Modeling Adolescent Substance Use Using Latent Transitional Analysis: The Health for Life Data Set

Zhiqun Tang
Stephanie Lanza
Linda M. Collins

The Pennsylvania State University

The Methodology Center
Technical Report Series
#01-44

College of Health and Human Development

The Pennsylvania State University

Modeling Adolescent Substance Use Using Latent Transition Analysis: The Healthy for Life Data Set

Zhiqun Tang

Stephanie Lanza

Linda M. Collins

The Pennsylvania State University

The Methodology Center

Abstract

The main purpose of the present study was fitting the onset use model and the recent use model to the Healthy for Life data set using latent transition analysis (LTA). Subgroup differences, where subgroups were defined according to gender and the timing of sexuality, were examined. Results from four studies were reported. A recent use model involving five latent statuses fits the HFL data set well. An onset use model does not fit the HFL data, which may be due to inconsistent responses across different measurement occasions by some adolescents. The same recent use model fits the data for both genders. There seem to be some interesting subgroup differences in adolescent substance recent use process related to the timing of sexuality, the strength and mechanisms of which need to be further examined.

Introduction

Substance use is a well-known risk factor for adolescent health and well-being in the United States (DiClemente, Hansen & Ponton, 1996), and some recent efforts have been made to understand adolescent substance use onset processes that can be conceptualized as a series of stages (Kandel & Faust, 1975; Kandel, 1989; Kandel, Yamaguchi & Chen, 1992; Kandel & Yamaguchi, 1993). A general pattern has been found in these studies that includes the following stages: no use of any kind of substance, use of alcohol and/or cigarettes, and then use of marijuana and/or other illicit substances (e.g., cocaine).

Recently, latent transition analysis (LTA) has been applied to questions about stage-sequential models (Graham, Collins, Wugalter, Chung, & Hansen, 1991; Collins, Graham, Long & Hansen, 1994; Collins, Graham, Rousculp, & Hansen, 1997; Collins, Hyatt, & Graham, 1998; Kam & Collins, 2000; Collins, Schafer, Hyatt, & Flaherty, in preperation). (LTA will be introduced in the next section). For example, a frequently asked question asked is whether the adolescent substance use onset process involves a hypothesized set of stages. LTA can be used to estimate and test such stage-sequential models. Although sequential patterns can be different and complicated (e.g., different (number of) stages, different transition restrictions among stages), some of them have been clearly evidenced in the above studies using various data sets such as the Adolescent Alcohol Prevention Trial (AAPT; Graham, Rohrbach, Hanson, Flay & Johnson, 1989), the National Longitudinal Survey of Youth (NLSY; Frankel, McWilliams & Spencer, 1983) and the National Longitudinal Study of Adolescent Health (Add Health; Resnick, et al., 1997). The results have suggested that adolescent substance use onset processes often start with alcohol or tobacco and/or experimentation with drunkenness, and then move upward to marijuana use. They also suggest that it is very unlikely for adolescents to experiment with marijuana without first trying alcohol or tobacco.

Generally, there are two types of stage-sequential models that have drawn attention from substance use researchers (Collins & Wugalter, 1992). One is the onset use model, which expresses the order in which individuals try new substances. In onset models development over time is generally cumulative with few developmental reversals. Such a model is illustrated in Figure 1.

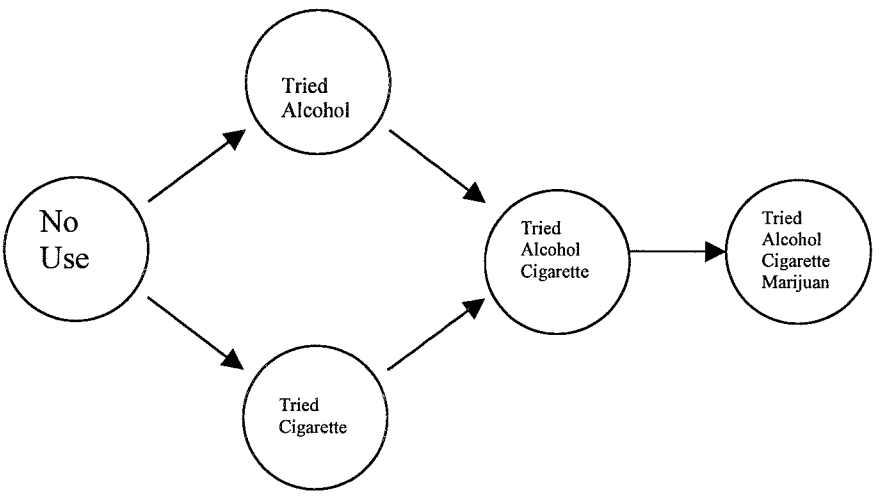

Figure 1 A Simple Onset Use Model

The circles in the figure represent stages in the onset process, and the arrows represent transitions between stages. The arrows indicate the directions of transitions that are permitted. Notice that only single-headed arrows appear in the figure. This indicates that only one-way transitions are permitted in the model. (Readers should note that the arrows in the model represent directions of stage transitions rather than regressions in structural equation models.) In other words, development takes place primarily in a forward direction, as individuals add new substances to their repertoire.

The other type of stage-sequential model is the recent use model, where there are developmental reversals or bidirectional changes, reflecting fluctuations in individuals' ongoing substance use behavior from one time to another. An extreme version of the recent use model is one that allows bidirectional changes between each pair of stages in the model (Figure 2). Most models that have been tested so far have been of the onset use variety, although they can include some bidirectional changes. For example, Collins et al.'s (1994, 1997) model includes a double-headed arrow between the "Tried Alcohol, Been Drunk" stage and the " Tried Alcohol, Been Drunk, Advanced Use" stage.

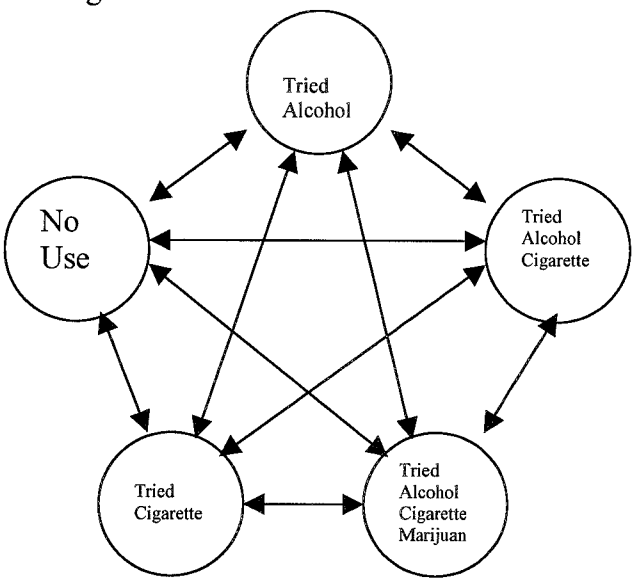

Figure 2 A Recent Use Model

Stage-sequential models can have important implications for prevention and intervention research once they have been established. Stage-sequential models with a reasonable goodness of fit for a particular data set can provide information about who, when and what should be the foci of prevention and intervention programs for adolescent substance use. Stage-sequential models enable researchers to define a population at risk, and to identify the specific characteristics or risk factors that are associated with each stage (Kandel, 1989). Programs may need to be designed differently in terms of content and strength for target individuals who are in different stages in the substance use sequence. For example, prevention programs for adolescent tobacco users may need to be different from programs for alcohol users or marijuana users. Graham et al. (1991) examined a stage-sequential model where six stages were involved. They found that adolescent substance use onset could start with either alcohol or tobacco, and those who tried tobacco first in the seventh grade appeared to have accelerated onset trajectories. In their study, the prevention program that had generally successful outcomes failed for the

individuals who started the onset process with tobacco. Therefore, the study suggested that there was a particular need for a prevention program targeting these individuals specifically.

