



Published in final edited form as:

Psychol Addict Behav. 2006 December ; 20(4): 373–384. doi:10.1037/0893-164X.20.4.373.

Comparison of Longitudinal Phenotypes Based on Number and Timing of Assessments: A Systematic Comparison of Trajectory Approaches II

Kristina M. Jackson and Kenneth J. Sher

Department of Psychological Sciences, University of Missouri—Columbia, and the Midwest Alcoholism Research Center

Abstract

The popularity of trajectory-based research to characterize developmental courses of alcohol (and other drug) involvement is growing rapidly. Given the increasing use of these methods, there is a need to identify methodological factors that affect course shape and prevalence. Using growth mixture models, the authors characterized the developmental course of 2 indices of alcohol involvement, alcohol use disorder and frequency of heavy drinking, with a prospective sample of 489 young adults (Year 1 age = 18.52; 55% female; 51% with family history of alcoholism) assessed 6 times over 11 years. Then, the authors explored the extent to which trajectory models that eliminated an assessment (at the beginning, middle, or end of the study interval) were similar to the full 6-wave model. Although classifications showed relatively high concordance, trajectory shape and predicted prevalences varied. Misclassification was associated with methodological factors such as probability of class membership and missing data. Findings suggest that researchers thoughtfully consider the nature of the phenomena being studied and the developmental period of interest when designing prospective studies.

Keywords

alcohol; drinking; trajectory; course; developmental

The popularity of trajectory-based research to characterize developmental courses of alcohol involvement is growing, but along with the allure of these methods arises a need to identify basic design factors that might affect course shape, prevalence, and membership. Bauer and Curran (2003) issued a general warning to exercise caution in drawing conclusions from these models with respect to distinguishing between a multiple class model and a single class model fit to nonnormal data (but see Muthén, 2003). With regard to the alcohol field, prior work has shown that the developmental course of alcohol involvement is conditional upon the specific indices used, suggesting that it may be hazardous to generalize across alternate indices of alcohol involvement (Jackson & Sher, 2005). We propose that many other design factors may affect the specific courses of alcohol involvement that are identified. A comparison across studies that identify developmental course reveals differences in course depending on age of the sample (with adolescent samples much more likely to reveal courses associated with escalation in drinking and young adult samples more likely to reveal courses associated with remission from drinking). Number of assessments, interval between assessments, and cohort

(i.e., year of study assessment) are other parameters that might function to alter what is observed with regard to developmental course (Faden et al., 2004). Although it is not possible to manipulate sample age or cohort or to manipulate the interval between assessments in the current study (without manipulating the number of assessments themselves), it is feasible to simulate the effect of the number of assessments and examine the extent to which alternate courses of alcohol involvement are observed.

Consistent with other psychometric research (Singer & Willett, 2003), studies examining reliability of growth (intercept and slope) parameters in latent growth models have shown that fewer assessments are associated with less reliable and more highly variable growth factors. This is analogous to the Spearman–Brown prophecy formula (Nunnally & Bernstein, 1994), which demonstrates that reliability is attenuated with a fewer number of items if we consider assessment occasions as items. Willett (1989) showed that reliability of change (the proportion of variance in the observed rate of change that is due to the true rate of change) increases with increasing numbers of assessments. Raudenbush and Chan (1993) demonstrated that one fewer assessment per participant (from five to four assessments) was associated with a relatively large loss in reliability of the slope (from .53 to .36) but not the intercept (.87 to .84). Finally, standard errors for growth factors and error parameters in a growth model, particularly slope parameters, were greater with increasingly fewer waves (Caskie, 1999), although minimal impact upon the parameter estimates themselves was observed on the basis of the number of waves alone. The analytic methods used to identify developmental trajectories, generally a form of mixture modeling, are based on latent growth modeling (as discussed below), and thus we would expect the number of assessments to play a role in their determination. However, no research that we identified has examined the impact of the number of assessments on the identification of discrete developmental trajectories.

In addition, the timing of assessments of drinking has implications for the resultant trajectories that are observed. As noted earlier, extant research on the developmental course of alcohol involvement suggests that, not surprisingly, younger samples have a greater proportion of individuals in a non- or low-using class, and with an increase in age, the increasing or later-onset class contains fewer individuals. Studies with samples that begin in adolescence (around age 12 or 13), in fact, are more likely to identify two distinct increasing classes (e.g., Chassin, Pitts, & Prost, 2002; Colder, Campbell, Ruel, Richardson, & Flay, 2002; Hill, White, Chung, Hawkins, & Catalano, 2000; Tucker, Orlando, & Ellickson, 2003) than studies with older (young adult) samples. Correspondingly, several studies with samples that begin in early adolescence (e.g., Chassin et al., 2002; Hill et al., 2000; White, Johnson, & Buyske, 2000) find no decreasing or developmentally limited class. Although the current study begins at age 18, we hypothesize that dropping an early assessment may result in smaller late-onset class (and a larger chronic class) due to the loss of information that serves to differentiate late-onset from chronic drinking courses. Likewise, dropping a late assessment may result in fewer individuals in the developmentally limited class (and more in the chronic class) because information is lost to differentiate chronic from developmentally limited drinking courses.

Determining Developmental Trajectory

There exist a variety of methods for determining the course of some behavior. Some of these techniques are designed to resolve the nature of “latent” or “unknown” subgroups in the data, including both nonparametric techniques (e.g., *K*-means and hierarchical cluster analysis) and parametric techniques (e.g., latent class and latent profile analysis). These are in contrast to techniques that use known subgroups such as sex and race as grouping variables. In addition, methods exist that are designed to chart the course of some behavior over time; these methods, mixture models, specifically account for the nature of prospective data. Mixture modeling approaches have recently gained popularity due to the user-friendly applications designed by

Nagin (Jones, Nagin, & Roeder, 2001; Nagin, 1999) and Muthén (Muthén, 2001a; Muthén & Muthén, 2000).

Study Goals

In the current study, we explored the extent to which the number of measurement occasions included in a growth mixture analysis of data from the same cohort altered the observed developmental courses (i.e., trajectories) obtained. We believe this is a critical issue because of the popular tendency to describe various growth mixture classes as fundamental properties of the underlying cohort. Such reification of these “trajectories” fails to appreciate the potential effects of critical design parameters, which are, at present, largely unknown. We systematically removed one assessment from a prospective study containing six waves (Years 1, 2, 3, 4, 7, and 11) at either the beginning (first assessment), the middle (fourth assessment), and the end (sixth assessment). This approach allows us to evaluate the consistency of results as a function of both the number of assessments as well as the total observation period (although we realize that the current study design precludes independent examination of each).

For the sake of efficiency, we opted to focus on only two indices of alcohol involvement for this objective (alcohol use disorder and frequency of heavy drinking). We selected these two measures in order to represent two distinct domains of alcohol involvement that are critical to adolescent and young adult drinking and associated problems. These items are consistent with much of the literature that examines the developmental course of heavy drinking as well as the general epidemiologic literature on drinking. In particular, we used heavy episodic drinking as a measure of alcohol consumption because it has been shown to be associated with the risk of experiencing a negative outcome such as injury, unplanned sexual behavior, or five or more alcohol-related problems (Wechsler, Lee, Kuo, & Lee, 2000) and alcohol use disorder (Knight et al., 2002).

