Development and Psychopathology **20** (2008), 291–318
Copyright © 2008 Cambridge University Press
Printed in the United States of America
DOI: 10.1017/S095457940800014X

A model-based cluster analysis approach to adolescent problem behaviors and young adult outcomes

EUN YOUNG MUN,[a] MICHAEL WINDLE,[b] AND LISA M. SCHAINKER[c]
[a]*Rutgers University;* [b]*Emory University; and* [c]*Iowa State University*

Abstract

Data from a community-based sample of 1,126 10th- and 11th-grade adolescents were analyzed using a model-based cluster analysis approach to empirically identify heterogeneous adolescent subpopulations from the person-oriented and pattern-oriented perspectives. The model-based cluster analysis is a new clustering procedure to investigate population heterogeneity utilizing finite mixture multivariate normal densities and accordingly to classify subpopulations using more rigorous statistical procedures for the comparison of alternative models. Four cluster groups were identified and labeled multiproblem high-risk, smoking high-risk, normative, and low-risk groups. The multiproblem high risk exhibited a constellation of high levels of problem behaviors, including delinquent and sexual behaviors, multiple illicit substance use, and depressive symptoms at age 16. They had risky temperamental attributes and lower academic functioning and educational expectations at age 15.5 and, subsequently, at age 24 completed fewer years of education, and reported lower levels of physical health and higher levels of continued involvement in substance use and abuse. The smoking high-risk group was also found to be at risk for poorer functioning in young adulthood, compared to the low-risk group. The normative and the low risk groups were, by and large, similar in their adolescent and young adult functioning. The continuity and comorbidity path from middle adolescence to young adulthood may be aided and abetted by chronic as well as episodic substance use by adolescents.

Adolescence is a developmental period characterized by increases in risk behaviors, mood fluctuations, and conflict with parents (Arnett, 1992, 1999), as well as when significant brain development occurs in the frontal lobes, influencing the development of better reasoning and decision-making capabilities (Dahl, 2004; Reyna & Farley, 2006). Recent national data from the 2005 Youth Risk Behavior Survey (Eaton et al., 2006) highlight that a substantial percentage of adolescents in 9th through 12th grades in the United States are involved in a number of risk behaviors, including current cigarette smoking (23.0%), alcohol use (43.3%), heavy drinking (25.5%), marijuana use (20.2%), aggression and violence (e.g., physical fighting; 35.9%), and extreme feelings of sadness (28.5%) and suicidal ideas (16.9%). The alarmingly high rates of substance use among adolescents and college students are echoed in the most recent report from the Monitoring the Future National Survey (Johnston, O'Malley, Bachman, & Schulenberg, 2006). For most adolescents, engagement in some level of risk behavior is part of a statistically normative process that is intertwined with age-appropriate developmental tasks associated with increases in autonomy and self-regulation.

However, researchers agree that there exist subsets of adolescents at elevated risk whose

This research was supported by the National Institute on Alcohol Abuse and Alcoholism Grant R37-AA07861 awarded to Michael Windle. We are grateful to Alexander von Eye for his helpful comments on an earlier version of this manuscript.

Address correspondence and reprint requests to: Eun Young Mun, Rutgers Center of Alcohol Studies, Rutgers University, 607 Allison Road, Piscataway, NJ 08854; E-mail: eymun@rci.rutgers.edu.

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

adolescent-typical behaviors may represent atypical developmental processes, and it remains a critical task to detect atypical symptomatic processes from the normative pathway in the field of developmental psychopathology (Cicchetti & Rogosch, 2002). Evidence of heterogeneity in developmental pathways consisting of normative as well as atypical processes can be found in depressive disorders and symptoms (Cicchetti & Toth, 1998; Kim, Capaldi, & Stoolmiller, 2003; Stoolmiller, Kim, & Capaldi, 2005), heavy drinking (Schulenberg, O'Malley, Bachman, Wadsworth, & Johnston, 1996; Windle, Mun, & Windle, 2005), smoking (Chassin, Presson, Pitt, & Sherman, 2000; Chassin, Presson, Sherman, & Edwards, 1991; Shiffman, Kassel, Paty, Gnys, & Zettler-Segal, 1994), alcohol and drug use (Chassin, Flora, & King, 2004), antisocial behavior (Moffitt, 1993; Nagin, Farrington, & Moffitt, 1995), and aggression (Bongers, Koot, van der Ende, & Verhulst, 2004; Brame, Nagin, & Tremblay, 2001; Loeber & Stouthamer-Loeber, 1998). From this corpus of research, variations in onset and desistance, as well as overall levels of the behaviors, have been identified as critical factors for identifying individuals at risk. In these studies, researchers found that the adolescents who are flagged for early onset tend to also show other elevated behavior problems and concurrent clinical symptoms. Thus, it is critical to identify adolescents who are at increased risk for developmental pathways marked by a prolonged period of multiple dysfunctions from other less risky or singular-problem pathways, and to understand how vulnerability develops and is expressed across time.

Cluster Analysis for Identifying Adolescents at Risk

A number of other studies have identified clusters of adolescents who can be distinguished from others based on behavioral domains and/or substance use (Crockett, Moilanen, Raffaelli, & Randall, 2006; Gorman-Smith, Tolan, Loeber, & Henry, 1998; Haselager, Cillessen, Van Lieshout, Riksen-Walraven, & Hartup, 2002; Miller & Plant, 2002; Potter & Jenson, 2003; Tubman, Vicary, von Eye, & Lerner, 1990). Although these studies have provided valuable insight into heterogeneous clusters of adolescents who vary on problem behaviors, major methodological limitations are twofold. First, the cluster analyses conducted in these studies used either a limited number of problem behaviors (Miller & Plant, 2002), a small sample size (Tubman et al., 1990), and/or consisted of a highly selective sample such as juvenile offenders (Potter & Jenson, 2003). Second, and more critically, the cluster analytic techniques used in these studies are considered to have substantial weaknesses. In particular, these clustering algorithms' lack of a statistical fit measure to determine the adequacy of the number of clusters, yields *indeterminacy of a cluster solution* (Dumenci & Windle, 2001; Tonidandel & Overall, 2004), and the clustering solution may be influenced by the *order of data input* (van der Kloot, Spaans, & Heiser, 2005). These issues are compounded, given that each clustering method has unique tendency to favor a certain solution. For instance, Ward's method, one of the most commonly used clustering methods, tends to result in spherical clusters of the same size (for more detail, see Everitt, Landau, & Leese, 2001; von Eye, Mun, & Indurkhya, 2004). Furthermore, spurious clusters can be artificially created by a clustering method that imposes a structure on the data even when there are no true clusters (Everitt et al., 2001; von Eye & Bergman, 2003; von Eye et al., 2004). Although a series of careful decisions (e.g., hierarchical vs. nonhierarchical clustering procedures, deterministic vs. stochastic clustering, choosing a linkage method, and deciding on including or excluding outliers) can limit the arbitrariness involved in heuristic clustering procedures (see von Eye et al., 2004), findings based on subjective classification rules and means in finding cluster patterns should be viewed cautiously (Hand & Bolton, 2004; Nagin, 2005).