A stage-sequential model can be tested across subgroups with LTA, providing further information for prevention and intervention programs to be tailored for target individuals. For example, in a recent study Collins et al. (1997) examined whether adolescents who were high or low caffeine users differed in initiation of the use of other substances. They suggested that heavy use of caffeine "is associated with an increased probability that an individual will initiate the onset process" (p95). Specifically, results showed that those who were heavy caffeine users in the seventh grade were 1.5 to 2.5 times less likely than those who were light caffeine users to be in the "No Use" stage in the seventh grade. Moreover, they were 1.2 to 1.9 times more likely than those light caffeine users to move out of the "No Use" drug stage and upward use stages involving substance use, particularly alcohol, between the seventh and the eighth grades. Although the results did not indicate any causal direction between caffeine use and substance use, it is possible that adolescents who are heavy caffeine users have some characteristics that can lead to substance use, i.e., they may be sensation-seekers. Thus, a high level of caffeine use can be a prospective risk factor for trying alcohol. The results could be informative for alcohol prevention researchers. For example, designing a prevention program for adolescent alcohol use that targets heavy caffeine users among children and adolescents would not only save money and manpower, but also might be more effective than a less targeted intervention.

Another example is a study by Hyatt and Collins (2000), where researchers examined the relationship between the onset of substance use and perceived parental permissiveness of adolescent alcohol use. The study found that the ninth grade adolescents who perceived their parents as highly permissive toward alcohol use were at higher risk for experimentation with various substances, and higher risk for upward to advanced substance use one year later. In contrast, adolescents who perceived their parents as strict were less likely to experiment with various substances in the ninth grade, and less likely to advance into high-level substance use stages. The study indicates that prevention and intervention programs may need to include family risk factors so as to facilitate parental monitoring and other helpful attitudes and behaviors, and thus, help make the programs more effective.

One of the major purposes of the present series of studies is to examine if fundamental sequential patterns similar to those found in prior literature can be replicated with a different data set—the Healthy for Life data set (HFL). In addition to modeling substance use onset and recent use, we also examine the HFL data for group differences in two domains. The first domain is gender. Studies on whether there are gender differences in substance use processes have provided mixed findings (Kandel, Yamaguchi & Chen, 1992, Kandel, 1989; Collins, Graham, Long, & Hansen, 1994; Kam & Collins, 2000). Kandel and colleagues found that progression to marijuana and other illicit substances among men who were from fifteen to thirty-five years old depended on their previous use of alcohol; whereas, either cigarettes or alcohol was a sufficient condition for progression to marijuana use among women of the same ages. Similar patterns held for boys in the twelfth grade in New York schools. For girls, alcohol *and* cigarettes preceded marijuana. These studies suggest some gender differences exist. On the other hand, Collins et al. (1994) found that there were no overall gender differences in the substance use patterns. These mixed findings are not necessarily contradictory because the different studies were based on different samples and methodologies. It is interesting to see if with LTA, certain patterns of substance use can be replicated in terms of gender differences with the Healthy for Life data set.

The second domain is sexual behavior, which has been found to be highly associated with substance use among adolescents. Rosenbaum & Kandel (1990) found that prior use of alcohol, cigarettes, marijuana and other illicit substances greatly increases the risk of early sex for adolescents. Moreover, the higher the stage of substance involvement, the greater the probability of early sex for adolescents. A similar relationship of substance use to sexual activity has also been demonstrated in other studies (Mott & Jean, 1988; Mensch & Kandel, 1992; Stewart, 1994; Graves & Leigh, 1995; Mott, Fondell, Hu, Kowaleski-Jones & Menaghan, 1996). On the other hand, some researchers (Graber, Petersen & Brooks-Gunn, 1996) argue that "entry into adolescence is marked by the physical changes of puberty" (pp23), and the biological changes can influence developmental trajectories in other domains (e.g., psychosocial, behavioral). In some sense, changes resulting from pubertal timing can be so prominent (e.g., being perceived as sexually attractive and a potential sex partner; Smith, Udry, & Morris, 1985) for adolescents that it could open the door to the onset of sexuality, which in turn may relate to substance use behaviors through motivational process (Capaldi, Crosby, & Stoolmiller, 1996) (e.g., wishing to show similar interests to that of their sex partners). It is interesting to see whether the substance use models fit for adolescents whose timing of sexuality is different (e.g., very early, early or normal timing of sexuality) when the LTA technique is used to analyze the HFL data set.

Methods

Latent Transition Analysis

Latent transition analysis (LTA)(Graham, Collins, Wugalter, Chung, & Hansen, 1991; Collins & Wugalter, 1992; Collins, Graham, Rousculp, & Hansen, 1997; Collins, Hyatt, & Graham, 1998; Hyatt & Collins, 1999; Hyatt, 1998; Kam & Collins, 2000) has been demonstrated to be a useful approach for testing stage-sequential models. LTA is rooted in latent class theory (refer to Goodman, 1974 and Clogg & Goodman, 1984 for details), which uses manifest variables or indicators (observed variables such as items in a questionnaire) to measure certain latent (unobserved or conceptualized) class membership of people (e.g., inhibited versus uninhibited, female versus male). The major difference is that LTA can model changes over time, in contrast to the traditional latent class approach.

Moreover, by modeling latent status membership nested in latent classes (e.g., gender), LTA can assess differences in changes in stage transitional across subgroups. Therefore, using manifest indicators (e.g., the item that measures sex, items that measure substance use), LTA can measure not only latent class memberships (the static part of LTA) but also changes of a specific latent stage sequence over time (the dynamic part of LTA). Specifically, if a latent variable is treated as static, that is, unchanging, over time, it is called a *static latent variable*; if a latent variable is treated as dynamic, that is, changing over time, it is called a *dynamic latent variable*. Transitions among *latent statuses*, i.e. stages, are the focus of LTA. In the present series of analyses, we are interested in how adolescent substance use onset or recent use changes from grade to grade, and how a five-latent-status model fits the HFL data set.

LTA models that include only the dynamic part provide two types of parameters to indicate change, which is the most informative part in LTA output. One is the transition probability matrix or the tau matrix (τ parameters) that provides specific transition probabilities from certain latent status at a given measurement occasion to other latent statuses at other given measurement occasions (e.g., the transition probabilities of "No Use" at time 1 to "Alcohol" or

"Cigarette" at time 2). The τ matrix is the most important part of LTA because it provides information about changes over time in detail. Another type of parameter is the delta (δ) parameters, which are probabilities of membership in certain latent statuses on a particular measurement occasion. Thus, the δ parameters also provide information about changes over time by allowing the user to compare how the sample is distributed across latent statuses at every measurement occasion.

