377	0.299	0.324
	Heavy Drinkers	0.213	0.091	0.696
Heavy Smokers	Non-drinkers	0.512	0.206	0.282
	Light Drinkers	0.241	0.247	0.512
	Heavy Drinkers	0.099	0.055	0.846

Note: Probabilities sum to 1.000 across each row.

Finally, the developmental transitions in drinking of those participants who both smoked lightly and drank lightly at time 1 were variable, but there were some interesting patterns. Light drinkers at time 1 were approximately equally likely to be light drinkers at time 2, regardless of their smoking behavior (0.281, 0.299, 0.247). In addition, non-smokers were most likely to transition from light drinking to non-drinking (0.512) and heavy smokers were most likely to transition from light drinking to heavy drinking (0.512); light smokers were approximately equally likely to transition to non-drinking and heavy drinking.

4.4 Discussion

The current study used LTA to identify types of adolescent smoking and drinking behaviors, and to describe change over time in these behaviors using data from the NLSY97. To address the primary objective, the current study modeled the relation between smoking and drinking development using MPM. The goal of the current study was to examine how drinking development relates to smoking behavior.

4.4.1 Smoking Development and Drinking Development

Using LTA, three types of smokers and three types of drinkers were identified: non-smokers, light smokers, heavy smokers, non-drinkers, light drinkers, and heavy drinkers. In addition, change over time in latent status membership was described for smoking and drinking separately. The majority of participants were non-smokers and non-drinkers at both times, but both heavy smoking and heavy drinking increased over time. Light smokers appeared to have the most variable behavior over time whereas non-smokers, non-drinkers, heavy smokers, and heavy drinkers appeared to have relatively stable behavior over time. Light smokers and light drinkers appeared to be experimenting with cigarettes and alcohol — light smokers and light drinkers at time 1 were likely to transition to any smoking and drinking latent statuses at time 2.

4.4.2 Relation Between Smoking and Drinking

The results discussed above that describe the cross-sectional relation between smoking and drinking suggest that knowing a participant smokes at any particular time provides limited information about his or her drinking behavior. Knowing a participant smokes lightly provides almost no clue as to what his or her drinking behavior is; knowing a participant smokes heavily provides only limited information in that the participant is likely to also be a heavy drinker, but he or she also has a substantial chance of having a different type of drinking behavior.

The results discussed above that describe the relation between developmental transitions in drinking and type of smoking behavior are particularly interesting. Transitioning from the non-drinking latent status to one of the drinking latent statuses between times 1 and 2 is progressively more likely for more serious types of smoking. Non-smokers at time 1 are least likely to transition from the non-drinking to the light or heavy drinking latent statuses between times 1 and 2, light smokers are more likely to make a transition, and heavy smokers are most likely to make a transition to light or heavy drinking. In addition, increasingly serious smoking behavior is related to progressively higher probabilities of transitioning from the non-drinking to the heavy drinking latent status between times 1 and 2: non-smokers have a low probability of making the transition from non-drinking to heavy drinking, light smokers have a slightly higher probability of making the transition, and heavy smokers have the highest probability of making the transition.

4.4.3 Implications for Prevention

Examining the way in which developmental transitions in drinking vary by smoking behavior provides interesting implications for prevention and treatment researchers. The current study provides additional evidence that it is particularly important for prevention programs to be implemented in late childhood or early adolescence in order to delay smoking and drinking onset, and to prevent heavy smoking and drinking for as long as possible: individuals who smoke and/or drink heavily are not likely to change their behaviors. Individuals who are light smokers and light drinkers, however, are a promising group of adolescents who may be highly likely to respond favorably to prevention. These adolescents appear to have variable behavior and this may be a good time at which to intervene. The current study reinforces the need for prevention researchers to continue implementing programs during this developmental period during which individuals may be at risk for initiating or increasing their smoking and/or drinking behaviors.

4.4.4 Exchanging Smoking and Drinking in the Multiprocess Models

The current study described how developmental transitions in drinking varied by type of smoking behavior. This is because it was hypothesized that knowing type of smoking provided comparatively more information about type of drinking than what knowing type of drinking provided about type of smoking (because more adolescents drink than smoke). However, this was a somewhat arbitrary decision. It would have been possible to exchange the direction of prediction and to describe how developmental transitions in smoking varied by type of drinking. It would also have been possible to look at the relation between the processes both ways.

When the smoking and drinking processes are exchanged, so that drinking predicts smoking, some of the results discussed here remain the same and some of them change. Consider the concurrent relation model (Model 3 - $S_1S_2, D_1D_2, S_1D_1, S_2D_2$). If the smoking and drinking processes were exchanged, the effects included in the loglinear model would remain the same because no new conditional dependencies are included in the model. This is most easily seen by examining the loglinear notation for the structural part of this new model: $(D_1D_2, S_1S_2, D_1S_1, D_2S_2)$. In this case, the fit statistics for the new model would be the same as those for Model 3.

Now consider the concurrent and lagged relation model (Model 5 - $S_1S_2, D_1D_2, S_1D_1, S_2D_2, S_1D_2$). If the smoking and drinking processes were exchanged in this model, the effects in the loglinear model are different because a new conditional dependency is included and an old conditional dependency is excluded. Again, this is most easily seen by examining the loglinear notation for the structural part of this new model: $(D_1D_2, S_1S_2, D_1S_1, D_2S_2, D_1S_2)$. In this new model D_1 and S_2 are conditionally dependent whereas S_1 and D_2 were conditionally dependent in Model 5. In this case, the fit statistics for the new model would be different from those for Model 5.

Finally, when the smoking and drinking processes are exchanged it is likely that the researcher is interested in addressing a slightly different research question. It is also likely, then, that the matrices selected and calculated to describe the nature of the relation between smoking and drinking are different. For example, the probability matrix describing how developmental transitions in smoking vary by type of drinking may be calculated instead of the matrix discussed above describing how developmental transitions in drinking vary by type of smoking.

4.4.5 A Note on the Practical Application of Multiprocess Models

The current study used a two-step process to address the research question. In the first step, LTA was used to explore the structure, measurement, and change over time for smoking and drinking individually. By so doing, an understanding of each developmental process was gained prior to trying to understand a more complex set of relations.

In general, the first step provides an opportunity to explore the individual processes in depth, and to explore the validity of restricting parts of the models for the individual processes. For example, measurement invariance over time, similar item-response probabilities across indicators or statuses, or stationarity (when three or more times are modeled) may be explored. When MPM is the ultimate goal, it may be desirable to have heavily restricted individual models, as long as the restrictions are appropriate. This is because multiprocess models may often be difficult to identify — the size of the contingency table for one of these models tends to be inordinately large. Restrictions may help in obtaining an identified solution.

In the second step, the two individual processes were modeled simultaneously using MPM. A set of models to be compared was selected based on hypothesized relations between smoking and drinking. The models were then fit and compared to determine the model best at balancing fit and parsimony in its description of the observed relation between smoking and drinking. The selected multiprocess model was then used to make substantive conclusions. In general, the second step models multiple developmental processes simultaneously using MPM.

In practice, employing the two-step process used here may be helpful when using MPM to address research questions. The two-step process builds multiprocess models incrementally, providing multiple opportunities to explore and assess the validity of models individually and jointly.

4.4.5.1 Substantively Interpreting Multiprocess Models

There are two main ways to make substantive conclusions using a multiprocess model. The first is to take the approach discussed here and calculate matrices describing the relation between the two processes in ways that directly address the research question. This may be done directly by using the vector of proportions for the latent class patterns based on the estimated model. In the current study the vector was used to calculate the entries in Table 4.7, probabilities that describe the cross-sectional relation between smoking and drinking at times 1 and 2, and Table 4.8, probabilities that describe the relation between developmental transitions in drinking and type of smoking. Depending on the research question, however, different matrices may be of interest and may be calculated. For example, it may be of interest to examine how developmental transitions in drinking vary by developmental transitions in smoking. The vector would then be used to calculate a matrix with entries corresponding to the probabilities of time 2 drinking latent status membership conditional on time 1 smoking, time 1 drinking, and time 2 smoking latent memberships jointly.

The second is to use odds ratios. This may be done by selecting and calculating odds ratios that describe some particular relations of interest. For example, in the current study, the odds of time 2 heavy drinking relative to time 2 non-drinking, for time 1 heavy drinking relative to time 1 non-drinking could have been calculated. In multiprocess models that do not include all multi-way interaction effects between the latent variables, restrictions are imposed on the odds ratios; which restrictions depend on which effects are left out of the model. The odds ratio approach to making substantive conclusions has been avoided in the current study because it seems less intuitive than the probability matrix approach.

4.4.6 Limitations

There are two main limitations to the current study. The first is that only participants aged 15 to 16 were included in the sample. Participants were selected primarily to avoid potential age differences in smoking and drinking behavior, and in the relation between smoking and drinking. It was desirable to keep the ages of the participants as homogeneous as possible. Fortunately, potential sample size was not a concern with the NLSY97 data so it was possible to select a small age range for the current study. However, this does limit the generalizability of the results presented here to a subset of adolescents.

The second limitation is that the indicators described smoking and drinking behavior only in the past 30 days, at two times tbat were approximately a year apart. Only to the extent that these "snapshots" provide a picture of typical adolescent smoking and drinking behavior is it possible to make conclusions about how smoking and drinking are related over longer periods of time.

4.4.7 Conclusions

The current study illustrates the application and importance of multiprocess models to developmental researchers. Multiprocess models provide a way to explore how developmental transitions in one behavior vary across levels of another behavior. To extend this to exploring how developmental transitions in one behavior vary across developmental transitions in another behavior, matrices of other relational probabilities may be calculated and interpreted. The results of the current study suggest that it may be particularly important to target prevention programs to adolescents who are currently engaging in low levels of smoking and drinking. These adolescents appear to have variable behavior and it may be a particularly good time to intervene in order to sway their behavior toward not smoking and/or not drinking. In addition, the results of the current study also suggest that adolescents who are already heavy smokers or heavy drinkers by age 15 are likely to continue smoking and drinking heavily, so it is important to intervene earlier in order to prevent heavy smoking and heavy drinking to the largest possible extent.

CHAPTER 5

Discussion

5.1 Three Empirical Studies

The three empirical studies were selected in order to: (1) address developmental questions in gambling and substance use research that are of particular importance to prevention and treatment scientists; and (2) illustrate the use of sophisticated latent class modeling methodologies. The three studies were presented in order of increasing model complexity. The first study illustrated the use of latent class and conditional latent class models for identifying and predicting groups of individuals within a population. The second illustrated the use of multivariable models to describe the nature of the relation between two developmental variables. The third illustrated the use of multiprocess models to parsimoniously describe the nature of the relation between two developmental processes and describe how developmental transitions in one process may vary across types of behavior in the other.