Method

Sample and Procedure

A sample of 489 (47% male; 94% Caucasian) freshmen at a large midwestern university took part in a prospective study examining variables related to alcohol involvement (see Sher, Walitzer, Wood, & Brent, 1991). Participants were screened for family history of alcoholism using the Short Michigan Alcoholism Screening Test (Selzer, Vinokur, & van Rooijen, 1975) adapted for assessing parental alcoholism (Crews & Sher, 1992) and sections of the Family History–Research Diagnostic Criteria Interview (Endicott, Andreasen, & Spitzer, 1978). Participants were assessed via clinical interview and self-administered survey at baseline (mean age = 18.5 years) and were followed up five times over the subsequent 10 years (Years 2, 3, 4, 7, and 11). By Year 11, 410 (90% of 455 participants targeted for follow-up) were reinterviewed. The proportion of missing data for the interview was 1.0% at Time 2, 3.5% at Time 3, 3.7% at Time 4, 6.5% at Time 5, and 16.2% at Time 6; the proportion of missing data for the questionnaire was 0.4% at Time 1, 1.6% at Time 2, 4.1% at Time 3, 4.3% at Time 4, 7.8% at Time 5, and 19.0% at Time 6. The mean number of interview and questionnaire assessments completed was 5.69 ($SD = 0.83$; range = 1–6) and 5.63 ($SD = 0.89$; range = 1–6), respectively. At all waves, informed consent was given for participation in the research.

Measures

Alcohol use disorder (AUD)—Past-year alcohol abuse and dependence diagnoses were assessed using the Diagnostic Interview Schedule (DIS; Robins, Helzer, Croughan, Williams, & Spitzer, 1985): DIS Version III-A (Robins et al., 1985) at Years 1 and 2; DIS Version III-R (Robins, Helzer, Cottler, & Goldring, 1989) at Years 3, 4, and 7; and DIS Version IV (Robins,

Cottler, Bucholz, & Compton, 1997) at Year 11. To maintain continuity, diagnoses were made according to criteria from the *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.; American Psychiatric Association, 1980). A single AUD diagnosis was assigned if a participant met criteria for alcohol abuse, dependence, or both. Full sample diagnosis with AUD ranged from 9.8% (Wave 6) to 24.3% (Wave 1).

Heavy drinking—Frequency of past-month heavy drinking was assessed using a single item: “In the past 30 days, how many times have you had five or more drinks at a single sitting, either of beer, wine, wine coolers, liquor, or some combination of these?” Item responses ranged from 0 (*didn’t drink 5 or more drinks at a single sitting in the past 30 days*) to 7 (*every day*). Reliability was moderately good, with test–retest correlations ranging from $r = .37$ (over a 10-year interval, from Time 2 to Time 6) to $r = .68$ (over a 1-year interval, between Times 3 and 4). Across waves, the full sample mean ranged from $M = 0.30$ ($SD = 0.74$; Wave 6) to $M = 0.74$ ($SD = 1.19$; Wave 2).

Analytic Technique

To identify trajectories of heavy drinking, we used a mixture modeling procedure (Jones et al., 2001; Muthén, 2001a; 2001b; Muthén & Muthén, 2000; Nagin, 1999) using Mplus 3.01 (L. K. Muthén & Muthén, 1998–2004). General growth mixture modeling (GGMM) is a form of latent growth modeling, but it includes an unobserved categorical variable that models variability around the latent growth factors via discrete homogeneous classes of individuals (vs. representing variability with a parameter, as in growth modeling). Basically, these models mix the continuous nature of a latent growth curve model with the categorical nature of group membership, in a single estimation procedure. Rather than obtaining a trajectory of drinking for each individual in the study, as might be observed via latent growth modeling, multilevel modeling, or generalized estimating equations, GGMM groups individuals into meaningful clusters or classes.

Typical latent growth curve models assume that respondents come from the same population, with the same basic growth function with respect to starting point (intercept) and growth (slope), with individual variation represented by the intercept and slope factor variances. GGMM, however, allows for different populations to have unique intercepts and slopes. In essence, GGMM estimates a unique latent growth curve, with individual variability, for each underlying population. This technique has some important advantages over other techniques used to derive developmental courses of substance use (e.g., cluster analysis) because it treats group membership as a latent (error free) variable and accounts for the temporal ordering of prospective data. For applications of the GGMM technique in the substance use area, see Chassin et al. (2002); Colder et al. (2002); and Li, Barrera, Hops, and Fisher (2002).¹

For deriving trajectories of AUD, we used latent class growth analysis (LCGA), which is very similar to GGMM but is appropriate for categorical data. This technique is based on latent class analysis but incorporates the time-ordered nature of the data (Muthén, 2001a). For both LCGA and GGMM, the base model included intercept and linear slope; for AUD, a quadratic slope was also necessary. Both LCGA and LCGA provide class prevalences, and each participant receives a probability of class membership for each class, ranging from 0 to 1.

¹We identified classes based on the mean of the growth factors alone (i.e., we did not allow the growth factor variances to differ across classes) because freeing the variances across classes typically resulted in model nonconvergence. Other applications of GGMM have also distinguished classes based on growth factor means only (e.g., Chassin, Flora, & King, 2004; Colder et al., 2002; Tucker et al., 2003). Although Li et al. (2002) were able to model growth factor variances across classes, their model was limited to two classes, which is a relatively simple analytic model.

Results

Identification of Trajectories

First, we characterized course of AUD and frequency of heavy drinking for the full model (including all six waves). We estimated a series of models ranging from two to five classes for both AUD and heavy drinking. For both measures, the four-class model was the best-fitting model in terms of model fit, which was evaluated using information criteria fit indices (Bayesian information criterion, Schwartz, 1978; and Akaike's information criterion, Akaike, 1987) and model convergence, as shown in Table 1. Selection of the four-class solution for all models also allowed us to maximize meaningful comparisons. Table 1 also presents entropy for each model, which is a parameter indexing precision of classification based on posterior probability values (L. K. Muthén & Muthén, 1998–2004); an entropy value close to 1.0 indicates clear classification (little overlap among trajectories).

In order to make the five-wave model comparable to the six-wave model, it is critical to have identical participants in both models. However, we were faced with sample attrition after Year 1. To maximize comparability, models were estimated using direct (full information) maximum likelihood, which employs all available data for parameter estimation.²

Figure 1 (upper left) presents the six-wave model for AUD and Figure 2 (upper left) presents the six-wave model for heavy drinking. Class membership was characterized by a nondrinking or nonproblematic drinking course (61% and 60% for AUD and heavy drinking, respectively); a developmentally limited or remitter course, which decreased over time (7% and 18%); a later onset course, which increased over time (19% and 2%); and a (relatively) chronic course (13% and 20%). Next, we systematically removed data from the first, fourth, and sixth assessment and reestimated the developmental trajectories for each of our two measures of alcohol involvement. We note that, although we would have liked to drop more than one assessment, the four-wave LCGA model is not empirically identified with four classes (L. K. Muthén, personal communication, January 16, 2004).

Comparison of Full Model With Five-Wave Models

Eliminating the first assessment (NO_EARLY)—The drinking trajectories for AUD failed to converge without the first assessment. In fact, we were able to estimate only a two-class model, with a nondiagnosing group and a group that began high but declined over time.

Figure 2 (top right) presents the heavy drinking trajectories without Year 1. The trajectories looked nearly identical to the six-wave model, and course prevalences were very similar. Next, we assigned respondents to their most likely class and examined agreement between the two models by examining a contingency table of class membership between the five- and six-wave trajectories (see Table 2). Agreement between the five- and six-wave trajectories was very good, $\chi^2(9, N = 484) = 1,303.34, p < .001; \Phi = 1.64; \text{Cramér's } V = .95; \text{Cohen's } \kappa = .92$. There was a great deal of congruence among corresponding (or concordant) classes (e.g., chronic with chronic, developmentally limited with developmentally limited, etc.), as can be seen in Table 2. However, a number of those in the nondrinking class for the NO_EARLY model were categorized as developmentally limited when based on six waves. The converse was also true: A small number of those in the nondrinking class based on six waves were categorized as chronic or developmentally limited in the NO_EARLY model. There was also a very small tendency for chronic and developmentally limited classes to be interchanged.