In addition to the above two types of parameters that provide information about changes over time, there is one more type of parameter called "big" rho (ρ) parameters. Big ρ parameters are probabilities that people give particular answers to items, conditional on latent status and time. For example, a person would have a relatively high probability of answering "no" to the question of cigarette use if he/she belongs to the "No Use" latent status, and have a relatively low probability of answering "yes" to the same question. Therefore, higher (close to one) or lower (close to zero) probabilities both represent good measurement for latent statuses. The function of the big ρ parameters is conceptually similar to that of factor loadings in factor analysis. Thus, LTA models are measurement models. LTA models may incorporate a static part, in which case they involve a latent class model. Latent class models have two types of parameters. One type is gamma (γ) parameters, which represent the probability of membership in each latent class. The other is the "little" ρ parameters, which are directly analogous to the "big" ρ parameters discussed above, in other words, they represent the probability of a particular response to a particular item, conditional on latent class membership.

LTA models where both the static part and the dynamic part are included involve all five of the abovementioned types of parameters. In such models, parameters such as τ's, δ's and big ρ's are conditional not only on latent status membership and measurement occasions, but also on latent class membership. LTA models where only the static part is included are traditional latent class models.

The Healthy for Life Data Set

Healthy for Life (HFL) was a NIDA-funded health promotion program for middle school, junior high and high school children (6-10th graders) that sought to influence the behavior of adolescents in many areas including substance use (Moberg, Piper, Wu & Serlin, 1993; Piper, Moberg & King, 1999). In 1987, roughly 2500 6th-8th graders attending 21 Wisconsin middle schools completed a "baseline" questionnaire[1]. In 1988, an additional 2483 6th grade students were assessed and followed until the study ended in 1992, when most of these students were in 10th grade. Thus, there were in total 5 waves from 1988 to1992. The overall attrition was approximately 8% in wave two, 14% in wave 3, 20% in wave 4, and 30-33% in wave 5 (refer to Moberg, Piper and Wu (1993) for more information about the project and subjects).

Measures

In previous studies (Graham, Collins, Wugalter, Chung, & Hansen, 1991; Collins, Hyatt, & Graham, 1998), researchers have used items reflecting lifetime substance use as manifest indicators for latent statuses (e.g., No Use, Tried Alcohol, Tried Cigarettes, Tried Alcohol + Tried Cigarettes, Tried Alcohol + Cigarettes + Marijuana). After being informed by the

[1] Data from the baseline questionnaire were not included in the dataset. The investigators used questionnaire data obtained prior to the intervention to establish baseline levels for the students who participated in the program.

literature reviewed in the introduction section, and exploring the available items[2] in the HFL data set, the present studies aimed at testing a basic five-latent-status substance use onset model and recent use model using LTA.

The onset use model (Figure 1) was similar to the model used by Kam & Collins (2000). This model suggests that the patterns of adolescent substance use change over time, and the changes involve a series of stages: adolescents typically start out from "No Use" at younger ages to either using "Alcohol", or using "Cigarettes", then to both "Alcohol + Cigarettes", and finally try everything in "Alcohol + Cigarettes + Marijuana" at older ages. Moreover, developments in substance over time are cumulative without any developmental reversals. One of the differences between the recent use model (Figure 2) and the onset use model, as shown in the previous two figures, is that the recent use model that is tested in the present studies is fully bidirectional, capturing fluctuations in individuals' substance use behavior from one time to another. In other words, adolescents could move either backwards or forwards between any two stages rather than advance only in a forward direction. Therefore, different sets of indicators are used to measure the two models.

The primary instrument for the HFL data set was the annual questionnaire, administered in classroom settings during regular classroom hours by research staff from local universities. Students were asked to answer the questions honestly. For the dynamic latent variables in the substance use onset model, three items about lifetime alcohol, cigarette and marijuana use were used as the manifest indicators for each measurement occasion:

1. How many times in **your whole life** have you used beer, wine or liquor?
2. How many times in **your whole life** have you used cigarettes?
3. How many times in **your whole life** have you used marijuana?

These three items were originally measured on a 5-point scale, where 1= not at all, 2 = 1-2 times, 3 = 3-4 times, 4 = 5-8 times, and 5 = 9 or more times. For better capturing the nature of the onset processes, those items were recoded in the present studies. The data that were originally coded as 1 stayed the same, meaning "no use"; the data that were originally coded as 2, 3, 4, or 5 were recoded as 2, standing for "have used"; the missing data were recoded as 0.

For the recent use model in the present studies, three items about recent alcohol, cigarette and marijuana use were used as the manifest indicators for each measurement occasion:

1. How often **in the past month** did you use beer, wine or liquor?
2. How often **in the past month** did you use cigarettes?
3. How often **in the past month** did you use marijuana?

The three recent use indicators were recoded in the same way as those of the lifetime items.

For the manifest indicator measuring gender, one item at time 1 (in the sixth grade) was used, "Are you a boy or girl?" For measuring the timing of sexuality, two indicators were used, "How many times in your life, if ever, have you had sexual intercourse?" 0=never, 1=1 time, 2=2-4 times, 3=5-9 times, and 4=10 or more times. This variable at both time 3 (in the eighth grade) and time 4 (in the ninth grade) were used and recoded. The data that were originally coded as 0 was recoded as 1, meaning that "have not had sexual intercourse". The data that were originally coded as 1, 2, 3, or 4 were recoded as 2, meaning that "have had sexual intercourse (s)". The missing data were recoded as 0. How these two items were combined to measure the timing of sexual intercourse is addressed in study 4.

---

[2] Items for measuring drunkenness were not available in HFL data. Although drunkenness was suggested to be an important part of latent statuses in the sequential stages of substance use (refer to Collins et al's 1994 model), we could not include it in our model.

Analytical Procedures

The main purpose of the present studies is fitting the onset use model (Figure 1) and the recent use model (Figure 2) using LTA with the HFL data set (all the analyses were done by WinLTA software version 3.0). The actual data analyzed by LTA were response patterns, arrays of answers to the target manifest indicators at different measurement occasions (please refer to Table A.1 in Appendix A).

In the dynamic part of LTA, there are three types of parameters: δ's, big ρ's and τ's. To fit the models using LTA, certain restrictions need to be imposed on some of the parameters. For all models tested in the present studies, the δ parameters were freely estimated. For the onset use model, all the transitions indicated by the single-headed arrows in Figure 1 were freely estimated, and others were all fixed to 0 so that it was impossible for the students to move backwards (refer to Table A.2 for the details about the restrictions on the τ matrix in the onset use model). For the recent use model, the entire τ matrix was freely estimated (refer to Table A.3). In other words, all the bidirectional transitions indicated by the double-headed arrows were possible and freely estimated. For all models tested in the present studies, there are two types of measurement errors that can be idenfied for each indicator. One is the probability for those who should answer "no" according to their latent status to answer "yes"(e.g., the probability of an adolescent in the "No Use" latent status answering "yes" to the alcohol item). The other is the probability for those who should answer "yes" according to their latent status to answer "no" (e.g., the probability of an adolescent in the "Alcohol" latent status to answer "no" to the alcohol item). Big ρ parameters were constrained to be equal for each type of measurement error within each indicator, so that there is a single big ρ parameter to be estimated for each type of measurement error for each indicator. Thus, there are six big ρ parameters in total to be estimated for all three indicators, two per indicator. In addition, big ρ parameters were constrained to be equal across all measurement occasions (refer to Table A.4 in Appendix A for the big ρ constraints in all studies). When the big ρ's are constrained equal across time, the meaning of the stages remains identical across times. This aids in interpretation of the transition probabilities and other estimated parameters (much like factor invariance across measurement occasions or groups in factor analysis).