5.1.1 Increased Knowledge About Gambling and Substance Use

The approaches taken in the first and second empirical studies provided opportunities to examine the nature of the relation between demographic characteristics, gambling, and drinking in ways that have not been used previously. The person-centered approaches and the ways in which the applied methodologies allowed specific patterns of behavior to be linked resulted in increased knowledge about gambling.

First, support was provided for the idea that it may be possible to identify types of gamblers based on similarities in behavioral characteristics that move beyond diagnostic criteria endorsement. In identifying latent classes of gambling behavior, specific gender differences in the distribution of men and women among gambling types, and specific gender differences in the probabilities of playing many different gambling activities were uncovered. Although there was already a great deal of evidence that gender differences exist in gambling, the current project was able to show the similarities and differences between men and women among individuals whose behavior was defined by the same primary gambling activities.

Second, a unique look at how gambling on the internet fits in with more traditional gambling activities like cards and the lottery was provided. Currently, policy makers and parents alike are becoming increasingly concerned about the availability of gambling on the internet, and it has been suggested that internet gambling is particularly prevalent among college students. Results of the first empirical study suggested that internet gambling is not particularly common among participants playing only one or two types of gambling activities, men and women who gambled on the internet appeared to also be engaging in a variety of other gambling activities. It may be that internet gambling is just part of a larger pattern of potentially problematic gambling behavior.

Third, a more nuanced look at the relation between Greek membership and gambling suggested that there is a more complex relation between Greek membership and gambling than has been shown by previous correlational research. For example, the effects of Greek membership were in opposite directions for men and women. Greek membership appeared to be a risk factor for certain types of gambling for men but a protective factor for women.

Fourth, more information about the relation between gambling and drinking was uncovered. While it is known that drinking tends to co-occur with gambling, the results of the first and second empirical studies showed that past-year drinking and longitudinal patterns of drinking over time are important risk factors for multi-game gambling but that they are not strongly related to lighter types of gambling. That is, it appears that infrequent or lighter drinking behaviors are not strongly associated with less varied and less potentially problematic gambling activities, perhaps because many individuals engage in both behaviors casually, whereas individuals who drink frequently or heavily are at increased risk for engaging in patterns of potentially problematic gambling behavior. The first empirical study also showed that the relation between gambling and drinking is much stronger for women than for men.

Overall, the current project places gambling in with other problem behaviors like drinking and unprotected sex. The findings discussed in the current project provide empirical evidence that gambling belongs to the same domain as other problem behaviors, an idea that has been suggested and discussed by Winters et al. (2002), Winters et al. (1993), Vitaro et al. (2001), and Petry and Tawfik (2001). To see how well gambling fits in the constellation of problem behaviors it would be interesting in future research to look closely at the relations among gambling and other less studied behaviors like unprotected sex.

Finally, the approach taken in the third empirical study provided opportunities to examine the nature of the relation between smoking and drinking in ways that have not been used previously. The person-centered approaches and the ways in which the applied methodologies allowed specific patterns of behavior to be linked resulted in increased knowledge about how developmental transitions in drinking vary by types of smoking behavior. The results showed that it is not the case that knowing type of smoking provides information about type of drinking. In addition, empirical evidence was provided to show that transitioning from non-drinking to drinking is progressively more likely for more serious types of smoking, and that more serious types of smoking are related to progressively higher probabilities of transitioning from non-drinking to heavy drinking. Empirical evidence also showed that non-smokers are most likely to transition from light to non-drinking while heavy smokers are most likely to transition from light to heavy drinking.

5.1.2 New Ways of Studying Gambling and Substance Use

The methodological approaches used in the current project are new to the study of gambling, and are somewhat new to the study of substance use. They provided new ways of looking at the data that resulted in some interesting knowledge being gained about gambling and substance use. Several types of sophisticated latent class modeling techniques were used in the current project, including: CLCA, RMLCA, MVLCM, LTA, and MPM. These methods provided a person-centered approach to examining the multidimensional nature of gambling and substance use. They also provided a way to model prospective longitudinal data for multiple processes simultaneously.

LCA and CLCA with a grouping variable provided ways to consider the multidimensional nature of gambling behavior in the identification of types of gamblers. These methods identified multiple types of gambling behavior in adolescents and young adults based on similar patterns of gambling. This may allow prevention and treatment researchers to: (1) move beyond the use of diagnostic criteria to identify individuals for prevention and treatment programs, and (2) identify individuals at increased risk for developing problems with gambling before their gambling behavior escalates. In addition, these methods provided a way to examine gender differences in the structure of gambling and in the way potential risk factors for gambling operate. For example, they provided unique evidence that gambling on the internet tends to occur with a variety of other types of gambling activities rather than alone. Additionally, they provided unique evidence that for college students Greek membership has a complex relation with gambling behavior that differs by gender, and that alcohol use may be a much stronger risk factor for women than for men.

MVLCM provided a way to link specific gambling behaviors to specific longitudinal patterns of drinking, which has not been done in previous research. Two particularly interesting relations between gambling and patterns of drinking were suggested by the results of this modeling approach. The first is that drinking frequency may be more predictive of heavier types of gambling than drinking quantity.

The second is that consistent drinking over time, whether it is frequent, heavy, or both, is more predictive of gambling in multiple types of activities than of gambling in only one or two common types of behaviors. These findings may be particularly important for prevention and treatment researchers because they suggest that any type of regular drinking, regardless of amount, may be a risk factor for problems with gambling.

MPM provided a way to model multiple developmental processes simultaneously over time. It also provided a way to describe how developmental transitions in one process are linked to behaviors in another process. This method was used to describe how developmental transitions in drinking were related to different types of smoking behavior. The results suggested that there is indeed an interesting and progressive relation between smoking and drinking. More severe types of smoking were related to transitions from less severe to more severe types of drinking; less severe types of smoking were related to transitions from more severe to less severe types of drinking. The results suggest that prevention and treatment researchers need to continue implementing programs for children and young adolescents to prevent or delay the onset of heavy smoking and drinking, as heavy smoking and heavy drinking appeared to be relatively stable across time. In addition, light smokers and light drinkers may be particularly good groups of individuals to target as their behavior may be more amenable to influence.

5.1.3 Connecting the Three Empirical Studies Methodologically

To understand how the three empirical studies are interconnected methodologically, it is helpful to consider the number of discrete latent variables in each model and how that relates to the discussion of loglinear modeling with latent variables presented in Chapter 1. The first empirical study had one discrete latent variable: gambling. In this study, loglinear modeling in the form of LCA with a grouping variable was used to understand the ways in which the manifest indicators were probabilistically related to the latent variable. This allowed the description of the latent classes comprising the discrete latent variable and how participants are distributed among the latent classes. In addition, CLCA with a grouping variable was used to predict latent class membership. In CLCA, the predictors are manifest covariates in a multinomial logistic regression where the outcome is latent class membership.

The second empirical study was more complex in the sense that it had two discrete latent variables: gambling (G) and drinking (D). In this study, loglinear modeling with latent variables was used to understand: (1) the ways in which manifest indicators of the latent variables are probabilistically related to their respective latent variables, and (2) the ways in which the latent variables are related. There are two possible loglinear models that may be used to model the relation between gambling and drinking. The first is the independence model (D, G), which specifies that there is no relation between gambling and drinking. The second is the dependence model (DG), which specifies that there is a relation between gambling and drinking. With two latent variables, the set of relevant models for consideration is small and manageable. The independence model was primarily used as a comparison model for the dependence model. Fitting the independence model in addition to the dependence model provided a way to explore the extent to which the dependence model fit better than the independence model. In addition, the dependence model provided a way to describe how gambling latent class membership is related to patterns of drinking behavior over time.

It would have been possible to include an empirical study that had three discrete latent variables. There are nine possible loglinear models that may be used to model the structural relations among three variables. These models include the complete independence model, one-factor independence models, conditional independence models, homogeneous association model, and the saturated model. With three latent variables, the set of relevant models for consideration is manageable even if all nine are of substantive interest.

The third empirical study was even more complex in the sense that it has four discrete latent variables: drinking at times 1 and 2 (D_1, D_2) and smoking at times 1 and 2 (S_1, S_2). Similar to the second study, in this study loglinear modeling with latent variables was used to understand: (1) the ways in which manifest indicators of the latent variables are probabilistically related to their respective latent variables, and (2) the ways in which the latent variables are related and interact. The number of possible loglinear models that may be used to model the structural relations among four variables is exponentially larger than that of two- and three-variable models. The complete set of possible models range in complexity from models including only the main effects to models that include two- and three-way interaction effects to the saturated model that includes the four-way interaction effect. With four latent variables, the set of relevant models for consideration must be carefully chosen to correspond to theoretically logical, hypothesized relations among the variables or processes. In this study, eight sets of relations were hypothesized, which corresponded to eight different multiprocess models. The selected model provided a way to describe how transitions in drinking were related to specific smoking behaviors.

5.1.4 Missing Data

Analyses for all three empirical studies were conducted with PROC LCA and M*plus*, as discussed in their respective chapters. In general, missing data was minimal in all three studies due to the way in which participants were selected in each study. As discussed in each empirical study, participants providing responses to any of the questions used as indicators for the latent class, latent class for repeated measures, and latent transition models were included in the analyses. PROC LCA (Lanza et al., 2007) and M*plus* (Muthén & Muthén, 1998-2006) both use full-information maximum likelihood techniques for model estimation, which avoids eliminating participants due to missing data. This technique, however, does require the assumption that data are missing at random (MAR).

The only models in which missing data may potentially be a concern are the conditional latent class models in the first empirical study where demographic characteristics and substance use behaviors are used as predictors. The full-information maximum likelihood technique used by PROC LCA does not incorporate missing data on the independent variable side of the multinomial logistic regression model. Participants with missing values on the predictors are list-wise deleted from analyses. Fortunately, the study had very low rates of missing data in general, and the list-wise deletion of participants from the conditional latent class models did not have any noticeable effect on the structure of the latent class model. The largest number of participants list-wise deleted in the conditional latent class models was in the model that included binge drinking as a predictor of latent class membership; 20 participants were deleted due to missing data.