²One outcome of modeling all missing data is that a respondent who partook in only one or two of the six assessments was still included in analyses. We estimated models requiring at least two waves of data and models requiring at least three waves of data. For both AUD and heavy drinking, we observed identical predicted prevalences and nearly identical trajectory shape for models estimated with the full sample, with respondents that had at least two waves of data, and with respondents that had at least three waves of data.

We presented the contingency table for the comparison of the NO_EARLY model with the full model (see Table 2), but presenting each of the pairwise contingency tables would be cumbersome. As such, we present a summary figure documenting trajectory agreement and disagreement by cell for each of our comparisons (see Figure 3) and report summary statistics (chi-squares, kappas) for classification agreement below. To determine what particular combinations of courses most contribute to agreement (or lack thereof) between two trajectories, the cell chi-square statistics for each cell were plotted using bar graphs. This is done for each of the six comparisons (three for AUD, three for heavy drinking). The four rows in Figure 3 refer to the full six-wave model trajectories, and the four columns refer to the (varying) five-wave model trajectories (depending on their placement, according to the legend). Within each cell of the table, there are six bars which will be referred to as *bar columns*. The Bar Columns A, B, and C refer to AUD and the Bar Columns D, E, and F refer to heavy drinking. So, the fourth bar column in Figure 3 (i.e., Bar Column D) corresponds to the heavy drinking comparison of the full model versus the model omitting Year 1. Note that there is no bar column a because the NO_EARLY model would not converge for AUD. Dark hatched bars reflect values that are greater than would be expected by chance, and light polka-dot bars indicate values that are lower than would be expected by chance, based on the marginals.

An examination of Bar Column D for the first row and first column (i.e., the concordant chronic classes) suggests that the value of the cell chi-square is 294, which corresponds to the table entry in Table 2 for the cell chi-square for that cell (cell $\chi^2 = 293.9$). Further, this bar column has dark hatching, indicating that this value is greater than would be expected by chance: An individual whose drinking is categorized as chronic in the full model is highly likely to be categorized as chronic in the NO_EARLY model. Likewise, Bar Column D for the first row and the second column (i.e., chronic drinking for the full model and developmentally limited drinking for the reduced model) shows a cell chi-square of 15, which corresponds to the table entry in Table 2 for that cell (cell $\chi^2 = 14.9$). In general, Figure 3 suggests that concordant courses were observed at a rate greater than would be expected by chance (i.e., agreement along the diagonal), and discordant courses were observed at a rate less than would be expected by chance (i.e., disagreement in the off-diagonal).

Eliminating the fourth assessment (NO_MIDDLE)—Figure 1 (lower left) presents the AUD trajectories without Year 4. There was similarity in shape to the six-wave model, but in the NO_MI-DDLE model, the chronic class remained high until the very last time point and the late-onset class had a much steeper slope, with a negative quadratic form. Also, the developmentally limited class had a much less steep decline during the middle years in the NO_MIDDLE model. Class prevalences showed many more individuals in the developmentally limited (17%) and nondiagnosing (70%) classes in the NO_MIDDLE model than in the six-wave model (7% and 61%, respectively). Correspondingly, there were fewer in the chronic (7%) and later-onset (7%) classes in the NO_MIDDLE model than in the six-wave model (13% and 19%). Still, classification agreement with the six-wave trajectories (see Figure 3, Bar Column B) was very good, $\chi^2(9, N = 489) = 704.18, p < .001; \Phi = 1.20$; Cramér's $V = .69$; Cohen's $\kappa = .67$. Correspondence was very high along the diagonal for all classes except for the nondiagnosing class. Some of those in the nondiagnosing class in one model were categorized as chronic or developmentally limited in the other model. In addition, the chronic class in the six-wave model was more likely to be categorized as the developmentally limited class in the NO_MIDDLE model, perhaps due to the similarity in their slopes between Years 7 and 11.

Figure 2 (lower left) presents the heavy drinking trajectories with Year 4 omitted. Again, the trajectories looked very similar to the six-wave model but were characterized by a somewhat smoother curve. Prevalences were similar for the chronic and later-onset classes but suggest a tendency for some nondrinkers in the six-wave model to be classified as developmentally

limited in the NO_MIDDLE model. Agreement with the six-wave trajectories (see Figure 3, Bar Column E) was excellent, $\chi^2(9, N = 489) = 1,227.05, p < .001$; $\Phi = 1.58$; Cramér's $V = .91$; Cohen's $\kappa = .87$. Concordant classes again showed high correspondence. However, those in the nondrinking class in one model had the tendency to be categorized as chronic or developmentally limited in the other model. Again, chronic and developmentally limited classes were somewhat interchanged.

Eliminating the final assessment (NO_END)—We expected omission of the last wave to be most informative in that it captures the greatest time span (with a 4-year interval since last assessment vs. other intervals which ranged from one to three years) as well as representing an important time point for resolving maturing out of alcohol involvement that would be expected to be observed by the end of the third decade of life (Dawson, Grant, Stinson, & Chou, 2004; Johnston, O'Malley, & Bachman, 2002; Muthén & Muthén, 2000). Figure 1 (lower right) presents the AUD trajectories. The trajectories looked remarkably similar to the six-wave model, with a slight tendency for the chronic class to show a greater decrease over time in the six-wave trajectory solution. Class prevalences were also similar to those in the six-wave model, with fewer in the chronic (11%) and later-onset (16%) classes and more in the nondiagnosing class (65%) than the six-wave model (13%, 19%, and 61%, respectively). Agreement between the five- and six-wave trajectories (see Figure 3, Bar Column C) was very good, $\chi^2(9, N = 489) = 1,194.88, p < .001$; $\Phi = 1.56$; Cramér's $V = .90$; Cohen's $\kappa = .90$. There was a great deal of congruence among corresponding classes, although nondiagnosing individuals in one model were frequently misclassified as chronic, developmentally limited, or later onset in the other model.

Figure 2 (lower right) presents the corresponding information for heavy drinking trajectories. In the five-wave model, the developmentally limited class showed a much more rapid decline, but the other courses appeared to be similar to the six-wave model. Prevalences, however, varied substantially: There were many more in the chronic class in the five-wave model than in the six-wave model (33% vs. 20%) and many fewer in the developmentally limited class in the five-wave model than in the six-wave model (6% vs. 18%), as was hypothesized. In addition, there were somewhat more in the later-onset class in the five-wave model than in the six-wave model (5% vs. 2%), and there were somewhat fewer in the nondrinker class in the five-wave model than in the six-wave model (57% vs. 60%). Agreement with the six-wave trajectories (see Figure 3, Bar Column F) was moderate, $\chi^2(9, N = 489) = 324.70, p < .001$; $\Phi = .81$; Cramér's $V = .47$; Cohen's $\kappa = .52$. Congruence along the diagonal was much lower for this comparison, particularly for the later-onset classes, which showed nearly no association. As with the prior five-wave model comparisons, those in the nondrinking class in one model had the tendency to be categorized as chronic or developmentally limited in the other model.

Understanding Misclassification: Inspection of Discordant Cases

In order to determine the extent to which method factors were associated with degree of concordance (being on vs. off the diagonal in Figure 3), we examined two factors for cases that were concordant versus discordant for both AUD and heavy drinking. The first factor was probability of class membership in the full sample model (which is, as we see it, the standard), and the second factor was presence of missing data. We expected that these factors might explain, in part, the misclassification of trajectories between the full model and the various five-wave models.