To examine further any gender differences and the relationship between substance use and the timing of sexuality, the static part of LTA was implemented as well as the dynamic part. There were two additional sets of parameters: γ parameters and little ρ parameters. For all the models tested in the present studies, the γ parameters were all freely estimated. In study 3, the little ρ parameters were fixed to either 1 or 0, where 1 represents perfect measurement for the corresponding class membership and 0 represents perfect measurement for the non-class membership (because gender is actually an observed variable). The little ρ parameters for measuring the timing of sexuality were freely estimated in study 4. In addition, big ρ parameters were constrained to be the same across subgroups, in order to allow meaningful comparisons. More information about parameter restrictions is also provided for each study in the next section.

Selected Results

Study 1: The onset use model

The results based on the analyses of two waves of data (grade eight in 1990 and grade nine in 1991) in the HFL data set are presented in this part. Individuals contributing partial data were included in the analyses.

The G-squared for the onset use model was 244.17 with 44 degrees of freedom. After adjusting for missing data, the G-squared was 147.759. The G-squared indicated a poor model fit.

The big ρ parameters (Table 1) represent the measurement quality of the manifest indicators. They were good with the exception of the marijuana measure for the latent status "alcohol + cigarettes + marijuana," which is highlighted in bold type. The big ρ parameter of this measure was 0.605, which was not very much different from 0.5—which is the weakest possible ρ when there are two response options. The meaning of this parameter is that 60.5% of those who had latent status membership "alcohol + cigarettes + marijuana" responded that they had used marijuana, and 39.5% responded that they had not used marijuana. Thus, this parameter was a weak parameter, suggesting that the answers to the indicator could not distinguish reliably whether or not the students were in this specific latent status. The other big ρ parameters were all good, consistent with the idea that there are five latent statuses underlying the response patterns given the indicators used in this study.

Table 1: Big ρ parameters in Study 1 (probability of responding "yes")

	ALC	CIG	MIJ
NOUSE	0.131	0.032	0.002
ALC	0.953	0.032	0.002
CIG	0.131	0.936	0.002
AC	0.953	0.936	0.002
ACM	0.953	0.936	**0.605**

Table 2: δ parameters in Study 1

	GRADE 8	GRADE 9
NOUSE	0.434	0.333
ALC	0.181	0.184
CIG	0.036	0.013
AC	0.264	**0.330**
ACM	0.085	**0.140**

The δ parameters (Table 2) are probabilities of latent status membership at each time. According to these estimates, more than half of the students in the sample had already tried at least one kind of substance by the eighth grade. Specifically, 43.4% of the sample in grade eight did not use any substance, but this number dropped to 33.3% when these students were in grade nine. In other words, 10.1% of the students moved out of the "No Use" stage between eighth grade and ninth grade. The proportions of students who were in "Alcohol + Cigarettes" and "Alcohol + Cigarettes + Marijuana" became larger at time two.

The transition probability matrix (Table 3) contains the τ parameters and shows the probabilities of latent status membership in grade nine (in columns) conditional on latent status membership in grade eight (in rows). The diagonal values are the probabilities that students remained in the same latent statuses. For example, among those who were "No Use" in grade eight, 76.6% remained in the same latent status, and 23.4% transitioned out of it, advancing to use of one or more substances one year later: 10.2% used alcohol, 10.6% used both alcohol and cigarette, and 2.2% tried all three kinds of substances. "Cigarettes" was the least stable status, only 29.3% remained in this status across the two times, and more than two-thirds moved to both alcohol and cigarettes (59.2%) or experimentation with all three substances (11.5%). Compared to those with "Alcohol" as the latent status membership in eighth grade, those in "Cigarettes" appeared to be on a faster track for advancing in substance use. Moreover, 12.5% of those who were alcohol and cigarette users in grade eight had already tried marijuana one year later.

Table 3: Transition probabilities in Study 1

	NOUSE	ALC	CIG	AC	ACM
NOUSE	0.766	0.102	0.005	0.106	0.022
ALC	0.000	0.775	0.000	0.179	0.046
CIG	0.000	0.000	**0.293**	0.592	0.115
AC	0.000	0.000	0.000	0.875	0.125
ACM	0.000	0.000	0.000	0.000	1.000

However, the poor fit of the onset use model suggested that a more flexible model might be more reasonable.

Study 2: The recent use model

In fact, we could not relax any restrictions imposed on the τ matrix in the first study because the indicators were all lifetime substance use measurements. For example, relaxing these restrictions would have allowed an individual to have tried a substance at one time and never to have tried it at a later time. Thus, it did not make any sense to relax the τ parameters that were fixed to zero in the first study, although this might have improved the model fit.

In this study, a different model, the recent use model, was tested using different indicators—substance use in the past month (refer back to the measures part in the methods section) for the same students. Again, individuals contributing partial data were included. For the parameter restrictions, the only difference between the recent use model and the onset use model is the τ matrix. For the recent use model, all τ parameters were freely estimated, indicating that students were allowed to transition between any two stages across the two times.

The G-squared for the recent use model was117.76 with 33 degrees of freedom, which is a considerably better fit than Model 1. After adjusted for missing data, the G-squared was 17.411, indicating a good fit of the model.

The big ρ parameters (Table 4) were all good for the recent use model. Therefore, the measurement quality for the recent use model was good, consistent with the idea that there were indeed five latent statuses underlying the response patterns given the indicators used to measure the latent statuses.

Table 4: Big ρ parameter estimates in Study 2 (probability of responding "yes")

	ALC	CIG	MIJ
NOUSE	0.002	0.006	0.001
ALC	0.966	0.006	0.001
CIG	0.002	0.955	0.001
AC	0.966	0.955	0.001
ACM	0.966	0.955	0.907

The δ parameters (Table 5) showed that 72.4% of the students in the sample in grade eight did not use any substance in *the past month* in grade eight, but this number dropped to 62.2% when these students were in grade nine in 1991. In other words, 10.2% of the students moved out of the "No Use" stage and proceeded to substance use between grade eight and grade nine.