5.1.5 Methodology Selection and Other Important Decisions

Selecting among the many different latent class modeling approaches discussed in the current project can be challenging when faced with complex research questions about development. The best piece of advice that helps make this task easier is to start with a well-formulated research question. Often, a well-formulated research question will point to the best methodology to use.

The overall goal of the third empirical study was to examine the nature of the relation between smoking development and drinking development. An alternative method to the one discussed was to: (1) model patterns of smoking over time using RMLCA; (2) model patterns of drinking over time using RMLCA; and (3) link smoking and drinking using MVLCM. This method would have linked longitudinal patterns of smoking behavior to longitudinal patterns of drinking behavior. This approach would have been similar to the growth mixture modeling approach discussed in the introduction to the third empirical study and to dual trajectory modeling (Nagin, 2005). This approach, however, would not have uncovered how developmental transitions in drinking between times 1 and 2 were linked to type of smoking. The research question directed which methodology to use.

When posing research questions that include prediction, it is important to think carefully about the direction of prediction. Decisions made about the direction of prediction help determine which latent variables are conditional on which other latent variables. Often, what you condition on is arbitrary. In these cases it is important to make a cohesive argument about which direction is chosen, or to examine prediction in both directions and note the differences between the models.

In the three empirical studies presented in the current project, decisions about the direction of prediction were not so arbitrary. In the first empirical example it was of interest to see how different demographic characteristics and substance use behaviors were related to gambling latent class membership; the easiest way to do this is to include the covariates into the model as predictors. In the second empirical example, gambling behavior was measured only at the third wave whereas drinking behavior was measured over time; it made the most logical sense to examine how longitudinal patterns of drinking predicted gambling. In the third empirical example it was assumed that because more adolescents drink than smoke, predicting drinking from smoking was relatively more informative, as discussed earlier.

When the direction of prediction is exchanged, some models will have the same fit and others will not. This was discussed in detail in Chapter 4. When the direction of prediction is exchanged, different parameters may be of interest. In addition, different parameters can be significant, even when the measurement model is the same.

5.2 General Use of Loglinear Models with Latent Variables

5.2.1 Advantages

There is a multitude of ways that loglinear models with latent variables may be applied to research questions about developmental variables. Models with two latent variables may be used with cross-sectional data to model the relation between two variables, or they may be used with longitudinal data to model development in a single construct across two times (LTA). Models with three latent variables may be used with cross-sectional data to model the relations among three variables, or they may be used with longitudinal data to model development in a single construct across three times, or to model the relation between a single developmental process across two times and a third variable. Similarly, models with four latent variables may be used with cross-sectional data to model the relation among four variables, or they may be used with longitudinal data to model the relation between two developmental processes across two times, and so on. The sky appears to be the theoretical limit, so to speak, with models including more than four latent variables. In the end, it comes down to the interpretation of the models selected for the set of relevant models for consideration.

The primary advantage of loglinear models with latent variables related to the current project is the ability to model the relations among multiple developmental processes over time. Other methods, like SEM, are well-developed and widely used, but these methods are primarily for modeling relations among continuous latent variables. In addition, loglinear models with latent variables provide a person-centered methodological approach, unlike methods like SEM. Without the use of loglinear models it is difficult to address research questions about multiple developmental variables conceptualized as discrete latent variables. In addition, loglinear models may provide a way to investigate direct and indirect effects, and mediation (Hagenaars, 1993).

5.2.2 Limitations

5.2.2.1 Contingency Table Sparseness

It is easy for multivariable and multiprocess models to become very large (defined as the size of the contingency table created by the manifest indicators) very quickly. These models, by definition, include many discrete latent variables and estimate many parameters. This presents two challenges of particular importance to the practical application of these models. The first is contingency table sparseness; the second is model estimation, identification, and interpretation.

Contingency table sparseness is a concern for all discrete latent variable models. It is often difficult to have sample sizes large enough to ensure that relatively few cells of the contingency table created by all the possible patterns of responses to the indicators will be empty. This is a concern particularly for models that contain many discrete latent variables, many indicators for any one variable, or indicators with many response options. When a contingency table is sparse, it is difficult to ascertain the absolute model fit because the G^2 fit statistic is no longer distributed as χ^2 (Lanza et al., 2003). In addition to issues with absolute model fit, it is not clear how well the AIC and BIC perform for model selection when they are being used to select among models of the size likely in MPM, regardless of contingency table sparseness.

The estimation of a large number of parameters, especially in the presence of a sparse contingency table, is also often a concern in discrete latent variable models. Typically, the estimation of discrete latent variable models uses a maximum likelihood approach. The likelihood function for models that estimate large numbers of parameters may contain many local maxima, making it difficult for the estimation procedure to identify the maximum likelihood solution. Model identification may be checked by using many different random starting values, as the solution should not vary based on the starting values used if the model is identified. When model identification is challenging it may be necessary to use parameter restrictions or fixing in order to identify a solution and stabilize parameter estimation.

Finally, models with many discrete latent variables provide many different ways in which to interpret the results. Often, there are too many parameters to reasonably interpret and there are too many possible relations to reasonably explore. In general, researchers should be judicious about what will be interpreted and relations of interest should be determined before attempting to interpret selected models.

5.2.2.2 Computation Time for Model Estimation

The biggest limitation to the practical application of these models to development may be computational resources. In order to fit multivariable and multiprocess models, specialized software is required. In addition, many potentially interesting models require long periods of time for model estimation. The computational resources and time required to fit multiprocess models at three or more times (or with more than two processes) may be prohibitive for some researchers. Some software packages may take less time than others, but these software packages likely run one set of starting values at a time. And, as discussed above, it is essential to check the identification of these models with multiple sets of starting values. Other software packages may run multiple sets of starting values in a single run of the program, but this requires the ability to let the computer run without interruption for long periods of time.

5.3 Alternatives to Loglinear Models with Latent Variables

Developmental processes are not always conceptualized as discrete latent variables. When it is not possible or desirable to conceptualize development as discrete, loglinear modeling with latent variables is not an appropriate methodological approach. Models of development for continuous latent variables include random effects models (Snijders & Bosker, 1999), latent growth curve models (Duncan et al., 1999), group-based trajectory models (Nagin, 1999, 2005), and growth mixture models (Muthén & Sheddon, 1999). In general, these techniques model growth in a developmental process over time where the outcome is continuous. Group-based trajectory models and growth mixture models also allow for the identification of groups or classes of individuals based on similar patterns of growth.

Methods for linking two developmental processes over time when development is conceptualized as a continuous latent variable include joint trajectory modeling with PROC TRAJ (Nagin, 2005), dual trajectory modeling with SEM (Willett & Sayer, 1996), and using a variable to predict an intercept and slope in latent growth curve modeling (Willett & Sayer, 1994). In general, these techniques predict growth in a developmental process over time from another variable.

5.4 Future Directions

The current project lays the groundwork for future research in several areas. The first is the dissemination of these methods via additional developmental applications. The methods discussed here are applicable to many areas of developmental research. It is planned that these methods will be used to continue to address research questions about gambling, substance use, and sexual behavior. The second is to reach a deeper understanding of the role of multi-way interactions in loglinear models with latent variables, specifically in developmental applications. It is not yet clear what types of situations would require the use of the saturated model in MPM, nor is it clear how to program such a model in commonly used software packages. A third area of future research concerns the issues of model identification, fit, selection, and estimation. All of these need to be more thoroughly explored in order to make practical recommendations to substantive researchers interested in these methods. A closely related topic also of interest is the issue of power in latent class models in general. A fourth area of future research is the application of multiprocess models to research questions investigating the idea of reciprocal relations between developmental processes, and to research questions addressing the gateway hypothesis.

5.5 Conclusions

The current project used loglinear models with latent variables to address a variety of research questions about development. It provides an example to substantive researchers of how sophisticated discrete latent variable methodologies may be particularly useful in the study of development, specifically in the study of problem behaviors like gambling and substance use. The current study provided support for the idea that loglinear models with latent variables are a particularly good way to model relations among developmental variables and processes. And, as a bonus, it laid the groundwork for a variety of substantive and methodological work in the future.

APPENDIX A

Programming Code for Empirical Study #1

All latent class and conditional latent class models with a grouping variable examined in the first empirical study were estimated using PROC LCA, as discussed in Chapter 2. The models examined in the first empirical study required the use of restrictions and starting values. The programs below provide examples of SAS programming code for: (1) a restrictions file, (2) a starting values file, (3) a latent class model with a grouping variable that uses restrictions and starting values, and (4) a conditional latent class model with a grouping variable that uses restrictions and starting values.

A.1 Creating a Restrictions File

```
DATA RESTRICT_4cmf;
INPUT param $ group variable $ respcat estlc1 estlc2 estlc3
      estlc4;
```

```
DATALINES;
GAMMA 1 .       .     1   1   1   1
GAMMA 2 .       .     1   1   1   1
BETA  1 .       .     1   1   1   1
BETA  2 .       .     1   1   1   1
RHO   1 cards   1     2   7   5   4
RHO   1 casino  1     3   1   1   4
RHO   1 skill   1     3   7   1   4
RHO   1 stock   1     3   1   6   1
RHO   1 bingo   1     3   1   6   1
RHO   1 dice    1     3   1   6   1
RHO   1 inter   1     3   1   6   1
RHO   1 sports  1     3   1   6   1
RHO   1 horses  1     3   1   6   1
RHO   1 bookie  1     3   1   6   1
RHO   1 lotto   1     2   1   5   4
RHO   1 slots   1     3   1   1   4
RHO   1 other   1     3   1   1   4
RHO   1 cards   1    12  17  15  14
RHO   1 casino  2    13   1   1  14
RHO   1 skill   2    13  17   1  14
RHO   1 stock   2    13   1  16   1
RHO   1 bingo   2    13   1  16   1
RHO   1 dice    2    13   1  16   1
RHO   1 inter   2    13   1  16   1
RHO   1 sports  2    13   1  16   1
RHO   1 horses  2    13   1  16   1
RHO   1 bookie  2    13   1  16   1
RHO   1 lotto   2    12   1  15  14
RHO   1 slots   2    13   1   1  14
RHO   1 other   2    13   1   1  14
RHO   2 cards   1     2   7   5   4
RHO   2 casino  1     3   1   1   4
```

```
RHO   2 skill  1     3   7   1   4
RHO   2 stock  1     3   1   6   1
RHO   2 bingo  1     3   1   6   1
RHO   2 dice   1     3   1   6   1
RHO   2 inter  1     3   1   6   1
RHO   2 sports 1     3   1   6   1
RHO   2 horses 1     3   1   6   1
RHO   2 bookie 1     3   1   6   1
RHO   2 lotto  1     2   1   5   4
RHO   2 slots  1     3   1   1   4
RHO   2 other  1     3   1   1   4
RHO   2 cards  1    12  17  15  14
RHO   2 casino 2    13   1   1  14
RHO   2 skill  2    13  17   1  14
RHO   2 stock  2    13   1  16   1
RHO   2 bingo  2    13   1  16   1
RHO   2 dice   2    13   1  16   1
RHO   2 inter  2    13   1  16   1
RHO   2 sports 2    13   1  16   1
RHO   2 horses 2    13   1  16   1
RHO   2 bookie 2    13   1  16   1
RHO   2 lotto  2    12   1  15  14
RHO   2 slots  2    13   1   1  14
RHO   2 other  2    13   1   1  14
;
RUN;
```