Probability of class membership—For AUD, there was a higher probability of class membership for concordant cases ($M = 0.90, SD = 0.13$) than for discordant cases ($M = 0.80, SD = 0.15$) for comparison of the full sample model and the NO_MIDDLE model, $F(1, 487)$

= 37.60, $p < .001$. Likewise, there was a higher probability of class membership for concordant ($M = 0.89$, $SD = 0.14$) than discordant cases ($M = 0.78$, $SD = 0.18$) in the comparison of the full versus NO_END models, $F(1, 487) = 14.40$, $p < .001$.

Next, we examined the specific concordant cells (4) and discordant cells (12) and the probability of class membership by each of the 16 cells. For comparison of the full sample model and the NO_MIDDLE model, no apparent pattern emerged, with a generally lower probability of class membership among all discordant cells, although the concordant cell for the decreasing classes exhibited lower probability than the other three concordant cells. Comparison of the full versus the NO_END model revealed that the cell misclassifying chronic drinkers according to the full model as increase drinkers according to the NO_END model had low probability of class membership ($M = 0.60$), as did the cell misclassifying increase drinkers according to the full model as decrease drinkers according to the NO_END model ($M = 0.57$). Both of these cells, however, occurred with low prevalence ($N = 4$ and $N = 3$, respectively).

For heavy drinking, there was a much higher probability of class membership for concordant ($M = 0.91$, $SD = 0.14$) than discordant cases ($M = 0.62$, $SD = 0.16$) for comparison of the full versus NO_EARLY models, $F(1, 482) = 82.89$, $p < .001$. This was true also for comparison of the full versus NO_MIDDLE models, $F(1, 487) = 216.24$, $p < .001$ ($M = 0.92$, $SD = 0.13$ and $M = 0.60$, $SD = 0.15$, respectively) and for comparison of the full versus NO_END models, $F(1, 487) = 50.12$, $p < .001$ ($M = 0.93$, $SD = 0.13$ and $M = 0.82$, $SD = 0.19$, respectively).

Again, we examined probability of class membership by cell. For comparison of the full versus NO_EARLY models, the cells misclassifying decrease drinkers or nondrinkers based on the full model as chronic drinkers according to the NO_EARLY model showed much lower probability of class membership ($M = 0.42$ and $M = 0.47$ for decreasing/chronic and nondrinking/chronic, - respectively); again, however, these cells occurred with low prevalence ($N = 2$ for both cells). For comparison of the full versus NO_MIDDLE models, the two cells that substituted chronic with - decreasing showed much lower probability of class membership ($M = .51$ and $M = 0.40$), as did the cell that misclassified chronic drinkers based on the full model as nondrinkers based on the NO_MIDDLE model ($M = 0.53$). However, the only cell of these that had more than one individual was the chronic-decrease misclassification ($N = 7$). Finally, for the comparison of the full sample model and the NO_END model, no apparent pattern emerged, with a generally lower probability of class membership among all discordant cells.

Presence of missing data—For AUD, there were more completed assessments (less missing data) for concordant ($M = 5.74$, $SD = 0.72$) than discordant cases ($M = 5.43$, $SD = 1.24$) for comparison of the full versus NO_MIDDLE models, $F(1, 487) = 9.46$, $p < .01$. However, for the full versus NO_END models, there was no difference in number of completed assignments for concordant ($M = 5.69$, $SD = 0.84$) versus discordant ($M = 5.62$, $SD = 0.58$) cases, $F(1, 487) < 1.0$, *ns*.

We examined the number of completed waves by cell for the comparison of the full versus NO_MIDDLE models. The cell that substituted the chronic group based on the full model with the decrease group based on the NO_MIDDLE model had fewer completed assessments ($M = 4.97$) than the remaining cells.

For heavy drinking, there were more completed assessments for concordant cases ($M = 5.68$, $SD = 0.75$) than for discordant cases ($M = 5.50$, $SD = 1.00$) for comparison of the full versus NO_EARLY models, although this difference failed to reach significance, $F(1, 482) = 1.12$, *ns*. Also, there were more completed assessments for concordant ($M = 5.70$, $SD = 0.78$) than discordant cases ($M = 4.75$, $SD = 1.52$) for comparison of the full versus NO_MIDDLE models,

$F(1, 487) = 40.93, p < .001$. This was not true for comparison of the full versus NO_END models: Concordant cases actually participated in (slightly) fewer assessments ($M = 5.61, SD = 0.92$) than discordant cases ($M = 5.69, SD = 0.79$), although this failed to reach significance, $F(1, 487) < 1.0, ns$.

We examined the number of completed waves by cell for comparison of the full versus NO_MIDDLE models. The four cells that substituted chronic with decrease drinking or chronic with nondrinking had many fewer completed assessments (M from 3.43 to 5.00) than the remaining cells.

Discussion

Our motivation for the current study was a simple one. When capturing the developmental course of alcohol involvement, what are the implications for the developmental stage based on when study is begun, the timing of assessments, and length of follow-up? When designing a developmentally sensitive longitudinal study, the researcher is often faced with having to strike a balance between these factors due to funding or logistical concerns. It is essential to capture key periods of differential growth, such as normative changes in onset, escalation, persistence, and desistence. Frequently, one wishes also to capture critical time points associated with changes for distinct subpopulations (latent or observed). Basically, we asked a very practical question, that is, whether we would observe the same developmental courses of drinking had we begun the study a year later (i.e., delaying the baseline assessment by one year), had participants not been assessed as frequently midstudy (e.g., due to budgetary constraints), or had we terminated the study four years early (e.g., had we lost funding).

To address this question, we manipulated assessments in order to omit time points at the beginning (Year 1), at the middle (Year 4), and at the end of the study (Year 11). These factors might influence observed findings for both normative change, via identification of the underlying growth model, and the identification of mixtures themselves. In doing so, we explored the extent to which timing influenced the developmental course for different indices of drinking (AUD and heavy drinking). Our findings have important implications for research not only in the alcohol field but also in developmental psychopathology, where use of growth mixture analysis is being used increasingly more frequently.

From some perspectives, the findings were somewhat reassuring, revealing generally high agreement between the full six-wave model and the models that manipulated number and timing of assessments. Contrary to expectation, dropping the first wave of assessment, roughly corresponding to age 18, did not affect the trajectory shape or predicted prevalences, although this model was only estimable for one of the two indices. Our failure to estimate a model for AUD with the first wave omitted is consistent with the fact that identification issues are more demanding for discrete variables such as AUD. According to Willett (1989), reliability of the latent growth parameters can be increased by grouping measurement occasions toward the beginning and end of the study. For the AUD models, perhaps the first assessment was necessary to distinguish classes beyond a two-class model. (In data not presented here, we observe a strong pattern of individuals who diagnose with AUD at the first wave but not at subsequent waves.)

Dropping the middle assessment resulted in different shapes and prevalences of developmental courses. This was particularly true for the developmentally limited class. This assessment corresponded to ages 21–22, a period during which many in our sample were preparing to transition from college into the workplace. The model without a middle assessment showed a higher prevalence for the developmentally limited class, and for AUD, a less steep course for this class. Also, for AUD, the late-onset class had a steeper slope for the model with the middle

assessment eliminated. Developmental courses that exhibit a positive or negative slope seem to require an assessment at the point at which growth occurs for accurate estimation of trajectory shape. This may be particularly true for assessment during a time of great transition, such as that surrounding college graduation and entry into the workplace. As a result of this misclassification, there was some blurring of chronic and developmentally limited course membership. A rapid decline in AUD prevalence and in heavy drinking rates is evident around the point of college graduation (for most participants, immediately after Year 4), and removing an assessment around this time has implications for the resultant trajectory shape (particularly for AUD; heavy drinking was estimated only with a linear slope, which may have limited the shape of the trajectory).