Table 5: δ parameter estimates in Study 2

	GRADE 8	GRADE 9
NOUSE	0.724	0.622
ALC	0.106	0.129
CIG	0.044	0.051
AC	0.096	0.147
ACM	0.030	0.051

The transition probability matrix (Table 6) provides information about the movements between substance use stages in detail. (Please note that in recent use models, the diagonal of the transition probability matrix cannot be interpreted as the conditional probability of remaining in a latent status, because it is possible to make several transitions and end up at the same latent status.) The τ parameter estimates showed that conditional on latent status membership in grade eight (except for "No Use" status), more than 50% of students were moving backward and forward in terms of their latent status membership of recent substance use. For example, of those who were "Alcohol" in grade eight, 43.7% were in the same latent status in grade nine, and 25.6% moved backward to "No Use", 26.3% advanced to both alcohol and cigarettes, and 4.4% even tried all three kinds of substances. There was no movement from "Alcohol" only to "Cigarettes" only.

Table 6: Transition probabilities in Study 2

	NOUSE	ALC	CIG	AC	ACM
NOUSE	0.779	0.087	0.042	0.072	0.020
ALC	0.256	0.437	0.000	0.263	0.044
CIG	0.292	0.089	**0.204**	0.316	0.099
AC	0.116	0.148	0.113	0.468	0.156
ACM	0.223	0.068	0.032	0.257	0.420

"Cigarettes" was still the least stable status in this study. Only 20.4% remained in this status across the two times, 29.2% moved backward to "No Use", 8.9% shifted to "Alcohol", 31.6% advanced to both "Alcohol+ Cigarettes", and 9.9% tried all three kinds of substances.

Compared to those in the "Alcohol" latent status in eighth grade, those in the "Cigarettes" latent status appeared to be on a faster track of advancing in recent substance use between eighth grade and ninth grade. This is consistent with the results for the onset use model. For those who were in "Alcohol+ Cigarettes" in grade eight, only 11.6% moved backward to "No Use", others have used one, two, or three kinds of substances in ninth grade. For those who were in "Alcohol+ Cigarettes+ Marijuana", 42% stayed in the same status, 22.3% moved backward to "No Use", and 25.7% still used "Alcohol+ Cigarettes" in grade nine.

Study 3: The recent use model with a static component: gender differences

In this study, the recent use model was examined across gender. Thus, the static part of LTA was incorporated in the model. There were two more sets of parameters in addition to the δ, big ρ and τ parameters. They were the little ρ and γ parameters (refer back to the analytical procedures part in the methods section for the parameter restrictions).

The big ρ parameters for both males and females had similar constraints as those in study 2. The only difference was that there were two sets of the same parameter constraints—one for males and another for females in this study. The big ρ parameters were constrained equal across genders.

The G-squared for the recent use model tested across gender was 215.195with 72 degrees of freedom. The G-squared adjusted for missing data was 32.061, indicating a good model fit.

The big ρ parameters (Table 7) were all good for the recent use model across gender. Thus, the measurement quality for this model was good after the grouping variable gender was included.

Table 7: Big ρ parameter estimates in Study 3

	ALC	CIG	MIJ
NOUSE	0.000	0.001	0.001
ALC	0.968	0.001	0.001
CIG	0.000	0.954	0.001
AC	0.968	0.954	0.001
ACM	0.968	0.954	0.911

The γ parameter estimates showed that there were 49.5% males and 50.5% females in the data used for this study.

The overall pattern of the δ parameter estimates (Table 8) for the two genders did not appear to be much different from each other, although the exact values were not identical. At both times males appear to be slightly more advanced in substance use.

Table 8: δ parameter estimates in Study 3

	GRADE 8		GRADE 9	
	Male	Female	Male	Female
NOUSE	0.690	0.748	0.610	0.625
ALC	0.108	0.105	0.134	0.125
CIG	0.053	0.043	0.062	0.048
AC	0.097	0.096	0.130	0.165
ACM	0.053	0.008	0.065	0.037

The transition probabilities for both males and females are provided in Table 9. It appears that there are some differences between males and females. For those who were in the same latent status in grade eight and grade nine (diagonal values – remember this does not mean they remained in this latent status for the entire time), more males appeared to be in "Alcohol+ Cigarettes+ Marijuana", " Alcohol", and "Cigarettes" than females did. In contrast, more females tended to be in "Alcohol+ Cigarettes" than males did. Given membership in the latent status "Alcohol+ Cigarettes+ Marijuana" in Grade 8, 24.7% of males moved backward to "No Use", but no female moved backward to "No Use". The probabilities for females to move to "Alcohol+ Cigarettes" in ninth grade were all much bigger than those of males except for those who were in "No Use" in eighth grade. In contrast, males seemed to be moving more to the "Alcohol+ Cigarettes+ Marijuana" stage. It seems that although the prevalence of the latent statuses for both genders was similar, the dynamic backward and upward movements among latent statuses for males and females differed to some degree. We were not able to test the significance of the above differences in the current analyses. This could be examined in future analyses using data augmentation.

Table 9: Transition probabilities in Study 3

	NOUSE		ALC		CIG		AC		ACM	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
NOUSE	0.778	0.770	0.082	0.093	0.042	0.049	0.076	0.071	0.023	0.017
ALC	0.237	0.290	0.487	0.372	0.008	0.000	0.196	**0.324**	**0.072**	0.014
CIG	0.355	0.284	0.094	0.083	0.229	0.152	0.188	**0.439**	**0.135**	0.043
AC	0.167	0.070	0.175	0.121	0.181	0.053	0.366	**0.562**	0.112	0.194
ACM	**0.247**	**0.000**	0.057	0.168	0.042	0.000	0.214	**0.614**	**0.441**	0.218

Study 4: The recent use model incorporating a static component: the relationship between substance use and the timing of sexuality

In this study, the analysis involved three waves of data in grade six (1988), grade seven (1989) and grade eight (1990), and the static part was the timing of sexuality using items asked in grade eight and nine. Two indicators measured the timing of sexuality of students; a question asked in grade eight and a question asked in grade nine. For those whose timing of sexuality was "very early", it was expected that they would have high probabilities of giving a "yes" response to the item both when they were in grade eight and in grade nine. For those whose timing of sexuality was "early", it was expected that they would have high probabilities of giving a "yes" response only when they were in grade nine. For those whose timing of sexuality was "normal", they would have low probabilities of giving a "yes" response to the item both when they were in grade eight in 1990 and grade nine in 1991. Note that the little ρ parameters were not fixed, in other words, the latent class membership in this study was not perfectly measured like that in study 3. Other parameter restrictions on the τ and big ρ matrices were the same as those in study 3. This study allowed us to examine the relationship between recent substance use in early years (in grade six, seven, and eight) and the timing of sexuality (in grade eight[3] and grade nine) for the students.

[3] This was the earliest year when measurements of sexual behavior were available in the HFL data set.

Table 10: Little ρ parameter estimates in Study 4

	Lifesex3	Lifesex4
vEARLYT	1.000	**0.721**
EARLYTM	0.000	1.000
NORMTM	0.009	0.000

The model had a good fit (the G-squared was1334.615 with 1901 degrees of freedom, and the adjusted G-squared was 286.933). The little ρ parameter estimates (Table 10) were good except for the "very early timing" group measured by the item in grade nine. The sexual item in grade nine was somewhat weak for the very early timing group; however, the overall pattern is consistent with the idea that there are three latent classes of students in terms of the timing of sexuality.