A.2 Creating a Starting Values File

```
DATA START_4CMF;
INPUT param $ group variable $ respcat estlc1 estlc2 estlc3
      estlc4;
DATALINES;
```

```
GAMMA 1 .      .    .30  .30  .30  .10
GAMMA 2 .      .    .30  .30  .30  .10
BETA  1 .      .      1    1    1    1
BETA  2 .      .      1    1    1    1
RHO   1 cards  1    .90  .10  .30  .20
RHO   1 casino 1    .95  .70  .60  .20
RHO   1 skill  1    .95  .10  .40  .20
RHO   1 stock  1    .95  .70  .90  .70
RHO   1 bingo  1    .95  .70  .90  .70
RHO   1 dice   1    .95  .70  .90  .20
RHO   1 inter  1    .95  .70  .90  .20
RHO   1 sports 1    .95  .30  .90  .20
RHO   1 horses 1    .95  .70  .90  .70
RHO   1 bookie 1    .95  .70  .90  .70
RHO   1 lotto  1    .90  .70  .30  .20
RHO   1 slots  1    .95  .70  .60  .20
RHO   1 other  1    .95  .70  .60  .20
RHO   1 cards  2    .10  .90  .70  .80
RHO   1 casino 2    .05  .30  .40  .80
RHO   1 skill  2    .05  .90  .60  .80
RHO   1 stock  2    .05  .30  .10  .30
RHO   1 bingo  2    .05  .30  .10  .30
RHO   1 dice   2    .05  .30  .10  .80
RHO   1 inter  2    .05  .30  .10  .80
RHO   1 sports 2    .05  .70  .10  .80
RHO   1 horses 2    .05  .30  .10  .30
RHO   1 bookie 2    .05  .30  .10  .30
RHO   1 lotto  2    .10  .30  .70  .80
RHO   1 slots  2    .05  .30  .40  .80
RHO   1 other  2    .05  .30  .40  .80
RHO   2 cards  1    .90  .10  .30  .20
RHO   2 casino 1    .95  .60  .70  .20
RHO   2 skill  1    .95  .10  .40  .20
```

```
RHO   2 stock  1    .95  .70  .90  .70
RHO   2 bingo  1    .95  .30  .90  .70
RHO   2 dice   1    .95  .60  .90  .30
RHO   2 inter  1    .95  .70  .90  .60
RHO   2 sports 1    .95  .30  .90  .60
RHO   2 horses 1    .95  .60  .90  .70
RHO   2 bookie 1    .95  .70  .90  .70
RHO   2 lotto  1    .90  .20  .30  .20
RHO   2 slots  1    .95  .30  .70  .20
RHO   2 other  1    .95  .30  .60  .20
RHO   2 cards  2    .10  .90  .70  .80
RHO   2 casino 2    .05  .40  .30  .80
RHO   2 skill  2    .05  .90  .60  .80
RHO   2 stock  2    .05  .30  .10  .30
RHO   2 bingo  2    .05  .70  .10  .30
RHO   2 dice   2    .05  .40  .10  .70
RHO   2 inter  2    .05  .30  .10  .40
RHO   2 sports 2    .05  .70  .10  .40
RHO   2 horses 2    .05  .40  .10  .30
RHO   2 bookie 2    .05  .30  .10  .30
RHO   2 lotto  2    .10  .80  .70  .80
RHO   2 slots  2    .05  .70  .30  .80
RHO   2 other  2    .05  .70  .40  .80
;
RUN;
```

A.3 Fitting a Latent Class Model with a Grouping Variable

This model fits the final model chosen to describe the gambling behavior of male and female college students at a large northeastern university, as discussed in Chapter 2. This model includes thirteen indicators of gambling and one grouping variable, and has four classes:

```
PROC LCA DATA = rutgers.mf_final RESTRICT = restrict_4cmf
         START = start_4cmf;
TITLE1 ‘Gambling - 4 Classes, Males and Females’;
TITLE2 ‘Partial Invariance Across Groups’;
NCLASS 4;
ITEMS cards casino skill stock bingo dice inter sports
      horses bookie lotto slots other;
CATEGORIES 2 2 2 2 2 2 2 2 2 2 2 2 2;
GROUPS sex;
GROUPNAMES male female;
RUN;
```

A.4 Fitting a Conditional Latent Class Model with a Grouping Variable

This model fits the final model and includes the covariate “school year” as a predictor of latent class membership, as discussed in Chapter 2:

```
PROC LCA DATA = rutgers.mf_final RESTRICT = restrict_4cmfcov
         START = start_4cmfcov;
TITLE1 ‘Gambling - 4 Classes, Males and Females’;
TITLE2 ‘Partial Invariance Across Groups & Covariate’;
NCLASS 4;
ITEMS cards casino skill stock bingo dice inter sports
      horses bookie lotto slots other;
CATEGORIES 2 2 2 2 2 2 2 2 2 2 2 2 2;
GROUPS sex;
GROUPNAMES male female;
COVARIATES scyrc;
REFERENCE 1;
RUN;
```

APPENDIX B

Programming Code for Empirical Study #2

All latent class models and latent class models for repeated measures examined in the second empirical study were estimated using PROC LCA; the multivariable latent class models that predicted gambling latent class membership from drinking latent trajectory class membership were estimated using M*plus*, as discussed in Chapter 3. The models examined in the second empirical study required the use of restrictions. The programs below provide examples of SAS and M*plus* programming code for: (1) a latent class model of gambling and its restrictions, (2) a latent class model for repeated measures of drinking and its restrictions, and (3) a multivariable latent class model that predicts gambling latent class membership from drinking latent trajectory class membership.

B.1 Fitting a Latent Class Model

This model fits the final model chosen to describe the gambling behavior of adolescents and young adults participating in Add Health, as discussed in Chapter 3. This model includes three indicators of gambling and has five classes. A restrictions file is first created, then the latent class model is fit:

```
DATA gamb_restr5;
INPUT param $ group variable $ respcat estlc1 estlc2 estlc3
      estlc4 estlc5;
DATALINES;
GAMMA 1 .      .   1 1 1 1 1
BETA  1 .      .   1 1 1 1 1
RHO   1 lotto  1   2 2 2 2 3
RHO   1 casino 1   2 3 2 3 3
RHO   1 other  1   2 2 3 3 3
RHO   1 lotto  2   12 12 12 12 13
RHO   1 casino 2   12 13 12 13 13
RHO   1 other  2   12 12 13 13 13
;
RUN;

PROC LCA DATA = addhlth.gamblers RESTRICT = gamb_restr5;
TITLE1 'Gambling LCA - 5 Classes';
NCLASS 5;
ITEMS lotto casino other;
CATEGORIES 2 2 2;
SEED 123456;
RUN;
```

This latent class model may also be fit with M*plus*:

```
TITLE: Gambling - All

DATA: FILE IS addhlthmplus_gamblers.dat;
```

```
VARIABLE: NAMES ARE newaid
                    lotto casino other
                    freq_1 freq_2 freq_3;
          USEV ARE lotto casino other;
          CATEGORICAL = lotto casino other;
          CLASSES = gamb(5);
          MISSING ARE ALL (999);

ANALYSIS: TYPE = MIXTURE MISSING;
          STARTS = 200 25;

MODEL:
%OVERALL%

MODEL gamb:
      %gamb#1%
      [lotto$1]   (4);
      [casino$1]  (4);
      [other$1]   (4);

      %gamb#2%
      [lotto$1]   (4);
      [casino$1]  (5);
      [other$1]   (4);

      %gamb#3%
      [lotto$1]   (4);
      [casino$1]  (4);
      [other$1]   (5);

      %gamb#4%
      [lotto$1]   (4);
```

```
    [casino$1]  (5);
    [other$1]   (5);

    %gamb#5%
    [lotto$1]   (5);
    [casino$1]  (5);
    [other$1]   (5);

!OUTPUT: TECH8;
```

B.2 Fitting a Latent Class Model for Repeated Measures

This model fits the final model chosen to describe the trajectories of drinking frequency over time of adolescents and young adults participating in Add Health, as discussed in Chapter 3. This model includes three indicators of drinking frequency and has four trajectory classes. A restrictions file is first created, then the latent class model for repeated measures is fit:

```
DATA drink_restr4;
INPUT param $ group variable $ respcat estlc1 estlc2 estlc3
      estlc4;
DATALINES;
GAMMA 1 .       .   1 1 1 1
BETA  1 .       .   1 1 1 1
RHO   1 freq_1  1   2 3 3 3
RHO   1 freq_2  1   2 2 3 3
RHO   1 freq_3  1   2 2 2 3
RHO   1 freq_1  2   12 13 13 13
RHO   1 freq_2  2   12 12 13 13
RHO   1 freq_3  2   12 12 12 13
;
RUN;
```

```
PROC LCA DATA = addhlth.drinkers RESTRICT = drink_restr4;
TITLE1 'Drinking LCA - 4 Classes';
NCLASS 4;
ITEMS freq_1 freq_2 freq_3;
CATEGORIES 2 2 2;
SEED 123456;
RUN;
```