Mixture Model-Based Cluster Analysis Using Finite Mixture Distributions

Newly emerging model-based cluster analysis[1] utilizes finite mixture densities (Everitt, 2005;

1. Model-based cluster analysis and mixture model-based cluster analysis are interchangeably used, and both terms indicate cluster analysis via a finite mixture model approach.

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

Everitt & Hand, 1981; McLachlan & Peel, 2000). The foci of clustering become that of estimating (a) the parameters of the assumed mixture (e.g., multivariate normal, multivariate *t*, multivariate multinomial, etc.) and (b) the posterior probabilities of cluster membership (Banfield & Raftery, 1993; Everitt et al., 2001). This analytic approach targets the identification of unobserved heterogeneity in a population based on the observed data. Model-based cluster analysis assumes that the distribution is made up of a number of multivariate component distributions (typically normal densities), and it utilizes the expectation–maximization (EM) algorithm for maximum likelihood (ML) estimation. The ultimate goal of model-based cluster analysis is similar to any of the heuristic clustering algorithms such as hierarchical agglomerative or the *k*-means clustering method. It aims to find a number of salient groups of objects (e.g., individuals, genes, etc.) that are similar to one another within a group but sufficiently different from members of other groups. However, mixture model-based cluster analysis is different from the existing heuristic clustering procedures in that it employs the explicit assumption that a population can be approximated as a weighted sum of component distributions a priori unknown and unobserved, which can be interpreted as evidence of distinctive subgroups or heterogeneity in a population. In addition, the cluster membership of an individual is expressed as a probability (range = 0–1), as opposed to hard, deterministic classification that typically assumes that cluster membership of an individual equals 1.

This emerging approach to cluster analysis via a finite mixture model has been better known and utilized in the field of gene expression data (McLachlan, Do, & Ambroise, 2004). To our knowledge, however, this model-based cluster analysis approach has not been introduced in the field of developmental psychopathology. Latent class analysis, latent class growth curve analysis, and latent class regression using the *Mplus* program (Muthén & Muthén, 1998–2006) can be considered as types of model-based cluster analysis as well (Everitt et al., 2001). From a different angle, mixture model-based cluster analysis can be viewed as a type of hybrid latent variable analysis with no measurement invariance and nonparametric factor distribution (i.e., nonparametric factor mixture analysis; Muthén, 2006). If a component in finite mixture densities is considered as a latent class (e.g., Dolan, Schmittmann, Lubke, & Neale, 2005; see Bollen, 2002, for in-depth discussion on various definitions of latent variables), the mixture model-based cluster analysis may also be understood as *latent* cluster analysis.

As expected, a mixture model-based cluster analysis has commonalities with recent methodological advances that utilize finite mixture distributions applied to structural equation models (Arminger & Stein, 1997; Arminger, Stein, & Wittenberg, 1999; Dolan & van der Maas, 1998; Jedidi, Jagpal, & DeSarbo, 1997; Lubke & Muthén, 2005) and longitudinal models (Muthén & Shedden, 1999; Nagin, 1999, 2005). Especially the growth mixture modeling and latent class growth curve analysis techniques have sparked an avalanche of empirical studies that have identified and compared groups of individuals who follow particular developmental courses and compared them with other patterns of change groups in terms of the antecedents, correlates, and consequences of their developmental trajectories. These studies have shown evidence of heterogeneous developmental trajectories for smoking behavior (Chassin et al., 2000; Colder, Mehta, et al., 2001; Orlando, Tucker, Ellickson, & Klein, 2004; Tucker, Ellickson, & Klein, 2002; White, Pandina, & Chen, 2002), alcohol use (Chassin, Pitts, & Prost, 2002; Colder, Campbell, Ruel, Richardson, & Flay, 2002; Hill, White, Chung, Hawkins, & Catalano, 2000; Tucker, Orlando, & Ellickson, 2003; Windle et al., 2005), marijuana use (Flory, Lynam, Milich, Leukefeld, & Clayton, 2004; Guo et al., 2002; Windle & Wiesner, 2004), depressive symptoms (Stoolmiller et al., 2005; Windle & Mun, 2006), and delinquent behavior (Nagin & Tremblay, 1999; Wiesner & Capaldi, 2003; Wiesner, Kim, & Capaldi, 2005).

Although these studies support heterogeneous subgroups in each of the problem behaviors previously mentioned, the rapidly growing literature has become difficult to synthesize because of its wide variation in measures (e.g., frequency versus quantity of substance use),

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

sample characteristics (e.g., community-based normative samples versus high-risk samples), and analytic models (e.g., conditional vs. unconditional models). Moreover, past studies utilizing latent growth mixture analysis have typically considered a single problem behavior (e.g., delinquency) per study when deriving groups, thereby overlooking the co-occurrence of multiple problem behaviors that is common in adolescence, and consequently limiting the understanding of general and specific causal mechanisms responsible for heterogeneity of pathways and comorbidity. Although a few recent studies have derived groups based on joint or parallel trajectory models (Chassin et al., 2004; Jester et al., 2005; Nagin & Tremblay, 1999), the scope of adolescent problem behaviors under investigation has been very limited in both number and kind. Furthermore, few existing studies of joint trajectories have investigated two closely related behaviors (Hix-Small, Duncan, Duncan, & Okut, 2004; Jester et al., 2005).