The big ρ parameter estimates (Table 11) were good except for the "Alcohol+ Cigarettes+ Marijuana" latent status as measured by marijuana use.

Table 11: Big ρ parameter estimates in study 4 (yes)

	ALC	CIG	MIJ
NOUSE	0.016	0.012	0.000
ALC	0.908	0.012	0.000
CIG	0.016	0.926	0.000
AC	0.908	0.926	0.000
ACM	0.908	0.926	**0.774**

The γ parameter estimates suggest that 14.2% of the students in this study belonged to the "very early timing of sexuality" latent class, 11.3% belonged to the "early timing of sexuality" and 74.6% belonged to the "normal timing of sexuality" who had not had sexual intercourse by the ninth grade.

Table 12: δ parameter estimates in study 4

	GRADE 6			GRADE 7			GRADE 8		
	V-early	early	norm	V-early	early	norm	V-early	Early	norm
NOUSE	0.682	0.839	0.940	0.596	0.696	0.896	0.253	0.482	0.849
ALC	**0.122**	0.053	0.042	0.119	0.101	0.070	**0.132**	**0.144**	0.087
CIG	0.071	0.050	0.010	0.086	0.109	0.013	0.057	0.090	0.025
AC	0.076	0.058	0.007	0.142	0.071	0.020	**0.301**	**0.272**	0.037
ACM	**0.049**	0.000	0.001	0.057	0.022	0.001	**0.257**	0.012	0.002

These data strongly suggest that there are associations between substance use latent status and the timing of sexuality. The δ parameters (Table 12) suggest that students in the "very early timing of sexuality" latent class were generally on a fast track in terms of their substance use. At grade 6, when there are 94% of the "normal timing of sexuality" students and 83.9% of the "early timing of sexuality" students in the "No Use" status, almost one-third of the "very early timing of sexuality" students already have tried alcohol and other substances. Among the "very early timing of sexuality" students, 4.9% have tried marijuana, as compared to virtually none of

the other students. At grade 8, when 84.9% of the "normal timing of sexuality" students remained in "No Use", only one-fourth of those in the "very early timing of sexuality" latent class remained in "No Use".

The τ matrices (Tables 13-18) provide detailed information about stage transitions. A big difference among the three latent classes was the backward transitions for the latent status "Alcohol+ Cigarette+ Marijuana". Much fewer of the "very early timing of sexuality" students transitioned backward from grade 6 to grade 7 and from grade 7 to grade 8 than the other two latent classes. In contrast, all the "early timing of sexuality" and the "normal timing of sexuality" students were transitioning backward either to "No Use" or to "Alcohol+ Cigarette". In other words, more than two-thirds of the "very early timing of sexuality" students remained using three kinds of substances across the three years, while none in the other two latent classes continued to use marijuana. Thus, it seems that previous marijuana use was highly associated with very early timing of sexuality.

For the "very early timing of sexuality" latent class

Table 13: GRADE 6 (rows) To GRADE 7 (columns)

	NOUSE	ALC	CIG	AC	ACM
NOUSE	0.721	0.087	0.097	0.070	0.025
ALC	0.177	0.460	0.000	0.363	0.000
CIG	0.520	0.000	0.194	0.198	0.088
AC	0.398	0.051	0.074	0.465	0.011
ACM	0.329	0.000	0.000	0.000	0.671

Table 14: GRADE 7 (rows) To GRADE 8 (columns)

	NOUSE	ALC	CIG	AC	ACM
NOUSE	0.400	0.145	0.060	0.163	0.232
ALC	0.058	0.355	0.000	0.380	0.207
CIG	0.000	0.000	0.199	0.801	0.000
AC	0.057	0.000	0.000	0.631	0.313
ACM	0.000	0.066	0.065	0.000	0.870

For the "early timing sexuality" latent class

Table 15: GRADE 6 (rows) To GRADE 7 (columns)

	NOUSE	ALC	CIG	AC	ACM
NOUSE	0.797	0.067	0.067	0.042	0.027
ALC	0.102	0.638	0.260	0.000	0.000
CIG	0.000	0.145	0.779	0.076	0.000
AC	0.366	0.077	0.000	0.557	0.000
ACM	1.000	0.000	0.000	0.000	0.000

Table 16: GRADE 7 (rows) To GRADE 8 (columns)

	NOUSE	ALC	CIG	AC	ACM
NOUSE	0.596	0.104	0.109	0.191	0.000
ALC	0.357	0.393	0.000	0.128	0.122
CIG	0.136	0.254	0.000	0.610	0.000
AC	0.000	0.049	0.192	0.760	0.000
ACM	0.745	0.000	0.000	0.255	0.000

For the "normal timing of sexuality" latent class

Table 17: GRADE 6 (rows) To GRADE 7 (columns)

	NOUSE	ALC	CIG	AC	ACM
NOUSE	0.935	0.045	0.007	0.012	0.001
ALC	0.320	0.571	0.043	0.066	0.000
CIG	0.000	0.222	0.236	0.542	0.000
AC	0.475	0.181	0.343	0.000	0.000
ACM	0.000	0.000	0.000	1.000	0.000

Table 18: GRADE 7 (rows) To GRADE 8 (columns)

	NOUSE	ALC	CIG	AC	ACM
NOUSE	0.911	0.049	0.016	0.021	0.002
ALC	0.288	0.548	0.063	0.101	0.000
CIG	0.248	0.100	0.456	0.196	0.000
AC	0.412	0.177	0.000	0.411	0.000
ACM	1.000	0.000	0.000	0.000	0.000

Another big difference among the three latent classes was the forward transitions out of the latent status "No Use". Many more of the "very early timing of sexuality" students were transitioning out of "No Use" from grade 6 to grade 7 and from grade 7 to grade 8 than the other two latent classes. For example, from grade 7 to grade 8, 60% of the "very early timing of sexuality" students were transitioning out of "No Use", moving to one or more kinds of substance use; for the "early timing of sexuality" students, this number was 40.4%; for the "normal timing of sexuality" students, this number was only 9.9%. In other words, the majority of the "normal timing of sexuality" students, who have not had any sexual intercourse by the ninth grade, remained "No Use" in terms of their substance use status across the three years. In contrast, only 40% of the "very early timing of sexuality" students, who had already tried sexual intercourse as early as in the eighth grade, remained in "No Use" by the ninth grade. Although the transition probabilities cannot tell anything about causality between substance use and the timing of sexuality for the students, they do suggest that there are strong associations between the two.

Discussion

The main purpose of the present study is fitting the onset use model and the recent use model to the HFL data set using LTA. Study 1 fitted the onset use model with the two waves of data from grade eight and grade nine. Although the model had a poor fit, the big ρ parameters are consistent with the idea that there were five possible latent statuses underlying the response patterns given the indicators used in study 1: No Use, Tried Alcohol, Tried Cigarettes, Tried Alcohol +Cigarettes, Tried Alcohol + Cigarettes + Marijuana. These latent statuses were consistent with previous findings. Moreover, there was a clear trend that substance use was advancing as age increased. This was indicated not only by the higher percentages of substance use at the latter time, but also by the advancements from legal substances (e.g., alcohol, cigarettes) to the illicit substance (marijuana) from time one to time two. Many studies (e.g., Graham et al., 1991; Kam & Collins, 2000;Yamaguchi & Kandel, 1984; Chen & Kandel, 1995) in this area have reported the similar age advancing effect for adolescent drug use.