This latent class model for repeated measures may also be fit with M*plus*:

```
TITLE: Drinking - All

DATA: FILE IS addhlthmplus_drinkers.dat;

VARIABLE: NAMES ARE newaid
                    lotto casino other
                    freq_1 freq_2 freq_3;
          USEV ARE freq_1 freq_2 freq_3;
          CATEGORICAL = freq_1 freq_2 freq_3;
          CLASSES = drink(4);
          MISSING ARE ALL (999);

ANALYSIS: TYPE = MIXTURE MISSING;
          STARTS = 200 25;

MODEL:
%OVERALL%

MODEL drink:
      %drink#1%
      [freq_1$1]  (2);
      [freq_2$1]  (2);
      [freq_3$1]  (2);
```

```
    %drink#2%
    [freq_1$1]  (3);
    [freq_2$1]  (2);
    [freq_3$1]  (2);

    %drink#3%
    [freq_1$1]  (3);
    [freq_2$1]  (3);
    [freq_3$1]  (2);

    %drink#4%
    [freq_1$1]  (3);
    [freq_2$1]  (3);
    [freq_3$1]  (3);

!OUTPUT: TECH8;
```

B.3 Fitting a Multivariable Latent Class Model

This model predicts gambling latent class membership from drinking latent trajectory class membership by modeling gambling and drinking simultaneously. In this model, gambling and drinking are specified to be dependent with drinking predicting gambling:

```
TITLE: Gambling and Drinking - Relation - All - Freq

DATA: FILE IS addhlthmplus_freq.dat;

VARIABLE: NAMES ARE newaid
                    lotto casino other
                    freq_1 freq_2 freq_3;
          USEV ARE lotto casino other
                   freq_1 freq_2 freq_3;
```

```
          CATEGORICAL = lotto casino other
                        freq_1 freq_2 freq_3;
          CLASSES = drink(4) gamb(5);
          MISSING ARE ALL (999);

ANALYSIS: TYPE = MIXTURE MISSING;
          STARTS = 200 25;

MODEL:
%OVERALL%

    gamb#1 on drink#1;
    gamb#1 on drink#2;
    gamb#1 on drink#3;

    gamb#2 on drink#1;
    gamb#2 on drink#2;
    gamb#2 on drink#3;

    gamb#3 on drink#1;
    gamb#3 on drink#2;
    gamb#3 on drink#3;

    gamb#4 on drink#1;
    gamb#4 on drink#2;
    gamb#4 on drink#3;

MODEL drink:
      %drink#1%
      [freq_1$1]  (2);
      [freq_2$1]  (2);
      [freq_3$1]  (2);
```

```
        %drink#2%
        [freq_1$1]  (3);
        [freq_2$1]  (2);
        [freq_3$1]  (2);

        %drink#3%
        [freq_1$1]  (3);
        [freq_2$1]  (3);
        [freq_3$1]  (2);

        %drink#4%
        [freq_1$1]  (3);
        [freq_2$1]  (3);
        [freq_3$1]  (3);

MODEL gamb:
        %gamb#1%
        [lotto$1]   (4);
        [casino$1]  (4);
        [other$1]   (4);

        %gamb#2%
        [lotto$1]   (4);
        [casino$1]  (5);
        [other$1]   (4);

        %gamb#3%
        [lotto$1]   (4);
        [casino$1]  (4);
        [other$1]   (5);

        %gamb#4%
        [lotto$1]   (4);
```

```
    [casino$1]  (5);
    [other$1]   (5);

    %gamb#5%
    [lotto$1]   (5);
    [casino$1]  (5);
    [other$1]   (5);

!OUTPUT: TECH8;
```

APPENDIX C

Programming Code for Empirical Study #3

All latent transition and multiprocess models examined in the third empirical study were estimated using M*plus*, as discussed in Chapter 4. The models examined in the third empirical study required the use of restrictions and starting values. The programs below provide examples of M*plus* programming code for: (1) a latent transition model of smoking and its restrictions and starting values, (2) a latent transition model of drinking and its restrictions and starting values, and (3) a multiprocess model.

C.1 Fitting a Latent Transition Model of Smoking

This model fits the final model chosen to describe the smoking behavior of adolescents participating in the NLSY97, as discussed in Chapter 4. This model includes two indicators of smoking and has three classes:

```
TITLE: Model 1B (Smoking LTA)
       Dissertation Example #3

DATA: FILE IS mplus1516cs.dat;

VARIABLE: NAMES ARE id gender age98 weight
                    c_mf_1 c_mf_2 c_mf_3
                    c_mi_1 c_mi_2 c_mi_3
                    a_mf_1 a_mf_2 a_mf_3
                    a_mi_1 a_mi_2 a_mi_3
                    a_d_1  a_d_2  a_d_3;
          USEV ARE  c_mf_1 c_mf_2
                    c_mi_1 c_mi_2;
          CATEGORICAL = c_mf_1 c_mf_2
                        c_mi_1 c_mi_2;
          CLASSES = t1_smk(3) t2_smk(3);
          MISSING ARE ALL (999);

ANALYSIS: TYPE = MIXTURE MISSING;
          STARTS = 500 50;
          PROCESSORS = 2;

MODEL:
%OVERALL%

    [t2_smk#1]  (3);
    [t2_smk#2]  (4);

    t2_smk#1 on t1_smk#1  (9);
    t2_smk#1 on t1_smk#2 (10);
    t2_smk#2 on t1_smk#1 (11);
    t2_smk#2 on t1_smk#2 (12);
```

```
MODEL t1_smk:
      %t1_smk#1%
      [c_mf_1$1*-3]  (22);
      [c_mf_1$2*2]   (23);
      [c_mi_1$1*-3]  (22);
      [c_mi_1$2*2]   (23);

      %t1_smk#2%
      [c_mf_1$1*-23]  (24);
      [c_mf_1$2*-3]   (25);
      [c_mi_1$1*-23]  (24);
      [c_mi_1$2*-3]   (25);

      %t1_smk#3%
      [c_mf_1$1*5.5]  (26);
      [c_mf_1$2*32]   (27);
      [c_mi_1$1*5.5]  (26);
      [c_mi_1$2*32]   (27);

MODEL t2_smk:
      %t2_smk#1%
      [c_mf_2$1*-3]  (22);
      [c_mf_2$2*2]   (23);
      [c_mi_2$1*-3]  (22);
      [c_mi_2$2*2]   (23);

      %t2_smk#2%
      [c_mf_2$1*-23]  (24);
      [c_mf_2$2*-3]   (25);
      [c_mi_2$1*-23]  (24);
      [c_mi_2$2*-3]   (25);

      %t2_smk#3%
```

```
      [c_mf_2$1*5.5]  (26);
      [c_mf_2$2*32]   (27);
      [c_mi_2$1*5.5]  (26);
      [c_mi_2$2*32]   (27);

!OUTPUT: TECH8;
```

C.2 Fitting a Latent Transition Model of Drinking

This model fits the final model chosen to describe the drinking behavior of adolescents participating in the NLSY97, as discussed in Chapter 4. This model includes three indicators of drinking and has three classes:

```
TITLE: Model 1A (Drinking LTA)
       Dissertation Example #3

DATA: FILE IS mplus1516cs.dat;

VARIABLE: NAMES ARE id gender age98 weight
                    c_mf_1 c_mf_2 c_mf_3
                    c_mi_1 c_mi_2 c_mi_3
                    a_mf_1 a_mf_2 a_mf_3
                    a_mi_1 a_mi_2 a_mi_3
                    a_d_1  a_d_2  a_d_3;
          USEV ARE  a_mf_1 a_mf_2
                    a_mi_1 a_mi_2
                    a_d_1  a_d_2;
          CATEGORICAL = a_mf_1 a_mf_2
                        a_mi_1 a_mi_2
                        a_d_1  a_d_2;
          CLASSES = t1_alc(3) t2_alc(3);
          MISSING ARE ALL (999);
```

```
ANALYSIS: TYPE = MIXTURE MISSING;
          STARTS = 500 50;
          PROCESSORS = 2;

MODEL:
%OVERALL%

    [t2_alc#1]  (1);
    [t2_alc#2]  (2);

    t2_alc#1 on t1_alc#1 (5);
    t2_alc#1 on t1_alc#2 (6);
    t2_alc#2 on t1_alc#1 (7);
    t2_alc#2 on t1_alc#2 (8);

MODEL t1_alc:
      %t1_alc#1%
      [a_mf_1$1*-5]  (13);
      [a_mf_1$2*1]   (14);
      [a_mi_1$1*-5]  (13);
      [a_mi_1$2*-1]  (14);
      [a_d_1$1*2]    (15);

      %t1_alc#2%
      [a_mf_1$1*5.5]  (16);
      [a_mf_1$2*6.5]  (17);
      [a_mi_1$1*5.5]  (16);
      [a_mi_1$2*6.5]  (17);
      [a_d_1$1*31]    (18);

      %t1_alc#3%
      [a_mf_1$1*-5]  (19);
```

```
      [a_mf_1$2*-2]  (20);
      [a_mi_1$1*-5]  (19);
      [a_mi_1$2*-2]  (20);
      [a_d_1$1*-2]   (21);

MODEL t2_alc:
      %t2_alc#1%
      [a_mf_2$1*-5]  (13);
      [a_mf_2$2*1]   (14);
      [a_mi_2$1*-5]  (13);
      [a_mi_2$2*1]   (14);
      [a_d_2$1*2]    (15);

      %t2_alc#2%
      [a_mf_2$1*5.5]  (16);
      [a_mf_2$2*6.5]  (17);
      [a_mi_2$1*5.5]  (16);
      [a_mi_2$2*6.5]  (17);
      [a_d_2$1*31]    (18);

      %t2_alc#3%
      [a_mf_2$1*-5]  (19);
      [a_mf_2$2*-2]  (20);
      [a_mi_2$1*-5]  (19);
      [a_mi_2$2*-2]  (20);
      [a_d_2$1*-2]   (21);

!OUTPUT: TECH8;
```