The current study intends to fulfill two inseparable goals. We aim to introduce a mixture model-based cluster analysis approach and to derive meaningful clusters of adolescents with different risk profiles from the person-oriented and pattern-oriented approaches (Bergman & Magnusson, 1997; Magnusson, 2000; von Eye & Bergman, 2003). One typical assumption involved in the variable-oriented research is that all relevant contributing factors to an outcome be included so that each causal factor's contribution can be understood when effects of all other variables are controlled (i.e., at their means). However, an average score is often atypical and uninformative for research on developmental psychopathology, and can be misleading in some situations (von Eye & Bergman, 2003). Furthermore, the entire model can be misleading if an important variable is excluded because the exclusion can potentially modify the entire network of relationships among variables. Another related assumption is that variable-oriented research attempts to identify common ("main effect") causal relations that apply to all individuals (unless an a priori model is hypothesized to account for the specificity of targeted relations for different groups of individuals, i.e., interaction terms). Thus, our understanding of individual vulnerability and resilience with regard to adolescent problem behaviors and their consequences will benefit from the person-oriented research with a within-individual, person-oriented focus, especially because adolescent problem behaviors are well known to co-occur with high frequency and with a higher frequency among some adolescents relative to others. In the present study, we sought to empirically identify clusters of adolescents who differ in their profiles of multiple problem behaviors using a new nonparametric method called model-based cluster analysis and to understand precursors and correlates from middle adolescence to young adulthood.

The Current Study

We focused on identifying distinctive groups of adolescents with different kinds and/or levels of risk, and comparing the empirically derived cluster groups on adolescent temperamental and academic risk factors and young adult outcomes. Adolescent temperamental characteristics (e.g., Chassin et al., 2004; Tubman & Windle, 1995; Windle, 1989, 1991) and academic functioning (e.g., Jessor, van den Bos, Vanderryn, Costa, & Turbin, 1995) have been extensively studied as correlates of adolescent problem behaviors. Furthermore, temperamental deviations, in particular, have been strongly suggested (Moffitt, 1993; Tarter & Vanyukov, 1994; Zucker, Ellis, Bingham, & Fitzgerald, 1996; Zucker, Fitzgerald, & Moses, 1995) as an early indicator of neuropsychological risk that may be manifested as deficits in behavioral and emotional regulation and academic difficulties. More broadly, individual differences in temperament may signal different levels of behavioral and emotional self-regulation and reactivity in expressing one's self in socially acceptable ways (Posner & Rothbart, 2000). A good number of publications have extended our understanding of early temperament in association with adult functioning and psychopathology (e.g., the Dunedin Multidisciplinary Health and Development Study), but relatively little attention has been given to more temporally proximal associations between adolescent measures and young adult functioning and

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

psychopathology (Cicchetti & Rogosch, 2002; Schulenberg, Sameroff, & Cicchetti, 2004). Finally, we examined whether derived clusters of adolescents were linked to young adult outcomes such as young adult substance use and abuse, depressive symptoms, overall perceived health, and education completed. High-risk groups of adolescents were hypothesized to have higher risk young adult profiles for the outcome measures in this study.

To cluster adolescents, we used the MCLUST program developed by Fraley and Raftery (1998, 1999, 2002a, 2002b, 2003) and designed for S-PLUS software program (version 6 or higher; Insightful Corporation, 1988–2006) and R language (available gratis at http://www.r-project.org/). This model-based clustering approach unambiguously postulates that a population consists of heterogeneous subpopulations that can differ in multidimensional data space calibrated in terms of volume (size of a cluster), orientation (an ellipsoid's orientation angles), and shape (a sphere or an ellipsoid). Appendix A provides graphic illustrations of volume, shape, and orientation in two-dimensional space. These geometric properties—volume, shape, and orientation—of a covariance matrix among observed variables are parameterized (see Appendix B for more information). A total of 10 possible models can be simultaneously compared using the Bayesian information criterion (BIC) as a statistical fit measure for the same number of extracted subcomponents (e.g., see Figure 1). The simplest model (Model A = same volume and shape, spherical clusters) has been shown to be very closely related to the *k*-means clustering algorithm (Celeux & Govaert, 1992; Everitt et al., 2001; Yeung, Fraley, Murua, Raftery, & Ruzzo, 2001). Because all models are compared at the same time using the BIC, a straightforward, statistical measure is available to draw a conclusion about the optimal number of clusters (e.g., component distributions) and the best clustering procedure reflected by the spatial characteristics of data. This is an important advantage of the MCLUST program because there generally exists a trade-off between the number of clusters and the number of estimated parameters. The more simplistic models may require a larger number of clusters to adequately fit the data, whereas a small number of clusters may be sufficient with increasingly more complex models to provide adequate fit to the data (Yeung et al., 2001; for an empirical demonstration applied to, for instance, latent class growth trajectory analysis, see Bauer & Curran, 2003, 2004). After identifying cluster groups, we analyzed adolescent temperamental and academic profiles by cluster groups, and young adult function profiles by cluster groups.

Methods

Participants

The data used in this study were collected as part of a larger, multiwave panel design focused on vulnerability factors and adolescent and adult substance use (Lives across Time; Windle & Wiesner, 2004). Of the maximum original sample of 1,205 adolescent participants (591 males, 614 females) who completed at least one of a maximum of four waves of adolescent assessment, we analyzed 1,126 participants (546 males, 580 females) who provided any information on the seven problem behaviors measured at Wave 2. The Wave 1 data collection during adolescence occurred during the fall semester of the teens' 10th- and 11th-grade years. Waves 2–4 data collection occurred every 6 months thereafter, with Wave 2 data being collected in the spring semester of Year 1, and Waves 3 and 4 being collected in the fall and spring semesters of Year 2, respectively, when adolescents were in 11th and 12th grades. The sample consisted of high school sophomores (53%) and juniors (47%) recruited from two homogeneous suburban public high school districts (a total of three high schools) in Western New York. Fifty-two percent of the sample was female, the average age of the respondents at the first occasion of measurement was 15.54 years ($SD = 0.66$), and 98% were White. Seventy percent of the sample was Catholic, 18% Protestant, and 12% other. Ninety-six percent of the fathers and 43% of the mothers were employed full-time outside the home (37% of mothers were employed part-time outside the home). Fathers completed an average of 13.79 years of education ($SD = 2.39$) and mothers completed an average of 13.55 years ($SD = 2.01$). The median