We think the strict restrictions imposed on the τ matrix, i.e. all the backward transitions were fixed to zero, may have contributed to the poor fit of the onset use model. After checking residuals (refer to Table A.5 in Appendix A for the details), we found some response patterns were illogical given lifetime substance use measures and resulted in poor residuals. For example, some students responded "no" to cigarette use item at time two after they had answered "yes" at time one. These response patterns present problems for an onset use model, and result in poor big ρ's and G-squared. Yet it is unclear whether or not the inconsistent answers to the same items across the two measurement occasions were due to participants' lying or due to a different understanding of the same question at different times. The obvious thing is that there were some measurement errors in using lifetime indicators to measure some participants' onset use behavior. Although relaxing some restrictions in the τ matrix might improve the model fit to some degree, it would not make any sense to do so.

Study 2 fits the recent use model with the same two waves of data. The recent use model is a substantially different model of substance use from the onset model, and uses different indicators. The model had good measurement quality and G-squared, which suggested that the recent use model had a good fit to the data. Although there have been no previous studies on the exact same model, it is reasonable to depict the adolescent substance use as multiple latent statuses, when recent use measurements (e.g., in past month) were used as indicators.

Again, study 2 suggested that fewer students remained in "No Use" in grade nine compared with that in grade eight, showing the age advancement trend in substance use during adolescence. Moreover, both study 1 and study 2 reported that "Cigarettes" was the least stable status, and those individuals in the latent status "Cigarettes" appeared to be on a faster track for advancing in substance use. The same finding has been reported by previous studies (e.g., Graham et al., 1991).

In general, the same model worked for both sexes in study 3. Gender differences were not very pronounced, which is consistent with previous findings by Collins and colleagues (Collins, Graham, Long, & Hansen, 1994; Kam & Collins, 2000) when they examined the substance use onset process among the AAPT adolescents (seventh and eighth grade) using LTA. In other words, the nature of the recent substance use process was the same for male and female students in eighth grade and ninth grade in this study. There were some small differences with respect to transition probabilities, however, the strength of which need to be further examined.

Study 4 suggested a strong relationship between previous substance use and the timing of sexuality. 14.2% of the sample belonged to the "very early timing of sexuality" and 11.3% belonged to the "early timing of sexuality" by the definition in the present study, which added up to more than one-fourth of the sample and was not a small number. More than half of these adolescents had previous or concurrent experiences with substance use. Adolescents in the "very early timing of sexuality" latent class were not only on a fast track of transitioning to advanced substance use, but also might be multiple substance users.

The findings were highly consistent with the previous findings mentioned in the introduction section. However, our results do not speak to the issue of whether it was substance use that contributed to early onset of sexuality, or these adolescents were generally deviants who had deviant pubertal timing and/or deviant attitude and value orientations, or whether other factors were operating (e.g., parental vs. peer influences; Hyatt & Collins, 2000; Kendel, 1989; 1996; Petraitis, Flay & Miller, 1995; ethnicity; Arredondo, Streit, Springer & Murry, 1993) to put these adolescents at higher risk for both substance use and sexual behaviors. The finding is subject to further studies with respect to both the strength of the relationship between substance use and the timing of sexuality and its mechanisms.

In conclusion, the recent use model of the five latent statuses fits the HFL data set well. The onset use model does not fit the HFL data well, which may be due to inconsistent responses across different measurement occasions by some adolescents. Further, the recent use models fit the data for both sexes. There seems to be some interesting subgroup differences (e.g., the timing of sexuality) in the adolescent recent substance use process, the strength and mechanisms of which need to be further examined.

References

Arredondo, R., Streit, K., Springer, N., & Murry, C. K. (1993). Ethnic and cultural factors in substance abuse: diagnosis, treatment, and prevention. Adolescent Medicine: State of the Art Reviews, 4(2), 263-276.

Capaldi, D. M., Crosby, L., & Stoolmiller, M. (1996). Predicting the timing of first sexual intercourse for at-risk adolescent males. Child Development, 67, 344-359.

Chen, K. & Kandel, D. B. (1995). The natural history of drug use from adolescence to the mid-thirties in a general population sample. American Journal of Public Health, 85(1), 41-47.

Clogg, C. C., & Goodman, L. A. (1984). Latent structure analysis of a set of multidimensional contingency tables. Journal of the American Statistical Association, 79, 762-771.

Collins, L. M., Graham, J. W., Long, J. D. & Hansen, W. B. (1994). Crossvalidation of latent class models of early substance use onset. Multivariate Behavioral Research, 29, 165-183.

Collins, L. M., Graham, J. W., Rousculp, S. S., & Hansen, W. B. (1997). Caffeine use and the beginning of the substance use onset process: An illustration of latent transition analysis. In K. Bryant, M. Windle, & S. West (Eds.), The Science of Prevention: Methodological Advances from Alcohol and Substance Abuse Research (pp79-99). Washington, DC: American Psychological Association.

Collins, L. M., Hyatt, S. L. & Graham, J. W. (1998). 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 (pp147-161). Hillsdale, NJ: Erlbaum.

Collins, L. M., Schafer, J. L., Hyatt, S. L., & Flaherty, B. F. (in preparation). Using latent class and latent transition models to analyze social science data: Testing hypotheses involving combinations of parameters.

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

DiClemente, R. J., Hansen, W. B., & Ponton, L. E. (1996). Handbook of Adolescent Health Risk Behavior. New York : Plenum Press.

Frankel, M. R., McWilliams, H. A. & Spencer, B. D. (1983). NLSY technical sampling report & addendum (1994). University of Chicago.

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

Graber, J. A., Petersen, A. C., & Brooks-Gunn, J. (1996). Pubertal processes: Methods, measures, and models. In J. A. Graber, J. Brooks-Gunn, & A. C. Peterson (Eds.), Transitions Through Adolescence: Interpersonal Domains and Context (pp23-53). Mahwah, NJ: Lawrence Earbaum Associates, Pubilishers.

Graham, J. W., Collins, L. M., Wugalter, S. E., Chung, N. K., & Hansen, W. B. (1991). Modeling transitions in latent stage-sequential processes: A substance use prevention example. Journal of Consulting and Clinical Psychology, 59, 48-57.

Graham, J. W., Rohrbach, L., Hanson, W. B., Flay, B. R., & Johnson, C. A. (1989). Convergent and discriminant validity for assessment of skill in resisting a role play alcohol offer. Behavioral Assessment, 11, 353-379.

Graves, K. L. & Leigh, B. C. (1995). The relationship of substance use to sexual activity among young adults in the United States. Family Planning Perspectives, 27(1), 18-22+33.

Hyatt, S. L., & Collins, L. M. (1999). Pubertal timing and the stages of substance use: Detailed results based on data augmentation. Technical report series #99-35. The Methodology Center, The Pennsylvania State University.

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. Hillsdale, NJ: Erlbaum, pp. 259-288.