Finally, removing an assessment toward the end of the study, a time during which respondents were approximately 28–30 years of age, appeared to have little effect on trajectory shape, although, as per expectation, there were fewer in the late-onset class (as well as fewer in the chronic class) for AUD. For heavy drinking, however, contrary to our hypotheses, there were more participants in the late-onset class. It is interesting to note that the late-onset class for the five-wave model showed a very low association with the late-onset class for the full six-wave model, suggesting that, despite a similar trajectory shape, these were largely different classes. Some of those in the AUD late-onset class were misclassified as nondiagnosing individuals, which we would have expected to be more likely when a later wave was dropped (because the two groups are distinguished by their drinking behavior at later waves). These findings suggest that extending the study might not provide additional information about course shape, but it might possibly affect how the course is characterized on the individual level, possibly affecting observed relations between the developmental course and risk factors for drinking. Unfortunately, our complex models and limited number of assessments prohibited examination of developmental course using fewer waves, such as also removing the fifth wave to limit data to college years only.

We determined that misclassification was in part due to methodological factors such as probability of class membership and presence of missing data. We noted that study attrition was particularly associated with misclassification that occurred when the middle assessment was omitted. As with most prospective research, attrition increased over the course of the study, and dropping an assessment at the midpoint in conjunction with missing data late in the study seems to make classification difficult. It is crucial to regain contact with a respondent who has been lost relatively early. We note here that attrition in the present study was very low (with 90% of the targeted respondents reinterviewed at Year 11, and < 8% attrition at any wave except the last). The findings presented here, then, likely provide a best-case scenario—samples with higher attrition would have even greater misclassification. Retaining a prospective sample is difficult, but clearly failure to do so can have critical implications.

In addition, degree of misclassification seems also to be a function of the extent to which the data fit the model well (e.g., an individual whose pattern of responses reflected consistent high drinking as opposed to an individual whose drinking fluctuated between moderate and high). These fluctuations may be a function of time-varying covariates that create short-term perturbations in drinking, perhaps due to illness or relocation that affords a greater or lesser opportunity to drink (e.g., due to military service). These inconsistent responses may also be due to measurement error. This underscores the importance of using highly reliable and valid measures and of retaining procedural consistency across the collection of panel data. Indeed, to parallel the misclassification among the developmentally limited class, we noted that those who were classified in a decreasing class exhibited lower probability of class membership. It is possible that we did not have the sample size to detect additional developmentally limited classes; in fact, in the full-sample five-class AUD model, the chronic group was split into a highly chronic group and a second developmentally limited group (data not shown), although

model fit was slightly better in the four-class model. This raises the possibility that the maturing out process does not, in fact, reflect a single trajectory but is perhaps a manifestation of multiple routes marking the decrease from high alcohol involvement.

The current study has implications for the timing of assessments. For example, regardless of the timing of the eliminated assessment, misclassification between the nondrinking (nondiagnosing) class and the drinking (diagnosing) classes was more evident than misclassification among drinking classes. This may simply be an issue of base rate, although it is somewhat surprising, because one might expect greater distinction between nondrinkers and nonabstainers than among different types of drinkers. Our findings also suggested that, when charting the developmental course of drinking over young adulthood, it is particularly important to include assessments that cover the mid-young adult years, which is a period during which many young adults are altering their drinking behavior (perhaps due to changes in career or social role status). For college student samples such as these, it may be less important to assess drinking during the college years frequently, as heavy drinking and AUDs are rather prevalent during these years due to the situational influence of being in college.

The interval between measurement occasions can have critical importance in a range of types of prospective data analyses. For example, using a portion of the current sample (Years 1 through 4), Sher and Wood (1997) observed prospective relations in the opposite direction when looking at alcohol use (a latent factor comprised of indices of quantity–frequency, drunkenness, and heavy drinking) and escape reasons for drinking (RFD) as a function of interval length. Specifically, alcohol use predicted subsequent RFD when the interval was one year in length (Years 1 to 2), but when the interval was three years (Years 1 to 4), RFD predicted subsequent alcohol use. These findings clearly show that timing of assessments is critical in resolving the directionality between drinking and risk factors such as reasons for drinking. In addition, whether interval length should be consistent for studies with multiple follow-ups is unclear. Although there may be a benefit to equal intervals for many applications, this can be design inefficient, particularly if the researcher is interested in carefully resolving certain developmental periods more than others and when the rate of change in observed behaviors is known to vary at different stages of development. We were not able to manipulate (independently) the construct of interval length in the current study; however, researchers might be able to use Monte Carlo studies (or more densely sampled cohorts assessed over extended time periods) to address the relative influence of the regularity of interval spacing, the temporal intervals themselves, and the number of assessment occasions.

From our perspective, our findings reveal both strengths and weaknesses of growth mixture-derived trajectories. On the positive side, we find that loss of a single assessment occasion (when six occasions are used in the base model) results in a relatively small loss of overall classification accuracy regardless of whether the assessment occurred at the beginning, middle, or end of our observation interval, although clearly, from our analysis of heavy drinking frequency, the final timepoint appeared influential in characterizing the composition of a later-onset group. However, although general classification was not severely impaired, there appeared to be changes in the shapes of some observed trajectories and in their prevalences. Perhaps more important, there was nontrivial instability in the membership of some classes, a finding that is somewhat counterintuitive when both shape and trajectory prevalences appear to be conserved. It is clear that both investigators conducting research with these techniques and consumers of this research need to be careful not to overreify the statistical abstractions we refer to as classes and to remain mindful that the trajectories observed are conditional upon the specific measure of a construct assessed (Jackson & Sher, 2005) as well as other design features such as number of assessments, timing of assessments, or length of the total observation period and factors such as inclusion of covariates in the model. In addition, because we did not validate the trajectories in each model with regard to risk factors known to be

associated with drinking course (e.g., gender, family history of alcoholism, alcohol outcome expectancies), we do not know the extent to which misclassification of class membership has implications for differentiating among trajectories that are distinctive on the basis of such risk factors. Another potentially important class of factors shown to be associated with the derivation of growth mixture-derived trajectories includes the statistical techniques used for identifying mixtures, a topic well beyond the focus of this paper (but see Muthén, 2001a; Muthén & Shedden, 1999; and Nagin, 1999, for useful discussions on this issue), as well as the distributional properties of the measures (Bauer & Curran, 2003). Admittedly, the results presented here are derived from a single study with its own design characteristics and sample using only two alternative measures of a broad alcohol involvement construct. However, even as a case study of the implications of these findings, we provide reason to be cautious in overgeneralizing and overreifying distinctions beyond the specific design features of the study under consideration.

Summary

Substance use disorders are often conceptualized as a chronic relapsing disease (Platt, 1995), and problems with greater severity typically are more chronic or progressive than those that are developmentally limited in nature (e.g., Schulenberg, Wadsworth, O'Malley, Bachman, & Johnston, 1996). Increasingly, clinicians are basing diagnoses on course as much as on symptomatology, consistent with the Krapelinian approach that emphasized syndrome description in part based on longitudinal course (Widiger & Clark, 2000). Unfortunately, formal diagnostic nosology has not kept up with either theory or data that highlight the importance of considering longitudinal course. The current study underscores the extent to which we might expect misclassification of an individual whose course of drinking is chronic and provides a discussion of the parameters that might influence such misclassification.

Acknowledgments

Preparation of this article was supported by National Institute on Alcohol Abuse and Alcoholism Grants K01 AA13938, R01 AA07231, R37 AA13987, and P50 AA11998. We wish to thank Bengt Muthén, Linda Muthén, Aesoon Park, and Phillip Wood for their assistance in analyses and helpful comments on the work presented herein.