C.3 Fitting a Multiprocess Model

This model predicts drinking latent status membership from smoking latent status membership by modeling drinking development and smoking development simultaneously. The programming code presented below fits Model 8, as discussed in Chapter 4. In this model, drinking at times 1 and 2 and smoking at time 1 are conditionally dependent, and smoking at times 1 and 2 and drinking at time 2 are conditionally dependent:

```
TITLE: Model 5C FIXED (Interaction -- Lagged and Concurrent
       -- All Times Concurrent)
       Dissertation Example #3

DATA: FILE IS mplus1516cs.dat;

VARIABLE: NAMES ARE id gender age98 weight
                    c_mf_1 c_mf_2 c_mf_3
                    c_mi_1 c_mi_2 c_mi_3
                    a_mf_1 a_mf_2 a_mf_3
                    a_mi_1 a_mi_2 a_mi_3
                    a_d_1  a_d_2  a_d_3;
          USEV ARE  a_mf_1 a_mf_2
                    a_mi_1 a_mi_2
                    a_d_1  a_d_2
                    c_mf_1 c_mf_2
                    c_mi_1 c_mi_2;
          CATEGORICAL = a_mf_1 a_mf_2
                        a_mi_1 a_mi_2
                        a_d_1  a_d_2
                        c_mf_1 c_mf_2
                        c_mi_1 c_mi_2;
          CLASSES = t1_smk(3) t1_alc(3) t2_smk(3) t2_alc(3);
          MISSING ARE ALL (999);

ANALYSIS: TYPE = MIXTURE MISSING;
          STARTS = 500 50;
          PROCESSORS = 2;
```

```
MODEL:
%OVERALL%

    [t2_alc#1]  (1);
    [t2_alc#2]  (2);
    [t2_smk#1]  (3);
    [t2_smk#2]  (4);

    t2_alc#1 on t1_alc#1 (5);
    t2_alc#1 on t1_alc#2 (6);
    t2_alc#2 on t1_alc#1 (7);
    t2_alc#2 on t1_alc#2 (8);
    t2_smk#1 on t1_smk#1 (9);
    t2_smk#1 on t1_smk#2 (10);
    t2_smk#2 on t1_smk#1 (11);
    t2_smk#2 on t1_smk#2 (12);

    t1_alc#1 on t1_smk#1;
    t1_alc#1 on t1_smk#2;
    t1_alc#2 on t1_smk#1;
    t1_alc#2 on t1_smk#2;

    t2_alc#1 on t2_smk#1;
    t2_alc#1 on t2_smk#2;
    t2_alc#2 on t2_smk#1;
    t2_alc#2 on t2_smk#2;

      t2_alc#1 on t1_smk#1;
      t2_alc#1 on t1_smk#2;
      t2_alc#2 on t1_smk#1;
      t2_alc#2 on t1_smk#2;

MODEL t1_alc:
```

```
      %t1_alc#1%
      [a_mf_1$1*-5]  (13);
      [a_mf_1$2*1]   (14);
      [a_mi_1$1*-5]  (13);
      [a_mi_1$2*-1]  (14);
      [a_d_1$1*2]    (15);

      %t1_alc#2%
      [a_mf_1$1*5.5]  (16);
      [a_mf_1$2*6.5]  (17);
      [a_mi_1$1*5.5]  (16);
      [a_mi_1$2*6.5]  (17);
      [a_d_1$1*31]    (18);

      %t1_alc#3%
      [a_mf_1$1*-5]  (19);
      [a_mf_1$2*-2]  (20);
      [a_mi_1$1*-5]  (19);
      [a_mi_1$2*-2]  (20);
      [a_d_1$1*-2]   (21);

MODEL t2_alc:
      %t2_alc#1%
      [a_mf_2$1*-5]  (13);
      [a_mf_2$2*1]   (14);
      [a_mi_2$1*-5]  (13);
      [a_mi_2$2*1]   (14);
      [a_d_2$1*2]    (15);

      %t2_alc#2%
      [a_mf_2$1*5.5]  (16);
      [a_mf_2$2*6.5]  (17);
      [a_mi_2$1*5.5]  (16);
```

```
      [a_mi_2$2*6.5]  (17);
      [a_d_2$1*31]    (18);

      %t2_alc#3%
      [a_mf_2$1*-5]  (19);
      [a_mf_2$2*-2]  (20);
      [a_mi_2$1*-5]  (19);
      [a_mi_2$2*-2]  (20);
      [a_d_2$1*-2]   (21);

MODEL t1_smk:
      %t1_smk#1%
      [c_mf_1$1*-3]  (22);
      [c_mf_1$2*2]   (23);
      [c_mi_1$1*-3]  (22);
      [c_mi_1$2*2]   (23);

      t2_alc#1 on t1_alc#1;
      t2_alc#1 on t1_alc#2;
      t2_alc#2 on t1_alc#1;
      t2_alc#2 on t1_alc#2;

      t2_alc#1 on t2_smk#1;
      t2_alc#1 on t2_smk#2;
      t2_alc#2 on t2_smk#1;
      t2_alc#2 on t2_smk#2;

      %t1_smk#2%
      [c_mf_1$1*-23]  (24);
      [c_mf_1$2*-3]   (25);
      [c_mi_1$1*-23]  (24);
      [c_mi_1$2*-3]   (25);
```

```
      t2_alc#1 on t1_alc#1;
      t2_alc#1 on t1_alc#2;
      t2_alc#2 on t1_alc#1;
      t2_alc#2 on t1_alc#2;

      t2_alc#1 on t2_smk#1;
      t2_alc#1 on t2_smk#2;
      t2_alc#2 on t2_smk#1;
      t2_alc#2 on t2_smk#2;

      %t1_smk#3%
      [c_mf_1$1*5.5]  (26);
      [c_mf_1$2*32]   (27);
      [c_mi_1$1*5.5]  (26);
      [c_mi_1$2*32]   (27);

MODEL t2_smk:
      %t2_smk#1%
      [c_mf_2$1*-3]  (22);
      [c_mf_2$2*2]   (23);
      [c_mi_2$1*-3]  (22);
      [c_mi_2$2*2]   (23);

      %t2_smk#2%
      [c_mf_2$1*-23]  (24);
      [c_mf_2$2*-3]   (25);
      [c_mi_2$1*-23]  (24);
      [c_mi_2$2*-3]   (25);

      %t2_smk#3%
      [c_mf_2$1*5.5]  (26);
      [c_mf_2$2*32]   (27);
      [c_mi_2$1*5.5]  (26);
```

```
      [c_mi_2$2*32]   (27);

!OUTPUT: TECH8;
```

References

Agresti, A. (2002). Loglinear models for contingency tables. In *Categorical data analysis* (chap. 8). Hoboken, NJ: Wiley-Interscience.

Aitkin, M., Anderson, D., & Hinde, J. (1981). Statistical modeling of data on teaching styles (with discussion). *Journal of the Royal Statistical Society, Series A*, *144*, 419-461.

Auerbach, K. J., & Collins, L. M. (2006). A multidimensional developmental model of alcohol use during emerging adulthood. *Journal of Studies on Alcohol*, *67*(6), 917-925.

Barnes, G. M., Welte, J. W., Hoffman, J. H., & Dintcheff, B. A. (1999). Gambling and alcohol use among youth: Influences of demographic, socialization, and individual factors. *Addictive Behaviors*, *24*(6), 749-767.

Barnes, G. M., Welte, J. W., Hoffman, J. H., & Dintcheff, B. A. (2002). Effects of alcohol misuse on gambling patterns in youth. *Journal of Studies on Alcohol*, *63*, 767-775.

Barnes, G. M., Welte, J. W., Hoffman, J. H., & Dintcheff, B. A. (2005). Shared predictors of youthful gambling, substance use, and delinquency. *Psychology of Addictive Behaviors*, *19*(2), 165-174.

Bien, T. H., & Burge, J. (1990). Smoking and drinking: A review of the literature. *International Journal of the Addictions*, *25*, 1429-1454.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge, MA: The MIT Press.

Blaszczynski, A., Huynh, S., Dumlao, V. J., & Farrell, E. (1998). Problem gambling within a Chinese speaking community. *Journal of Gambling Studies*, *14* (4), 359-380.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York, NY: John Wiley.

Brown, S., & Coventry, L. (1997). *Queen of hearts: The needs of women with gambling problems.* Melbourne, AU: Financial and Consumer Rights Council.

Bulik, C. M., Sullivan, P. F., & Kendler, K. S. (2000). An empirical study of the classifications of eating disorders. *American Journal of Psychiatry*, *157*, 886-895.

Bureau of Labor Statistics, U.S. Department of Labor. (n.d.). *National Longitudinal Survey of Youth 1997 cohort, 1997-2003 (rounds 1-7) [computer file].* Produced by the National Opinion Research Center, the University of Chicago and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH: 2005.

Chung, H. (2003). *Latent-class modeling with covariates.* Unpublished doctoral dissertation, The Pennsylvania State University, University Park, PA.

Chung, H., Flaherty, B. P., & Schafer, J. L. (2006). Latent class logistic regression: Application to marijuana use and attitudes among high school seniors. *Journal of the Royal Statistical Society, Series A*, *169* (Part 4), 723-743.

Chung, H., Park, Y., & Lanza, S. T. (2005). Latent transition analysis with covariates: Pubertal timing and substance use behaviors in adolescent females. *Statistics in Medicine*, *24*, 2895-2910.

Clogg, C. C. (1995). Latent class models. In G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (p. 311-359). New York, NY: Plenum Press.

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 [Monograph]. In L. M. Collins & L. Seitz (Eds.), *Advances in data analysis for prevention research.* Bethesda, 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., Hyatt, S. L., & Graham, J. W. (2000). Latent transition analysis as a way of testing models of stage-sequential change in longitudinal data. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multiple-group data: Practical issues, applied approaches, and specific examples* (p. 147-161). Hillsdale, NJ: Erlbaum.

Collins, L. M., & Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. *Multivariate Behavioral Research*, *27*, 131-157.

Derevensky, J. L., Gupta, R., & Winters, K. C. (2003). Prevalence rates of youth gambling problems: Are the current rates inflated? *Journal of Gambling Studies*, *19*, 405-425.

Dube, D., Freestone, M. H., & Ladouceur, R. (1996). Potential and probable pathological gamblers: Where do the differences lie? *Journal of Gambling Studies*, *12*(4), 419-430.

Duncan, T. E., Duncan, S. C., Stycker, L. A., Fuzhong, L., & Alpert, A. (1999). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications*. Mahwah, NJ: Erlbaum.

Everitt, B. S., & Dunn, G. (1988). Log-linear modeling, latent class analysis, or correspondence analysis: Which method should be used for the analysis of categorical data? In R. Langeheine & J. Rost (Eds.), *Latent trait and latent class models* (p. 109-128). New York, NY: Plenum Press.

Feigelman, W., Wallisch, L. S., & Lesieur, H. R. (1998). Problem gamblers, problem substance users, and dual-problem individuals: An epidemiological study. *American Journal of Public Health*, *88*, 467-470.

Fergusson, D. M., Horwood, L. J., & Lynskey, M. T. (1994). The comorbidities of adolescent problem behaviors: A latent class model. *Journal of Abnormal Child Psychology*, *22*(3), 339-354.

Flaherty, B. P., & Collins, L. M. (1999). *Modeling transitions in two stage-sequences simultaneously* (Tech. Rep. No. 99-33). University Park, PA: The Methodology Center, Penn State.