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

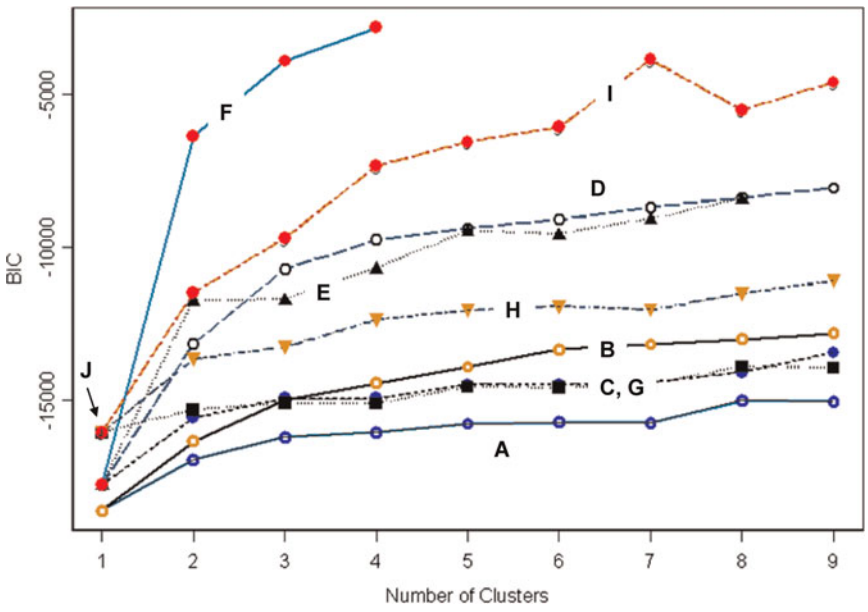


Figure 1. Bayesian information criteria (BIC) of different cluster solutions. A, spherical, equal volume, equal shape; B, spherical, unequal volume, equal shape; C, diagonal, equal volume, equal shape; D, diagonal, unequal volume, equal shape; E, diagonal, equal volume, varying shape; F, diagonal, unequal volume, varying shape; G, ellipsoidal, equal volume, equal shape, equal orientation; H, ellipsoidal, equal volume, equal shape, varying orientation; I, ellipsoidal, unequal volume, equal shape, varying orientation; J, ellipsoidal, unequal volume, varying shape, varying orientation. The BIC for the J model was −16,056.65, the same BIC value as the Model I one cluster solution, and it stopped converging beginning at two clusters. [A color version of this figure can be viewed online at www.journals.cambridge.org]

family income was about $40,000, with only 3% of the sample reporting a family income less than $12,000. Eighty-eight percent of the adolescents' primary caregivers were currently married, 12% were divorced, and 1% were widowed. The retention rate at the second occasion of measurement (6 months later) was 93%. Sample retention across the first four waves of measurement was uniformly high, in excess of 90%. The prevalence of drinking and substance use among adolescents in this sample was highly similar to findings in national survey studies (Windle, 1996).

Subsequent to the four waves of assessment during adolescence, Wave 5 data collection occurred when the average age of the young adults was 24.2 years. The attempt was made to contact all of the original 1,205 participants and an additional 10 young adult children whose primary caregivers participated in the study at Waves 1–4, although their adolescent children (who were surveyed in school settings) were unable to participate because of competing demands (e.g., scheduling conflicts with participation in sports events). At least one (of three possible) participant from 940 households participated in the study, including 829 young adults at Wave 5. Attrition analyses conducted at Wave 5 among participants (n = 829), refusers (n = 153), and nonlocatables (n = 233) revealed that the nonlocatable females had significantly higher levels of delinquency and stressful life events during adolescence than the participating females. However, the nonlocatable females did not differ from the participating females with respect to substance use, deviant peers, depressive symptoms, a grade point average (GPA), or family relationships (for more information, see Windle & Wiesner, 2004).

Data collection

Subsequent to receiving approval from school administrators to conduct the study, schools provided a mailing list of the addresses of

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

10th and 11th graders. A packet of materials, including a letter of introduction by the principal, a description of the study, and informed consent forms, was mailed to adolescents and their parents. Those individuals willing to participate in the study were requested to sign the informed consent form (both adolescent and one parent) and return it to the investigator in a self-addressed, stamped envelope. Confidentiality was also assured with a Department of Health and Human Services Certificate of Confidentiality. Teachers made announcements about the study in home classrooms. Adolescents completed survey materials in large groups (e.g., 40–50 students) in their high school setting. A trained survey research team administered the survey to adolescents, and neither teachers nor school administrators were in the room during the time the students completed the surveys. The survey took about 45–50 min to complete, and subjects received $10 for their participation. A make-up date for testing was arranged for participants who were absent or unable to participate on the regularly scheduled day of testing. A similar procedure was used at each of the first four waves of measurement.

The young adulthood interviews were conducted via one-on-one interviews either in the subjects' home or at the host institute of the investigators. Subjects were paid $40 to complete an interview that lasted approximately 2 hr. The data collection procedures were computerized and administered via computer-assisted personal interviews. Because some of the young adults resided out of the state for a variety of reasons (e.g., military service, college attendance, jobs), the complete interview was administered to 762 young adults who were available for face-to-face interviews. A reduced-protocol telephone interview was administered to the remaining 67 participating young adults.

Measures

Adolescent problem behaviors

Adolescent problem behaviors measured at Wave 2 were utilized in this study because not all of the problem behavior measures were administered at Wave 1. The measurement interval between Wave 1 and Wave 2 was about 6 months within the same school year.