References

- Akaike H. Factor analysis and AIC. *Psychometrika* 1987;52:317–332.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 3. Washington, DC: Author; 1980.
- Bauer DJ, Curran PJ. Distributional assumptions of growth mixture models: Implications for over-extraction of latent trajectory classes. *Psychological Methods* 2003;8:338–363. [PubMed: 14596495]
- Caskie, GIL. Latent growth curve model performance under varying incomplete data conditions. Poster session presented at the annual meeting of the American Psychological Association; Boston, MA. 1999 Aug.
- Chassin L, Flora DB, King KM. Trajectories of Alcohol and Drug Use and Dependence from Adolescence to Adulthood: The Effects of Familial Alcoholism and Personality. *Journal of Abnormal Psychology* 2004;113:483–498. [PubMed: 15535782]
- Chassin L, Pitts SC, Prost J. Binge drinking trajectories from adolescence to emerging adulthood in a high-risk sample: Predictors and substance abuse outcomes. *Journal of Consulting and Clinical Psychology* 2002;70:67–78. [PubMed: 11860058]
- Colder CR, Campbell RT, Ruel E, Richardson JL, Flay BR. A finite mixture model of growth trajectories of adolescent alcohol use: Predictors and consequences. *Journal of Consulting and Clinical Psychology* 2002;70:976–985. [PubMed: 12182281]
- Crews TM, Sher KJ. Using adapted Short MAST's for assessing parental alcoholism: Reliability and validity. *Alcoholism: Clinical and Experimental Research* 1992;16:576–584.

- Dawson DA, Grant BF, Stinson FS, Chou PS. Another look at heavy episodic drinking and alcohol use disorders among college and noncollege youth. *Journal of Studies on Alcohol* 2004;65:477–488. [PubMed: 15378804]
- Endicott, J.; Andreasen, N.; Spitzer, RL. *Family History–Research Diagnostic Criteria (FH-RDC)*. New York: New York State Psychiatric Institute; 1978.
- Faden VB, Day NL, Windle M, Windle R, Grube JW, Molina BSG, et al. Collecting longitudinal data through childhood, adolescence, and young adulthood: Methodological challenges. *Alcoholism: Clinical and Experimental Research* 2004;28:330–340.
- Hill KG, White HR, Chung I, Hawkins JD, Catalano RF. Early adult outcomes of adolescent binge drinking: Person- and variable-centered analyses of binge drinking trajectories. *Alcoholism: Clinical and Experimental Research* 2000;24:892–901.
- Jackson KM, Sher KJ. Similarities and differences of longitudinal phenotypes across alternate indices of alcohol involvement: A methodologic comparison of trajectory approaches. *Psychology of Addictive Behaviors* 2005;19:339–351. [PubMed: 16366806]
- Johnston, LD.; O'Malley, PM.; Bachman, JG. *National survey results on drug use from the Monitoring the Future study, 1975–2001: Vol. I. Secondary school students. Vol. II. College students and young adults (NIH Publication No. 02–5106 & 02–5107)*. Bethesda, MD: National Institute on Drug Abuse; 2002.
- Jones BL, Nagin DS, Roeder KA. SAS procedure based on mixture models for estimating developmental trajectories. *Sociological Methodology Research* 2001;29:374–393.
- Knight JR, Wechsler H, Kuo M, Seibring M, Weitzman ER, Schuckit MA. Alcohol abuse and dependence among U.S. college students. *Journal of Studies on Alcohol* 2002;63:263–270. [PubMed: 12086126]
- Li F, Barrera MJR, Hops H, Fisher KJ. The longitudinal influence of peers on the development of alcohol use in late adolescence: A growth mixture analysis. *Journal of Behavioral Medicine* 2002;25:293–315. [PubMed: 12055779]
- Muthén, B. Latent variable mixture modeling. In: Marcoulides, GA.; Schumacker, RE., editors. *New developments and techniques in structural equation modeling*. Mahwah, NJ: Erlbaum; 2001a. p. 1–33.
- Muthén, B. Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class-latent growth modeling. In: Collins, LM.; Sayer, AG., editors. *New methods for the analysis of change. Decade of behavior*. Washington, DC: American Psychological Association; 2001b. p. 291–322.
- Muthén B. Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran (2003). *Psychological Methods* 2003;8:369–377. [PubMed: 14596497]
- Muthén B, Muthén LK. Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research* 2000;24:882–891.
- Muthén B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* 1999;55:463–469. [PubMed: 11318201]
- Muthén, LK.; Muthén, BO. *2004 Mplus user's guide. 3*. Los Angeles: Authors; 1998.
- Nagin DS. Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods* 1999;4:139–157.
- Nunnally, JC.; Bernstein, IH. *Psychometric theory*. St. Louis: McGraw-Hill, Inc; 1994.
- Platt, JJ. *Heroin addiction: Theory, research, and treatment: Vol. 2. The addict, the treatment process, and social control*. Melbourne, FL: Krieger; 1995.
- Raudenbush SW, Chan WS. Application of a hierarchical linear model to the study of adolescent deviance in an overlapping cohort design. *Journal of Consulting and Clinical Psychology* 1993;61:941–951. [PubMed: 8113495]
- Robins LN, Cottler L, Bucholz KK, Compton W. *Diagnostic Interview Schedule for DSM–IV (DIS-IV)*. Unpublished manuscript. 1997
- Robins, LN.; Helzer, JE.; Cottler, L.; Goldring, E. *National Institute of Mental Health Diagnostic Interview Schedule, Version III (Rev. ed.)*. St. Louis, MO: Washington University; 1989.
- Robins, LN.; Helzer, JE.; Croughan, J.; Williams, JBW.; Spitzer, RL. *PHS Publication No. ADDM-T-42–3*. Washington, DC: Institute of Mental Health; 1985. *National Institute of Mental Health Diagnostic Interview Schedule, Version III-A*; p. 4–85.

- Schulenberg J, Wadsworth KN, O'Malley PM, Bachman JG, Johnston LD. Adolescent risk factors for binge drinking during the transition to young adulthood: Variable- and pattern-centered approaches to change. *Developmental Psychology* 1996;32:659–674.
- Schwartz G. Estimating the dimension of a model. *Annals of Statistics* 1978;6:461–464.
- Selzer M, Vinokur A, van Rooijen L. A self-administered Short Michigan Alcoholism Screening Test (SMAST). *Journal of Studies on Alcohol* 1975;36:117–126. [PubMed: 238068]
- Sher KJ, Walitzer KS, Wood PK, Brent EE. Characteristics of children of alcoholics: Putative risk factors, substance use and abuse, and psychopathology. *Journal of Abnormal Psychology* 1991;100:427–448. [PubMed: 1757657]
- Sher, KJ.; Wood, PK. Methodological issues in conducting prospective research on alcohol-related behavior: A report from the field. In: Bryant, KJ.; West, SG.; Windle, M., editors. *The science of prevention: Methodological advances from alcohol and substance abuse research*. Washington, DC: American Psychological Association; 1997. p. 3-41.
- Singer, JD.; Willett, JB. *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press; 2003.
- Tucker JS, Orlando M, Ellickson PL. Patterns and correlates of binge drinking trajectories from early adolescence to young adulthood. *Health Psychology* 2003;22:79–87. [PubMed: 12558205]
- Wechsler H, Lee JE, Kuo M, Lee H. College binge drinking in the 1990s: A continuing problem. Results of the Harvard School of Public Health 1999 College Alcohol Study. *Journal of American College Health* 2000;48:199–210. [PubMed: 10778020]
- White HR, Johnson V, Buyske S. Parental modeling and parenting behavior effects on offspring alcohol and cigarette use: A growth curve analysis. *Journal of Substance Abuse* 2000;12:287–310. [PubMed: 11367605]
- Widiger TA, Clark LA. Toward *DSM-V* and the classification of psychopathology. *Psychological Bulletin* 2000;126:946–963. [PubMed: 11107884]
- Willett JB. Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth. *Educational & Psychological Measurement* 1989;49:587–602.