Frank, M. L., Lester, D., & Wexler, A. (1991). Suicidal behavior among members of Gamblers Anonymous. *Journal of Gambling Studies*, *7*, 249-254.

Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, *61*, 215-231.

Grant, J. E., Kusher, M. G., & Kim, S. W. (2002). Pathological gambling and alcohol user disorder. *Alcohol Research and Health*, *26*(2), 143-150.

Griffiths, M., & Sutherland, I. (1998). Adolescent gambling and drug use. *Journal of Community and Applied Social Psychology*, *8*, 423-427.

Hagenaars, J. A. (1993). *Loglinear models with latent variables* (No. 07-094). Newbury Park: Sage Publications.

Hagenaars, J. A. (1994). Latent variables in log-linear models of repeated observations. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (p. 329-352). Thousand Oaks, CA: Sage Publications.

Hardoon, K. K., & Derevensky, J. L. (2002). Child and adolescent gambling behavior: Current knowledge. *Clinical Child Psychology and Psychiatry*, *7*(2), 263-281.

Heinen, T. (1996). Log-linear models and latent class analysis. In *Latent class and discrete latent trait models: Similarities and differences* (p. 30-64). Thousand Oaks, CA: Sage Publications.

Horn, J. L. (2000). Comments on integrating person-centered and variable-centered research on problems associated with the use of alcohol. *Alcoholism: Clinical and Experimental Research*, *24*(6), 924-930.

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. Sherman (Eds.), *Multivariate applications in substance use research: New methods for new questions.* Hillsdale, NJ: Erlbaum.

Ibáñez, A., Blanco, C., Moreryra, P., & Saiz-Ruiz, J. (2003). Gender differences in pathological gambling. *Journal of Clinical Psychiatry*, *64*(3), 295-301.

Istvan, J., & Matarazzo, J. D. (1984). Tobacco, alcohol, and caffeine use: A review of their relatiosnhips. *Psychological Bulletin*, *95*, 301-326.

Jackson, K. M., Sher, K. J., & Schulenberg, J. E. (2005). Conjoint developmental trajectories of young adult alcohol and tobacco use. *Journal of Abnormal Psychology*, *114*(4), 612-626.

Jackson, K. M., Sher, K. J., & Wood, P. K. (2000a). Trajectories of concurrent substance use disorders: A developmental, typological approach to comorbidity. *Alcoholism: Clinical and Experimental Research*, *24*(6), 902-913.

Jackson, K. M., Sher, K. J., & Wood, P. K. (2000b). Prospective analysis of comorbidity: Tobacco and alcohol use disorders. *Journal of Abnormal Psychology*, *109*(4), 679-694.

Jacobs, D. F. (2000). Juvenile gambling in North America: An analysis of long-term trends and future prospects. *Journal of Gambling Studies*, *16*(2/3), 119-151.

John, U., Meyer, C., Rumpf, H., & Hapke, U. (2003). Probabilities of alcohol high-risk drinking, abuse, or dependence esimated on grounds of tobacco smoking and nicotine dependence. *Addiction*, *98*, 805-814.

Kaminer, Y., & Haberek, R. (2004). Pathological gambling and substance use. *Journal of the American Academy of Child and Adolescent Psychiatry*, *43*(11), 1326-1327.

Kassinove, J. I., Doyle, K. A., & Milburn, N. G. (2000). Gambling and alcohol use in adolescence. *Journal of Social Behavior and Personality*, *15*(1), 51-66.

Kim, J.-O., & Mueller, C. M. (1978a). *Introduction to factor analysis: What it is and how to do it* (No. 13). Thousand Oaks, CA: Sage Publications.

Kim, J.-O., & Mueller, C. M. (1978b). *Factor analysis: Statistical methods and practical issues* (No. 14). Thousand Oaks, CA: Sage Publications.

Kline, R. B. (1998). *Principles and practice of structural equation modeling.* New York, NY: Guilford Press.

Korn, D. A., & Shaffer, H. J. (1999). Gambling and the health of the public: Adopting a public health perspective. *Journal of Gambling Studies*, *15*(4), 289-365.

LaBrie, R. A., Shaffer, H. J., LaPlante, D. A., & Wechsler, H. (2003). Correlates of college student gambling in the United States. *Journal of American College Health*, *52*(2), 53-62.

Ladd, G. T., & Petry, N. M. (2003). A comparison of pathological gamblers with and without substance abuse treatment histories. *Experimental and Clinical Psychopharmacology*, *11*(3), 202-209.

Ladouceur, R., Dubb, D., & Bujold, A. (1994). Prevalence of pathological gambling and related problems among college students in the Québec Metropolitan Area. *Canadian Journal of Psychiatry*, *39*, 289-293.

Lanza, S. T., & Collins, L. M. (2002). Pubertal timing and the onset of substance use in females during early adolescence. *Prevention Science*, *3*(1), 69-82.

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*(4), 552-561.

Lanza, S. T., & Collins, L. M. (under review). A new SAS procedure for latent transition analysis: The development of dating and sexual risk behavior. *Developmental Psychology*.

Lanza, S. T., Collins, L. M., & Lemmon, D. (in press). PROC LCA: A SAS procedure for latent class analysis. *Structural Equation Modeling*.

Lanza, S. T., Flaherty, B. P., & Collins, L. M. (2003). Latent class and latent transition analysis. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Vol. 2, research methods in psychology* (p. 663-685). Hoboken, NJ: Wiley.

Lanza, S. T., Lemmon, D., Schafer, J. L., & Collins, L. M. (2007). *PROC LCA & PROC LTA user's guide, version 1.1.3.* University Park, PA.

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis.* Boston, MA: Houghton Mifflin.

Lesieur, H. R. (1979). The compulsive gambler's spiral of options and involvement. *Psychiatry*, *42*, 79-87.

Lesieur, H. R., & Blume, S. B. (1991). When lady luck loses: Women and compulsive gambling. In N. Van Den Bergh (Ed.), *Feminist prespectives on addictions* (p. 181-197). New York, NY: Springer.

Lesieur, H. R., Blume, S. B., & Zoppa, R. M. (1986). Alcoholism, drug abuse, and gambling. *Alcoholism: Clinical and Experimental Research*, *10*(1), 33-38.

Lesieur, H. R., Cross, J., Frank, M. L., Welch, M., White, C. M., Rubenstein, G., et al. (1991). Gambling and pathological gambling among university students. *Addictive Behaviors*, *16*, 517-527.

Lesieur, H. R., & Rosenthal, R. J. (1991). Pathological gambling: A review of the literature. *Journal of Gambling Studies*, *7*(1), 5-39.

Loken, E. (2004). Using latent class analysis to model temperament types. *Multivariate Behavioral Research*, *39*(4), 625-652.

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.

Miller, M. A., & Westermeyer, J. (1996). Gambling in Minnesota. *American Journal of Psychiatry*, *153*, 845.

Muthén, B. O., & Sheddon, K. (1999). Finite mixture modeling with mitxure outcomes using the EM algorith. *Biometrics*, *55*, 463-469.

Muthén, L. K., & Muthén, B. O. (1998-2006). *M*plus *user's guide* (Fourth ed.). Los Angeles, CA: Muthén & Muthén.

Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods*, *4*, 139-157.

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

National Gambling Impact Study Commission. (1999). *National Gambling Impact Study Commission final report*. Washington, DC: National Gambling Impact Study Commission.

National Research Council. (1999). *Pathological gambling: A critical review*. Washington, DC: National Academy Press.

Neighbors, C., & Larimer, M. E. (2004). Self-determination and problem gambling among college students. *Journal of Social and Clinical Psychology*, *23*(4), 565-583.

Nower, L., Derevensky, J. L., & Gupta, R. (2004). The relationship of impulsivity, sensation seeking, coping, and substance use in youth gamblers. *Psychology of Addictive Behaviors*, *18*(1), 49-55.

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

Petry, N. M. (2002). A comparison of young, middle-aged, and older adult treatment-seeking pathological gamblers. *The Gerontologist*, *42*(1), 92-99.

Petry, N. M. (2005). *Pathological gambling: Etiology, comorbidity, and treatment*. Washington, DC: American Psychological Association.

Petry, N. M., Stinson, F. S., & Grant, B. F. (2005). Comorbidity of DSM-IV pathological gambling and other psychiatric disorders: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry*, *66*(5), 564-574.

Petry, N. M., & Tawfik, Z. (2001). A comparison of problem gambling and non-problem gambling youths seeking treatment for marijuana abuse. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 1324-1331.

Phillips, D. P., Welty, W. R., & Smith, M. M. (1997). Elevated suicide levels associated with legalized gambling. *Suicide and Life-Threatening Behavior*, *27*, 373-378.

Potenza, M. N., Steinberg, M. A., McLaughlin, S. D., Wu, R., Rounsaville, B. J., & O'Malley, S. S. (2001). Gender-related differences in the characteristics of problem gamblers using a gambling helpline. *American Journal of Psychiatry*, *158*, 1500-1505.

Ramirez, L. F., McCormick, R. A., Russo, A. M., & Taber, J. I. (1984). Patterns of substance abuse in pathological gamblers undergoing treatment. *Addictive Behaviors*, *8*, 425-428.

Raylu, N., & Oei, T. P. (2002). Pathological gambling: A comprehensive review. *Clinical Psychology Review*, *22*(7), 1009-1061.

Reboussin, B. A., & Anthony, J. C. (2001). Latent class marginal regression models for modeling youthful drug involvement and its suspected influences. *Statistics in Medicine*, *20*, 623-639.

Reboussin, B. A., Reboussin, D. M., Liang, K., & Anthony, J. C. (1998). Latent transition modeling of progression of health-risk behavior. *Multivariate Behavioral Research*, *33*(4), 457-478.

Rindskopf, D. (1990). Non-standard log-linear models. *Psychological Bulletin*, *108*(1), 150-162.

Rohde, P., Lewinsohn, P. M., & Seeley, J. R. (1995). Psychiatric comorbidity with problematic alcohol use in high school students. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*, 101-109.

Rosenthal, R. J., & Lorenz, V. C. (1992). The pathological gambler as criminal offender: Comments on evaluation and treatment. *Psychiatric Clinics of North America*, *15*, 647-660.

Schulenberg, J., O'Malley, P. M., Bachman, J. G., Wadsworth, K. M., & Johnston, L. D. (1996). Getting drunk and growing up: Trajectories of frequent binge drinking during the transition to young adulthood. *Journal of Studies on Alcohol*, *57*, 289-304.