Delinquent and sexual behaviors. Delinquent behaviors were measured using 13 items selected from prior delinquency research (Elliott, Huizinga, & Menard, 1989). Thirteen items measured delinquent behavior encompassing property damage (e.g., *destroyed school or other public property*), aggression (e.g., *beat up someone on purpose*), and theft (e.g., *stole from a store something valued at $20 or less)*. Three items measured sexual behavior (e.g., *engaged in sexual activity not involving actual intercourse, engaged in sexual activity involving actual intercourse*, and *the number of different persons a participant has had intercourse with*). Adolescents were asked to indicate the number of times they were involved in each of the 16 behaviors during the last 6 months prior to assessment. Response options were coded: 0 = *never*, 1 = *once*, 2 = *2–3 times*, 3 = *4–5 times*, 4 = *6–9 times*, 5 = *10 or more*. Scores from individual items were summed to create two scale scores for *delinquent behavior* and *sexual behavior*. The internal consistency measures for the delinquent and sexual behaviors were .77, and .79, respectively. Means, standard deviations, and internal consistency estimates of all measures used in this study are shown in Table 1.

Heavy drinking, smoking, and other substance use. Participants were asked how many times they drank six or more cans or bottles of beer, glasses of wine, or drinks of liquor at one time during the last 6 months (*heavy drinking*). The three responses for the different alcoholic beverages were summed and divided to create a composite score of heavy drinking that reflected the average number of times participants engaged in *heavy drinking per month* over the past 6 months. *Cigarette smoking* was assessed using the average quantity that individuals smoked per day during the last 6 months with a seven-option response format: *None, less than 1 cigarette per day, 1–5 cigarettes per day, about one-half pack per day, about 1 pack per day*, and a*bout 2 packs or more per day*. The measures of *marijuana use and other illicit drug use* (e.g., cocaine, stimulants,

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

Table 1. *Descriptive statistics of the variables measured at three time points in adolescence and young adulthood*

Variable	No. of Items	$\hat{\alpha}$	Mean	*SD*
Problem Behaviors Measured at Wave 2[a]				
Delinquent behavior[b]	13	.77	4.21	5.80
Sexual behavior[b]	3	.79	3.97	4.00
Depressive symptoms[b]	20	>.90	15.99	10.99
Heavy drinking[b]	3	—	.86	2.45
Cigarette smoking[b]	1	—	2.83	6.51
Marijuana use[b]	1	—	1.02	3.89
Illicit drug use ex. marijuana[b]	6	—	.63	3.61
Adolescent Correlates Measured at Wave 1[c]				
Usual school grades	1	—	4.96	1.42
Expected years of education	1	—	15.93	1.76
Rhythmicity				
Eating	5	.77	13.48	3.64
Sleep	6	.67	14.84	3.65
Activity				
General	7	.79	19.43	4.54
Sleep	4	.82	10.94	3.55
Low distractibility	5	.74	11.50	3.01
Persistence	3	.70	8.12	1.98
Mood quality	7	.87	23.98	4.20
Rigidity–flexibility	5	.66	14.86	2.73
Young Adult Outcomes Measured at Wave 5[d]				
Education (years)	1	—	15.31	1.90
Perceived physical health	1	—	3.80	0.80
Perceived mental health	1	—	3.90	0.92
Heavy drinking[b]	3	—	1.64	3.44
Cigarette smoking[b]	1	—	4.45	8.19
Marijuana use[b]	1	—	1.47	5.05
Illicit drug use ex. marijuana[b]	6	—	0.25	1.88
Depressive symptoms[b]	20	>.90	10.92	9.00

[a]Average age = 16.0 years.
[b]Reported values are in original scale units; these variables were log-transformed in subsequent analyses.
[c]Average age = 15.5 years.
[d]Average age = 24.2 years.

barbiturates, hallucinogens) assessed how often participants had used these substances during the previous 6 months (*frequency*). The response options were *never*, *a few times in past 6 months*, *about once a month*, *2–3 days a month*, *about once a week*, *2–3 days a week*, *4–5 days a week*, and *every day*. The validity of adolescent self-reports of cigarette smoking (e.g., Wills & Cleary, 1997), drinking and other substance use (e.g., Winters, Stinchfield, Henly, & Schwartz, 1991) has been supported in existing research studies.

Depressive symptoms. Depressive symptoms were assessed using the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977). The CES-D consists of 20 self-report items and provides a unitary measure of current depressive symptomatology, with an emphasis on the affective component, depressed mood. Participants are asked to indicate how many days during the past week they experienced the emotions or behaviors indicated in each of the items. The response options for these items ranged from 0 = *rarely or none of the time* to 3 = *most or all of the time*.

Adolescent correlates measured at Wave 1

Temperament. The Revised Dimensions of Temperament Survey (DOTS-R) was used to assess adolescent temperament at Wave 1 (Windle, 1992; Windle & Lerner, 1986). The DOTS-R is a 54-item self-report measure that assesses 10 temperament attributes. A four-choice response format was used for each item, with responses ranging from 1 = *usually false* to 4 = *usually true*. Concurrent and predictive validity for the DOTS-R have been documented in previous studies (see Windle, 1992). Higher subscale scores indicate higher levels of that temperamental characteristic. Of the 10 temperament attributes, the following eight scales were used in the study: rhythmicity-sleep (regularity of sleeping habits in timing and amount), rhythmicity-eating (regularity of eating habits in timing and amount), activity level-general (overt motor activity, energy, getting fidgety), activity level-sleep (tossing and turning in bed), low distractibility (concentrating and maintaining perceptual focus despite distractions), persistence (continuing in an activity for a relatively long period of time), mood quality (levels of positive affect—smiling

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

and being cheerful), and rigidity–flexibility (responding to changes in environment).

Usual grades in school and years of education expected to complete. Adolescents were asked at Wave 1 what grades they usually got in school. Responses ranged from 1 = *mostly Ds and Fs* to 6 = *mostly As*. A higher score reflected better grades. The correlation between adolescents' reports of their GPA and official high-school records (that used a somewhat different measurement scale) was .78. Therefore, this self-report item was judged as valid. Dornbusch, Mont-Reynaud, Ritter, Chen, and Steinberg (1991) reported a similar correlation (r = .79) with a sample of 5,000 students. Adolescents were also asked how many years of education they thought they would complete in their lifetime.

Young adult outcome variables measured at Wave 5

Years of education completed, and perceived physical and mental health. Young adults reported on the number of years of education that they had completed. For physical and mental health, young adults were asked two questions: "How would you rate your overall physical health?" and "How would you rate your overall mental or emotional health?" Response options ranged from 1 to 5: 1 = *poor*, 2 = *fair*, 3 = *good*, 4 = *very good*, and 5 = *excellent*.