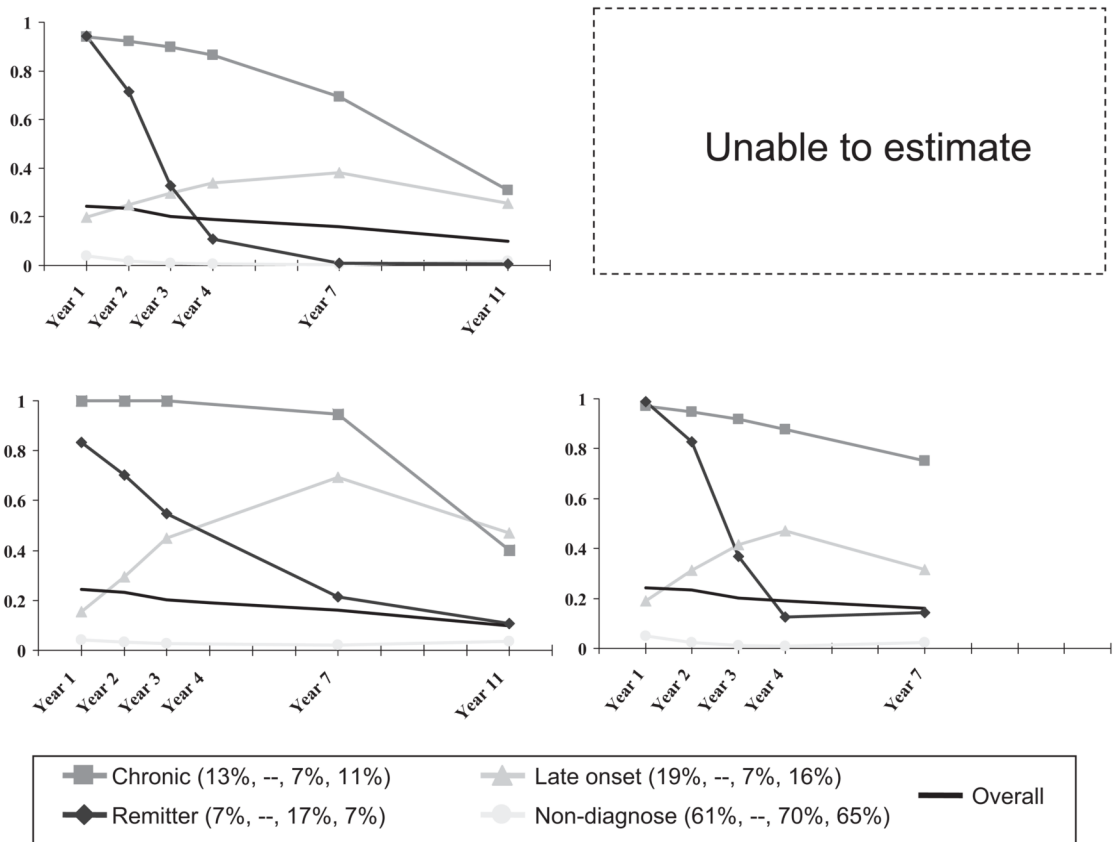


Figure 1. Latent class growth analysis (LCGA) models for alcohol use disorder (AUD) weighted by estimated class probabilities and modeling missing data. Upper left: With all six assessments. Upper right: With the Year 1 assessment omitted. Lower left: With the Year 4 assessment omitted. Lower right: With the Year 11 assessment omitted. Class prevalences in the figure correspond to all six assessments, Year 1 omitted, Year 4 omitted, and Year 11 omitted, respectively. $N = 489$.

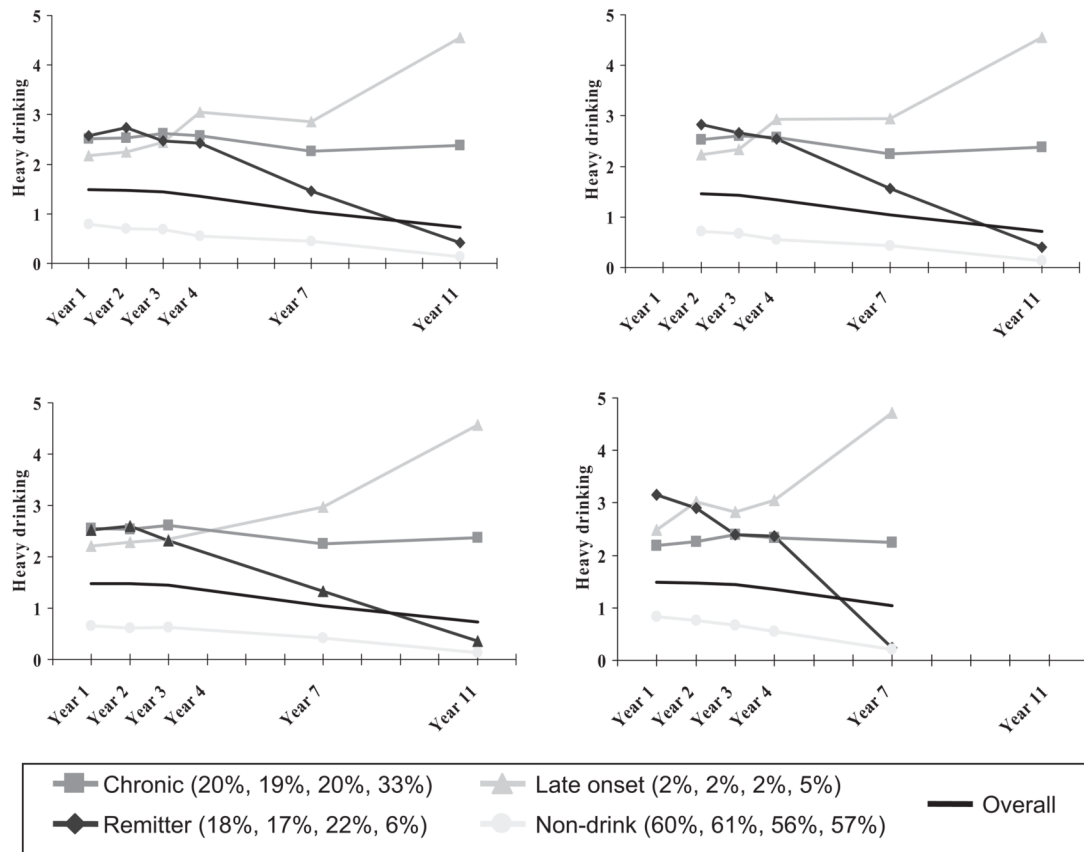


Figure 2. Mixture models for heavy drinking weighted by estimated class probabilities and modeling missing data. Upper left: With all six assessments. Upper right: With the Year 1 assessment omitted. Lower left: With the Year 4 assessment omitted. Lower right: With the Year 11 assessment omitted. Class prevalences in the figure correspond to all six assessments, Year 1 omitted, Year 4 omitted, and Year 11 omitted, respectively. $N = 489$.