Schulenberg, J., Wadsworth, K. M., O'Malley, P. M., Bachman, J. G., & Johnston, L. D. (1996). Adolescent risk factors for binge drinking during the transition to young adulthood: Variable-centered and pattern-centered approaches to change. *Developmental Psychology*, *32*, 659-674.

Shaffer, H. J., & Hall, M. N. (1996). Estimating the prevalence of adolescent gambling disorders: A quantitative synthesis and guide toward standard gambling nomenclature. *Journal of Gambling Studies*, *12*, 193-214.

Shaffer, H. J., & Hall, M. N. (2001). Updating and refining prevalence estiamtes of disordered gambling behavior in the United States and Canada. *Canadian Journal of Public Health*, *92*, 68-172.

Shaffer, H. J., Hall, M. N., & Vander Bilt, J. (1999). Estimating the prevalence of disordered gambling behavior in the United States and Canada: A research synthesis. *American Journal of Public Health*, *89*, 1369-1376.

Shaffer, H. J., LaBrie, R. A., LaPlante, D. A., Nelson, S. E., & Stanton, M. V. (2004). The road less traveled: Moving from distribution to determinants in the study of gambling epidemiology. *Canadian Journal of Psychiatry*, *49*(8), 504-516.

Sher, K. J., Gotham, H. J., Erickson, D. J., & Wood, P. K. (1996). A prospective, high-risk study of the relationship between alcohol and tobacco dependence and alcohol use disorders. *Alcoholism: Clinical and Experimental Research*, *20*, 485-492.

Smart, R. G., & Ferris, J. (1996). Alcohol, drugs and gambling in the Ontario adult population. *Canadian Journal of Psychiatry*, *41*, 36-45.

Snijders, T. A. B., & Bosker, R. (1999). *Multilevel analysis*. Thousand Oaks, CA: Sage.

Specker, S. M., Carlson, G. A., Edmonson, K. M., Johnson, P. E., & Marcotte, M. (1996). Psychopathology in pathological gamblers seeking treatment. *Journal of Gambling Studies*, *12*, 67-81.

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.

Stinchfield, R. D., Cassuto, N., Winters, K. C., & Latimer, W. (1997). Prevalence of gambling among Minnesota public school students in 1992 and 1995. *Journal of Gambling Studies*, *13*, 25-48.

Stinchfield, R. D., & Winters, K. C. (1998). Gambling and problem gambling among youth. *Annals of the American Academy of Political and Social Science*, *556*, 172-185.

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.

Tang, Z. (2002). *Using two-sequence latent transition analysis to model transitions in adolescent substance use and sexual behavior simultaneously* (Tech. Rep. No. 02-53). University Park, PA: The Methodology Center, Penn State.

Tavares, H., Martins, S. S., Lobo, D. S. S., Silveira, C. M., Gentil, V., & Hodgins, D. C. (2003). Factors at play in faster progression for female pathological gamblers: An exploratory analysis. *Journal of Clinical Psychiatry*, *64*(4), 433-438.

Tavares, H., Zilberman, M. L., Beites, F. J., & Gentil, V. (2001). Gender differences in gambling progression. *Journal of Gambling Studies*, *17*(2), 151-159.

Udry, J. R. (2003). *The National Longitudinal Study of Adolescent Health (Add Health), waves I & II, 1994-1996; wave III, 2001-2002* [Machine-readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill.

Velicer, W. F., Martin, R. A., & Collins, L. M. (1996). Latent transition analysis for longitudinal data. *Addiction*, *91*, S197-S209.

Vermunt, J. K. (1996). Causal log-linear modeling with latent variables and missing data. In U. Engel & J. Reinecke (Eds.), *Analysis of change: Advanced techniques in panel data analysis* (p. 35-60). Berlin/New York: Walter de Gruyter.

Vitaro, F., Brendgen, M., Ladouceur, R., & Tremblay, R. E. (2001). Gambling, delinquency, and drug use during adolescence: Mutual influences and common risk factors. *Journal of Gambling Studies*, *17*(3), 171-190.

Vitaro, F., Ferland, F., Jacques, C., & Ladouceur, R. (1998). Gambling, substance use, and impulsivity during adolescence. *Psychology of Addictive Behaviors*, *12*(3), 185-194.

Vitaro, F., Wanner, B., Ladouceur, R., Brendgen, M., & Tremblay, R. E. (2004). Trajectories of gambling during adolescence. *Journal of Gambling Studies*, *20*(1), 47-69.

Volberg, R. A. (1996). Prevalence studies of problem gambling in the United States. *Journal of Gambling Studies*, *12*(2), 111-128.

Volberg, R. A. (2007). Population surveys. In G. Smith, D. C. Hodgins, & R. J. Williams (Eds.), *Research and measurement issues in gambling studies* (p. 33-54). Burlington, MA: Elsevier.

von Eye, A., & Clogg, C. C. (Eds.). (1996). *Categorical variables in developmental research: Methods of analysis*. New York, NY: Academic Press.

Wallisch, L. (1993). *The 1992 Texas Survey of Adult Gambling Behavior.* Report to the Texas Commission on Alcohol and Drug Abuse.

Welte, J. W., Barnes, G. M., Wieczorek, W. F., Tidwell, M. O., & Parker, J. C. (2001). Alcohol and gambling pathology among U.S. adults: Prevalence, demographic patterns, and comorbidity. *Journal of Studies on Alcohol*, *62*(5), 706-712.

Welte, J. W., Barnes, G. M., Wieczorek, W. F., Tidwell, M. O., & Parker, J. C. (2004). Risk factors for pathological gambling. *Addictive Behaviors*, *29*(2), 323-335.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, *116*(2), 363-381.

Willett, J. B., & Sayer, A. G. (1996). Cross-domain analyses of change over time: Combining growth modeling and covariance structure analysis. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling* (p. 125-157). Mahwah, NJ: Erlbaum.

Winters, K. C., & Anderson, N. (2000). Gambling involvement and drug use among adolescents. *Journal of Gambling Studies*, *16*(2/3), 175-198.

Winters, K. C., Stinchfield, R. D., Botzet, A., & Anderson, N. (2002). A prospective study of youth gambling behaviors. *Psychology of Addictive Behaviors*, *16*(1), 3-9.

Winters, K. C., Stinchfield, R. D., Botzet, A., & Slutske, W. S. (2005). Pathways of youth gambling problem severity. *Psychology of Addictive Behaviors*, *19*(1), 104-107.

Winters, K. C., Stinchfield, R. D., & Fulkerson, J. (1993). Patterns and characteristics of adolescent gambling. *Journal of Gambling Studies*, *9*(4), 371-387.

Yusko, D. A., Buckman, J. F., White, H. R., & Pandina, R. J. (under review). Alcohol, tobacco, illicit drugs and performance enhancers: A comparison of use by college student athletes and non-athletes. *TBD*.

Zitzow, D. (1996). Comparative study of problematic gambling behaviors between American Indian and non-Indian adolescents within and near a Northern Plains reservation. *American Indian and Alaskan Native Mental Health Research*, *7*(2), 14-26.

Vita

Bethany Cara Bray

GENERAL INFORMATION

Office: The Methodology Center, College of Health and Human Development, The Pennsylvania State University; **Address:** 204 E. Calder Way, Suite 400, State College, PA 16801; **Email:** bcbray@psu.edu; **Telephone:** 814-865-0568; **Fax:** 814-863-0000.

EDUCATION

Ph.D.	Human Development and Family Studies, Penn State, August 2007
M.A.S.	Statistics, Penn State, August 2006
M.S.	Human Development and Family Studies, Penn State, May 2005
B.S.	Mathematics and Statistics, University of Michigan - Dearborn, December 2000

RESEARCH EXPERIENCE

Research Associate, The Methodology Center	07/2007 - present
Graduate Research Fellow, The Methodology Center	07/2005 - 06/2007
Graduate Research Assistant, The Methodology Center	01/2001 - 06/2005
Research Assistant, University of Michigan - Dearborn	11/1999 - 12/2000

SELECTED PUBLICATIONS

Ridenour, T. A., **Bray, B. C.**, & Cottler, L. B. (in press). Reliability of use, abuse, and dependence of four types of inhalants in adolescents and young adults. *Drug and Alcohol Dependence.*

Bray, B. C., Almirall, D., Zimmerman, R. S., Lynam, D., & Murphy, S. A. (2006). Assessing the total effect of time-varying predictors in prevention research. *Prevention Science, 7*, 1, 1-17.

Bray, B. C. (2005). *Modeling the development of substance use and delinquency simultaneously: An associative latent transition analysis example* (Tech. Rep. No. 03-59). University Park, PA: The Methodology Center, The Pennsylvania State University.

Wong, P., McAuslan, P., & **Bray, B. C.** (2000). *Survey of problem gambling in the Metropolitan Detroit Area, 2000.* Conducted under contract for the United Way Community Services.

IN PREPARATION

Bray, B. C. & Ridenour, T. A. Examining late childhood predictors of gambling participation and the desire to gamble in the future. In preparation for *Psychology of Addictive Behaviors.*

Ridenour, T. A., **Bray, B. C.**, Scott, H. S., & Cottler, L. B. Classification and assessment of substance use disorders in adolescents. In preparation for C. A. Essau (Ed.), *Substance abuse and dependence in adolescence.* New York, NY: Taylor & Francis.

SELECTED PRESENTATIONS

Bray, B. C. & Collins, L. M. *The structure of gambling behavior in adulthood.* Presented at the Society for Prevention Research 14th Annual Conference: "Applying Prevention Science to Reduce Health Disparities" (June 1, 2006).

Ridenour, T. A. & **Bray, B. C.** *Do correlates of preadolescent gambling resemble correlates of early substance use targeted in prevention programs?* Presented at the Society for Prevention Research 14th Annual Conference: "Applying Prevention Science to Reduce Health Disparities" (June 1, 2006).

Bray, B.C. *Rambling, gambling Willie: What Bob Dylan left out about pathological gambling.* Presented at the Centre County Correctional Facility, Bellefonte, PA (February 21, 2006).

RESEARCH INTERESTS

Discrete Latent Variable Methods: Loglinear modeling with latent variables including latent class analysis, conditional latent class analysis, multivariable latent class modeling, latent transition analysis, conditional latent transition analysis, and multiprocess modeling; model fit assessment, selection, and interpretation.

Problem and Pathological Gambling: Normative development of gambling behavior; risk and protective factors for gambling and problems with gambling; prevention and treatment of problems with gambling; measurement of gambling and problems with gambling.