Heavy drinking, smoking, and other substance use. The items assessing heavy drinking, smoking, and other substance use in young adulthood were the same as those used in adolescence.

Psychiatric and substance abuse disorders. Positive lifetime and 12-month psychiatric diagnoses of alcohol use disorder (AUD), cannabis use disorder (CUD), and major depressive disorder (MDD) were derived in the young adulthood interview via the World Health Organisation (WHO) Composite International Diagnostic Interview (WHO, 1997) based on the *DSM-IV* criteria (American Psychological Association, 1994). The prevalence rates of these disorders in our sample are similar to those reported in the National Cormorbidity Study (NCS; Kessler et al., 1994; for more information see Windle & Wiesner, 2004).

Results

Model-based cluster analysis: Identification of groups

The total sample of 1,126 participants had a very small fraction of the data missing (less than 0.1%) on the seven problem behavior variables. The EM algorithm for ML was used for imputation using the SPSS Missing Value Analysis (SPSS, 1989–2004). Little's Missing Completely at Random Test (MCAR; Little, 1988) resulted in nonsignificant, χ^2 (df = 22) = 30.194, p = .114, thus indicating random missing values and subsequent unbiased imputation of missing values. The seven problem behaviors were log-transformed to normalize skewed distributions using the natural logarithm after adding a constant of one. Data were assessed for the degree to which the data fit multivariate normal distributions both before and after subjecting them to data transformations. Although model-based clustering methods work best when the data follow the multivariate normal distribution, the model-based clustering methods are reasonably robust to deviations from the multivariate normal distribution (Hardin & Rocke, 2004; Yeung et al., 2001).

The log-transformed variables were then *centered for males and females separately* to account for a priori known gender differences in the means of these behaviors. This procedure takes into account any known gender differences in the level of these behaviors a priori so that derived cluster membership probabilities would mostly reflect individual differences or relative standings within, but not across, the same gender. If males and females differ only in terms of mean levels, the results would be equivalent to separate analyses for males and females (Venables & Ripley, 2002, p. 308). The results of the model-based cluster analysis are shown in Figure 1. The best fitting model involved a four-cluster solution BIC = −2,828.181. The best model consisted of four clusters of different volumes and shapes. The mixing proportion was .183 (n = 204), .235

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

($n = 265$), .350 ($n = 394$), and .232 ($n = 263$) in the order of Clusters 1–4. The average posterior probabilities for the most likely cluster classification were .996, .989 .994, and .993, respectively, for Clusters 1–4 (see Appendices B and C for more information). This indicates a high degree of classificatory certainty for the four cluster solution.

We observed that the same model did not converge for five and more clusters. We relaxed the default tolerance level for the purpose of evaluating alternate cluster solutions (see Appendix C). The five-cluster model converged with the relaxed tolerance level and yielded a substantially better BIC, −2,362.02, with the mixing proportion .177 ($n = 198$), .144 ($n = 156$), .099 ($n = 116$), .347 ($n = 393$), and .232 ($n = 263$) in the order of Clusters 1–5. We statistically compared the overlap, or the lack thereof, between the two clustering solutions (i.e., the four- and five-cluster solutions) using the variation of information (VI) criterion proposed by Meilă (2003, 2007), which measures the amount of information lost in changing from clustering "a" to clustering "b," and vice versa. This measure has a lower limit of 0 and an upper limit of 1, and is symmetric. The VI criterion, based on information theory, quantifies the amount of the conditional entropy not shared by the mutual information between the clusterings "a" and "b." It is conceptually analogous to $1 - R^2$ in analysis of variance or regression. Unlike Cohen's kappa, this statistic does not require that two clusterings have the same number of categories, nor is it necessary that two clusterings are ordered. The comparison of two clustering solutions yielded a value of .031, a relatively small value for the amount of information lost in comparing the four- and five-cluster solutions. Subsequently, we deemed the five cluster solution as a case of overextraction of components or overfitting, especially given that we greatly relaxed the tolerance level for relative convergence. A common pitfall associated with mixture models is that spurious maxima can be achieved that border on singularities for small but tightly clustered sets of data points (Hipp & Bauer, 2006; Leisch, 2004; McLachlan & Chang, 2004; McLachlan & Peel, 2000). If a component (a cluster or class) is near empty, or two or more components have the same parameters, then it is generally considered as a red flag for the existence of a spurious cluster. One can counter the potential problem of finding a solution based on local optima by using different starting values (Hipp & Bauer, 2006; Steinley, 2003), or by dividing a sample into random subsamples (McLachlan & Chang, 2004).

For the purpose of validation, we randomly selected half of the sample with a selection probability = 0.5 and ran the same mixture model-based cluster analysis for 563 individuals. The same clustering model (diagonal, unequal volumes, and shapes) with four clusters was the best model, resulting in BIC = −1,889.229. The overlap between the full sample and the random half sample was substantial as indicated by the VI criterion = .085, and Cohen $\kappa = .844$, asymptotic standard error = .018, $t = 34.532$, $p < .05$. A Cohen κ exceeding .8 is widely considered as excellent chance-corrected agreement (von Eye & Mun, 2005, pp. 5–6). In addition, we divided the sample into four subsamples of an equal sample size ($n = 281$ or 282) and ran separate cluster analyses. Three or four clusters of the same clustering model (diagonal, unequal volumes, and shapes) were identified. When compared against the full sample, the VI criterion statistics showed evidence of small discrepancy, ranging from .052 to .166. In addition, we also took steps to evaluate that multivariate normality was observed in each cluster by examining the marginal distributions (univariate normality), and also by examining variance for univariate data and generalized variance for multivariate data. Based on the additional validation analyses involving subsamples of different sizes, we deemed the four cluster solution satisfactory.