Kam, C. & Collins, L. M. (2000). Latent transition analysis of substance use among adolescent in the National Longitudinal Survey of Youth. Technical report series #00-37. The Methodology Center, The Pennsylvania StateUniversity.

Kandel, D. B. (1996). The parental and peer contexts of adolescent deviance: An algebra of interpersonal influences. Journal of Drug Issues, 26(2), 289-315.

Kandel, D. B. (1989). Issues of sequencing of adolescent drug use and other problem behaviors. Drugs and Society, 3, 55-76. New York, NY: Hawthorne Press.

Kandel, D. B., & Faust, R. (1975). Sequences and patterns of adolescent drug use. Archives of General Psychiatry, 32, 923-932.

Kandel, D. B., & Yamaguchi, K. (1993). From beer to crack: Developmental patterns of drug involvement. American Journal of Public Health, 83, 851-855.

Kandel, D. B., Yamaguchi, K. & Chen, K. (1992). Stages of progression in drug involvment from adolescence to adulthood: further evidence for the gateway theory. Journal of Studies on Alcohol, 447-457.

Mensch, B. & Kandel, D. B. (1992). Drug use as a risk factor for premarital teen pregnancy and abortion in a national sample of young white women. Demography, 29(3) 409-429.

Moberg, D. P., Piper, D. L., & Wu, J. (1993). When total randomization is impossible: Nested randomized assignment. Evaluation Review, 17(3), 271-291.

Mott, F. L., Fondell, M. M., Hu, P. N. Kowaleski-Jones, L., & Menaghan E. G. (1996). The determinants of first sex by age 14 in a high-risk adolescent population. Family Planning Perspectives, 28(1), 13-18.

Mott, F. L. & Jean R. H. (1988). Linkages between sexual activity and alcohol and drug use among American adolescents. Family Planning Perspectives, 20(3), 128-136.

Petraitis, J., Flay, B. R., & Miller, T. Q. (1995). Reviewing theories of adolescent substance use: Organizing pieces in the puzzle. Psychological Bulletin, 117(1), 67-86.

Resnick, M. D., Bearman, P. S., Blum, R. W., Bauman, K. E., Harris, K. M., Jones, J., Tabor, J., Beuhring, T., Sieving, R. E., Shew, M., Ireland, M., Bearinger, L. H., Udry, J. R. (1997). Protecting adolescents from harm: Findings from the National Longitudinal Study of Adolescent Health. Journal of the American Medical Association, 278, 823-832.

Rosenbaum, E. & Kandel, D. (1990). Early onset of adolescent sexual behavior and drug involvement. Journal of Marriage and the Family, 52, 783-798.

Smith, E. A., Udry, J. R., & Morris, N. M. (1985). Pubertal development and friends: A biosocial explanation of adolescent sexual behavior. Journal of Health and Social Behavior, 26(3), 183-192.

Stewart M. (1994). High school students who use drugs are more likely than nonusers to engage in risky sexual. Family Planning Perspectives, 26(6), 279-280.

Yamaguchi, K. & Kandel, D. B. (1984). Patterns of drug use from adolescence to young adulthood: III. predictors of progression. American Journal of Public Health, 74(7), 673-681.

Appendix A

Data example, parameter restrictions and residuals

Table A.1: Selected data (or the response patterns) for study 1

000000	320
000111	9
000121	3
000211	13
000221	27
.	
.	
021222	1
101000	1
101111	3
101211	1
110111	4
.	
.	
121222	7
200221	1
201111	1
201211	1
201221	1
.	
.	
222121	3
222211	4
222212	2
222221	21
222222	52

Table A.2: The Tau matrix for the onset use model

TRANSITION PROBABILITIES
ROWS REPRESENT LATENT STATUS MEMBERSHIP AT TIME 1
COLUMNS REPRESENT LATENT STATUS MEMBERSHIP AT TIME 2

	NOUSE	ALC	CIG	AC	ACM
NOUSE	1	1	1	1	1
ALC	0	1	0	1	1
CIG	0	0	1	1	1
AC	0	0	0	1	1
ACM	0	0	0	0	1

Note: 1 refers to freely estimated
0 refers to being fixed to zero

Table A.3: The matrix of τ constraints for the recent use model

	NOUSE	ALC	CIG	AC	ACM
NOUSE	1	1	1	1	1
ALC	1	1	1	1	1
CIG	1	1	1	1	1
AC	1	1	1	1	1
ACM	1	1	1	1	1

Table A.4: The constraints imposed on the big ρ parameters' in all 4 studies

	Alcohol (yes)	Cigarette (yes)	Marijuana (yes)
NOUSE	2	3	4
ALC	5	3	4
CIG	2	6	4
AC	5	6	4
ACM	5	6	7

	Alcohol (no)	Cigarette (no)	Marijuana (no)
NOUSE	105	106	107
ALC	102	106	107
CIG	105	103	107
AC	102	103	107
ACM	102	103	104

Note: 2 OR GREATER MEANS CONSTRAINED EQUAL TO ANY OTHER PARAMETER WITH THE SAME DESIGNATION

Table A.5: Expected cell frequencies and residuals for the onset use model in study 1

	OBS	EXP	RESID	PEARSON
11111	451	432.6277	18.3723	0.8833
21111	68	76.3888	-8.3888	-0.9598
12111	7	15.4526	-8.4526	-2.1503*
22111	21	3.9519	17.0481	8.5758*
11211	1	0.7547	0.2453	0.2824
22211	7	0.1072	6.8928	21.0551*
11121	150	143.0232	6.9768	0.5834
21121	257	242.4739	14.5261	0.9329
12121	2	8.2068	-6.2068	-2.1666*
22121	32	33.9017	-1.9017	-0.3266
22221	4	2.0896	1.9104	1.3216
11112	24	21.7681	2.2319	0.4784
21112	4	7.5224	-3.5224	-1.2843
12112	15	15.6252	-0.6252	-0.1582
22112	21	20.7647	0.2353	0.0516
22212	3	1.502	1.498	1.2223
11122	78	79.1385	-1.1385	-0.128
21122	102	97.5785	4.4215	0.4476
12122	53	52.8544	0.1456	0.02
22122	356	380.5568	-24.5568	-1.2588
21222	1	2.2006	-1.2006	-0.8094
22222	21	30.3437	-9.3437	-1.6962
11111	1	0.7836	0.2164	0.2445
21111	1	0.1765	0.8235	1.9605
22121	1	4.1288	-3.1288	-1.5398
22221	2	3.1108	-1.1108	-0.6298
11122	10	8.78	1.22	0.4117
21122	15	12.9727	2.0273	0.5629
12122	7	6.5344	0.4656	0.1822
22122	63	60.1505	2.8495	0.3674
22222	52	45.4712	6.5288	0.9682

* Pearson residual is bigger than 2.

Acknowledgements

Sincere acknowledgements go to the other members of LTA group: David Wagstaff, Brian Flaherty, Chi-Ming Kam, Mildred Maldonado-Molina. The analyses would be less fruitful without their help: original HFL data preparation, constructive suggestions and comments with respect to model building, etc. This work was supported by NIDA grant P50 DA10075.