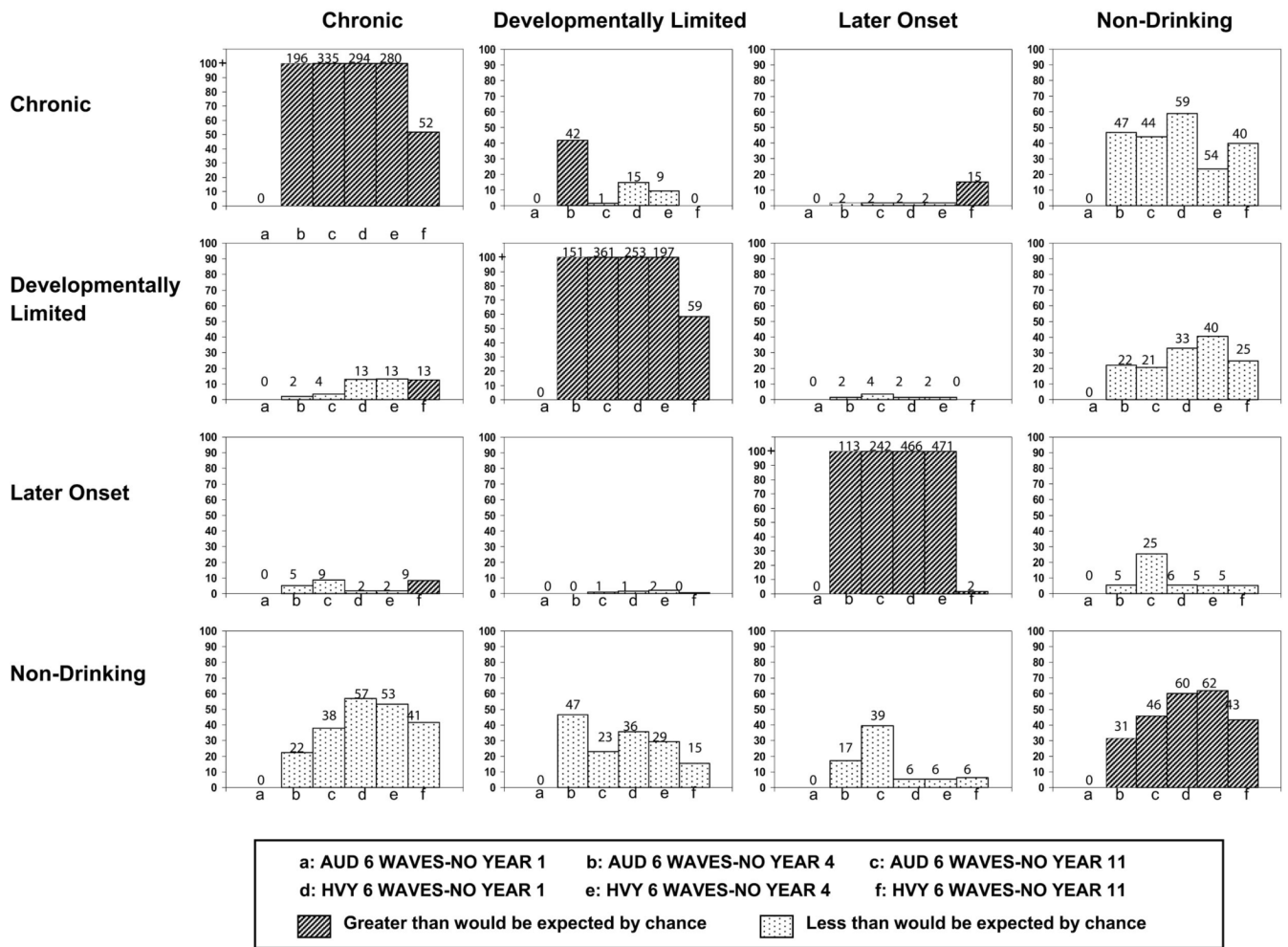


Figure 3. Summary of trajectory agreement represented by cell chi-square statistics for comparison between six-wave and five-wave models, for alcohol use disorder (AUD) and heavy drinking. The four rows refer to the six-wave models and the four columns refer to the (varying) five-wave models (depending on their placement, according to the legend; e.g., the fourth bar column, labeled D, corresponds the six-wave model vs. the five-wave model with Year 1 omitted, for heavy drinking). Dark hatched bars reflect values that are greater than would be expected by chance and light polka-dot bars indicate values that are lower than would be expected by chance, based on the marginals. Chi-square values for each comparison are above each bar. Note that a chi-square value with a single degree of freedom of 3.84 is significant at $p < .05$ and a chi-square value with a single degree of freedom of 6.64 is significant at $p < .01$.

Table 1

Model Fit for Five Indices of Alcohol Involvement

	Alcohol Use Disorder (AUD)			
	2-class	3-class	4-class	5-class
Full assessments (Years 1, 2, 3, 4, 7, 11)	AIC = 2103.93	AIC = 2067.60	AIC = 2051.76	AIC = 2054.06
	BIC = 2133.28	BIC = 2113.72	BIC = 2114.64	BIC = 2133.72
	Entropy =.86	Entropy =.75	Entropy =.77	Entropy =.76
Missing Year 1	AIC = 1725.74			
	BIC = 1755.02	Would not converge	Would not converge	Would not converge
	Entropy =.84			
Missing Year 4	AIC = 1780.29	AIC = 1773.81	AIC = 1765.03	
	BIC = 1809.64	BIC = 1819.93	BIC = 1827.92	Would not converge
	Entropy =.86	Entropy =.85	Entropy =.81	
Missing Year 11	AIC = 1865.02	AIC = 1838.56	AIC = 1827.68	
	BIC = 1809.64	BIC = 1884.67	BIC = 1890.57	Would not converge
	Entropy =.86	Entropy =.76	Entropy =.79	
Heavy drinking				
	2-class	3-class	4-class	5-class
Full assessments (Years 1, 2, 3, 4, 7, 11)	AIC = 8107.10	AIC = 8068.86	AIC = 7975.66	
	BIC = 8165.79	BIC = 8140.14	BIC = 8059.51	Would not converge
	Entropy =.86	Entropy =.77	Entropy =.82	
Missing Year 1	AIC = 6670.90	AIC = 6626.04	AIC = 6530.56	AIC = 6530.66
	BIC = 6725.27	BIC = 6692.95	BIC = 6610.02	BIC = 6622.67
	Entropy =.87	Entropy =.79	Entropy =.83	Entropy =.86
Missing Year 4	AIC = 6894.42	AIC = 6863.09	AIC = 6769.22	
	BIC = 6948.92	BIC = 6930.17	BIC = 6848.88	Would not converge
	Entropy =.86	Entropy =.75	Entropy =.80	
Missing Year 11	AIC = 7160.80	AIC = 7132.74	AIC = 7034.02	
	BIC = 7215.30	BIC = 7199.81	BIC = 7113.67	Would not converge
	Entropy =.77	Entropy =.82	Entropy =.87	

Note. BIC: Bayesian information criterion; AIC: Akaike's information criterion. $N = 489$ ($N = 484$ for the heavy drinking model without Year 1).

Table 2

Cross-Tabulation of Frequency, Cell Percentage, and Cell Chi-Square for Group Membership in Classes for the Full Model Crossed With the Model That Has Year 1 Eliminated

Full six-wave model	Model with Year 1 eliminated			
	Chronic	Develop. ltd.	Late onset	Nondrinker
1. Chronic	95 (19.6%) <i>293.9</i>	0 (0%) <i>14.9</i>	0 (0%) <i>1.8</i>	0 (0%) <i>58.9</i>
2. Developmentally limited	2 (0.4%) <i>13.0</i>	70 (14.5%) <i>253.4</i>	0 (0%) <i>1.5</i>	10 (2.1%) <i>32.8</i>
3. Late onset	0 (0%) <i>1.8</i>	0 (0%) <i>1.4</i>	9 (1.9%) <i>466.2</i>	0 (0%) <i>5.6</i>
4. Nondrinker	2 (0.4%) <i>57.0</i>	6 (1.2%) <i>35.6</i>	0 (0%) <i>5.5</i>	290 (59.9%) <i>60.0</i>
Marginals	99 (20.4%)	76 (15.7%)	9 (1.9%)	300 (62.0%)
				484

Note. $\chi^2(9, N = 484) = 1,303.34, p < .001$; $\phi = 1.64$; Cramér's $V = .95$; Cohen's $\kappa = .92$. Cell percentages are indicated by parentheses; cell chi-squares are indicated by italicized font.