Model-based cluster analysis: Interpretation of groups

The four clusters included a *multiproblem high-risk* group (Cluster 1, $n = 204$; 99 males, 105 females) characterized by the highest levels of all problem behaviors, including high levels of multiple substance use including marijuana and other illicit substances; a *smoking high-risk* group (Cluster 2, $n = 265$; 97 males, 168 females) characterized by high levels of

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

delinquent behavior, cigarette smoking, marijuana use, and depressive symptoms, and moderate levels of heavy drinking and sexual behavior; a *normative* group (Cluster 3, $n = 394$; 214 males, 180 females) characterized by overall low levels of delinquent and sexual behaviors, depressive symptoms, and heavy drinking; and a *low-risk* group (Cluster 4, $n = 263$; 136 males, 127 females) that showed almost nonexistent or very low levels of problem behaviors. The descriptive statistics of the seven behaviors for the four cluster groups are shown in Table 2 for males and females. The profile plots of seven problem behaviors for males and females are shown in Figure 2. Males and females were unequally represented across the four cluster groups, $\chi^2 = 21.43$, $df = 3$, $p < .05$. Configural frequency analysis (CFA; von Eye, 2002) showed that the data deviated substantially from the assumption of independence (i.e., equal representation of males and females in each cluster) in the following two cluster groups: males were more represented in the Normative group (Cluster 3; 54.3% males) and females were more highly represented in the smoking high-risk group (Cluster 2; 63.4% females) than expected by a model of proportional representation.

Adolescent temperamental correlates

We analyzed how the cluster groups were different across measures of their usual school grades, educational expectations, and the temperamental characteristics assessed at Wave 1. Approximately 19% of the observations were missing for adolescent temperamental and educational variables at Wave 1. We conducted Little's MCAR test (Little, 1988), resulting in nonsignificant χ^2 ($df = 74$) $= 80.316$, $p = .288$. We used the EM algorithm for ML for imputation of the missing values. We included cluster membership, gender, and their interaction as sources of variation (see Table 3). In terms of usual school grades, all three other cluster groups reported lower usual school grades than the low-risk group. The multiproblem high-risk and the smoking high-risk expected to complete fewer years of education than the low risk. Temperamental characteristics were also different across the four cluster groups. The multiproblem high risk and the smoking high risk reported lower levels of rhythmicity in eating and sleeping habits, were more distractible, and less persistent, and reported higher levels of activity compared to the low risk. The multiproblem high risk, however, was less flexible and more active in general activities than the smoking high risk. The Normative was, by and large, similar to the low risk in terms of temperamental characteristics with a few exceptions. The normative was more active and distractible. In addition to group differences because of cluster membership, females reported higher levels of usual school grades and positive mood quality, lower levels of rhythmicity, and were more distractible and less persistent.

Young adult follow-up

Of the 1,126 adolescents included in the model-based cluster analysis, we subsequently analyzed the 762 young adults who were retained at the Wave 5 follow-up.[2] The cluster membership information derived using data from Wave 2 and their continued participation in Wave 5 data collection were not independent, χ^2 (3) $= 9.25$, $p < .05$. The nonparticipation rate was higher for the multiproblem high-risk group (40.7% vs. 28.7, 32.5, and 29.3%

2. There are two conditions under which the missing-data mechanism is said to be *ignorable*. The first condition is *missing at random* (MAR) and the second condition is *distinctness of parameters* (Little & Rubin, 1987; Schafer, 1997). The first condition requires that the missing values may be conditional on the missing data, provided that they are *only indirectly* via observed quantities. The second condition requires that the parameters θ of the data model and the parameters ξ of the mechanism for "missingness" are *distinct*. With imputed data using the EM algorithm under the assumption of ignorable missing-data mechanism, the results reported in Table 4 slightly improve in terms of overall magnitude of η^2 and significant F statistics, because of, in major part, increased sample size and power. The imputed data resulted in additional significant group differences across clusters in perceived mental health and depressive symptoms and between males and females in perceived mental health. Because incomplete data were attributed to dropouts at Wave 5 ($n = 364$, 32.3%), however, we deemed that ignorability was not plausible for Wave 5 data missingness, and opted to report results based on the available 762 cases in Table 4. Likewise, we analyzed the available 723 cases in Table 5.

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

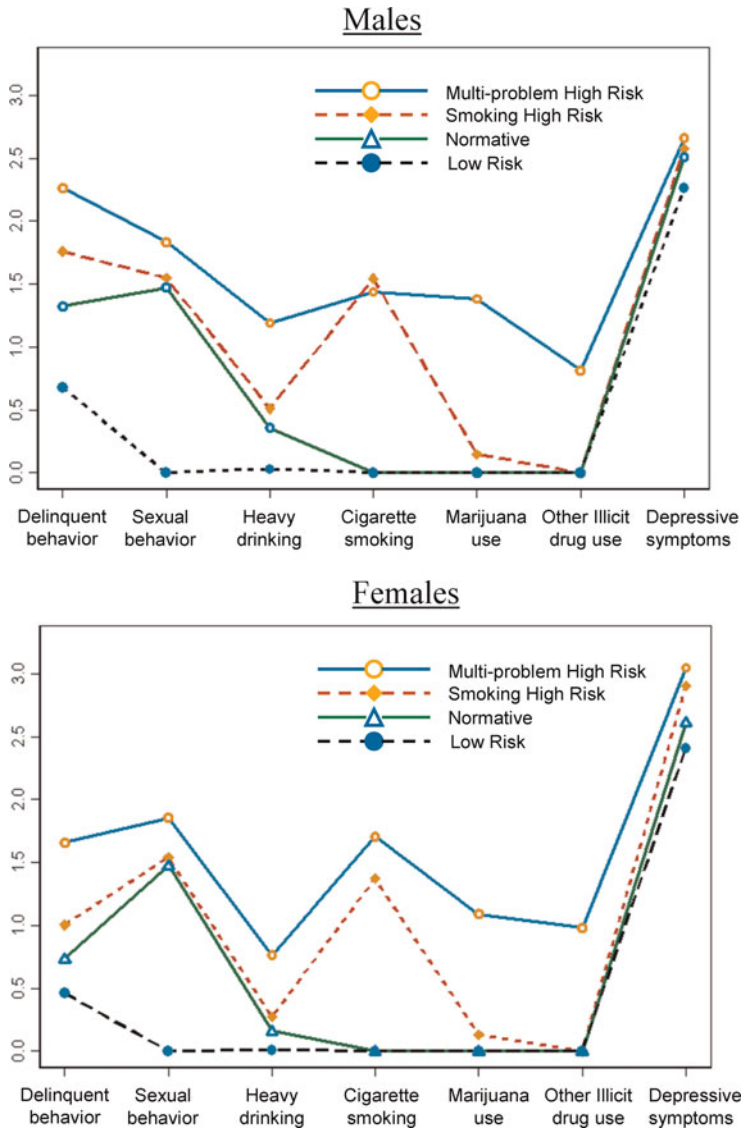


Figure 2. Profile plots of adolescent problem behaviors by cluster groups. The figure illustrates Table 2 in visually, so that the numbers along the *y* axis represent the averages of the log-transformed values reported in Table 2. [A color version of this figure can be viewed online at www.journals.cambridge.org]

compared to the other cluster groups). Individuals who did not participate in the young adult assessment reported higher levels of involvement in delinquent behavior, means = 1.27 versus 1.11, F (df = 1124) = 6.53, $p <$ .05; sexual behavior, means = 1.32 versus 1.16, F (df = 1124) = 7.57, $p <$ .05; cigarette smoking, means = 0.71 versus 0.58, F (df = 1124) = 3.94, $p <$ .05; and marijuana use, means = 0.31 versus 0.27, F (df = 1124) = 4.25, $p <$ .05 among adolescent measures. Therefore, the subsequent findings in young adulthood should be interpreted along with these differential attrition rates that negatively affect the association between cluster membership and young adult variables.

We compared the four groups on the number of years of completed education, overall perceived physical and mental health, depressive symptoms, heavy drinking, cigarette smoking, marijuana use, and other illicit drug use (Table 4). We used cluster membership, gender,

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

Table 2. *Descriptive statistics of adolescent problem behaviors measured at Wave 2 (mean age = 16) by cluster groups (N = 1,126)*

	1. Multiprob. High Risk (n = 99 M, 105 F)	2. Smoking High Risk (n = 97 M, 168 F)	3. Normative (n = 214 M, 180 F)	4. Low Risk (n = 136 M, 127 F)	Total	F	η^2
Delinquent behavior	2.263 (0.807)	1.759 (0.894)	1.323 (0.903)	0.678 (0.795)	1.410 (1.011)	65.98	.292
	1.659 (0.963)	1.006 (0.837)	0.732 (0.755)	0.459 (0.657)	0.919 (0.893)		
Sexual behavior	1.835 (0.779)	1.550 (0.838)	1.471 (0.705)	0.000 (0.000)	1.185 (0.954)	178.70	.528
	1.854 (0.697)	1.540 (0.796)	1.477 (0.666)	0.000 (0.000)	1.240 (0.925)		
Depressive symptoms	2.659 (0.632)	2.576 (0.607)	2.510 (0.631)	2.260 (0.750)	2.486 (0.672)	18.45	.104
	3.047 (0.675)	2.904 (0.585)	2.613 (0.706)	2.408 (0.802)	2.731 (0.727)		
Heavy drinking	1.191 (0.927)	0.511 (0.525)	0.356 (0.475)	0.028 (0.089)	0.453 (0.664)	80.61	.335
	0.761 (0.716)	0.273 (0.340)	0.161 (0.326)	0.005 (0.027)	0.268 (0.470)		
Cigarette smoking	1.436 (1.302)	1.542 (1.004)	0.000 (0.000)	0.000 (0.000)	0.534 (0.997)	191.29	.545
	1.705 (1.160)	1.373 (0.828)	0.000 (0.000)	0.000 (0.000)	0.706 (1.007)		
Marijuana use	1.383 (1.084)	0.147 (0.211)	0.000 (0.000)	0.000 (0.000)	0.277 (0.703)	156.55	.495
	1.090 (1.093)	0.129 (0.207)	0.000 (0.000)	0.000 (0.000)	0.235 (0.626)		
	0.813 (0.930)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.147 (0.504)	127.63	.444
Illicit drug use ex. marijuana	0.981 (0.911)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.178 (0.540)		

Note: The values, which are means (standard deviations), are averages of log-transformed values using the natural logarithm after adding a constant of 1. The top and bottom values in each pair are for males and females, respectively. See Figure 2 for a graphic illustration of this table.

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

Table 3. *Adolescent temperamental correlates at Wave 1 (mean age = 15.5) by cluster groups (N = 1,126)*

	1. Multiprob. High Risk ($n = 99$ M, 105 F)	2. Smoking High Risk ($n = 97$ M, 168 F)	3. Normative ($n = 214$ M, 180 F)	4. Low Risk ($n = 136$ M, 127 F)	Total	F	η^2
Usual school grades	$4.23\ (1.46)_a$	$4.57\ (1.23)_a$	$4.97\ (1.31)_b$	$5.17\ (1.15)_c$	4.82 (1.33)	12.61	.073
	$4.75\ (1.04)_a$	$4.82\ (1.29)_a$	$5.24\ (1.18)_b$	$5.51\ (1.14)_c$	5.09 (1.21)		
Expected years of education	$15.48\ (1.81)_a$	$15.71\ (1.23)_{a,b}$	$15.99\ (1.38)_{b,c}$	$16.11\ (1.52)_c$	15.88 (1.49)	2.87	.018
	$15.70\ (1.85)_a$	$15.87\ (1.63)_{a,b}$	$16.08\ (1.66)_{b,c}$	$16.21\ (1.55)_c$	15.98 (1.67)		
Rhythmicity							
Eating	$13.64\ (2.99)_a$	$13.15\ (3.35)_a$	$14.08\ (3.08)_b$	$14.26\ (3.12)_b$	13.88 (3.14)	6.91	.041
	$12.38\ (3.56)_a$	$12.70\ (3.00)_a$	$13.10\ (3.42)_b$	$14.22\ (3.24)_b$	13.10 (3.35)		
Sleep	$14.51\ (2.98)_a$	$15.02\ (3.39)_a$	$15.41\ (3.21)_{a,b}$	$15.28\ (3.14)_b$	15.14 (3.19)	5.16	.031
	$14.07\ (3.42)_a$	$14.11\ (2.91)_a$	$14.45\ (3.28)_{a,b}$	$15.68\ (3.63)_b$	14.55 (3.33)		
Activity							
General	$20.58\ (4.07)_a$	$19.88\ (3.56)_b$	$19.60\ (4.13)_b$	$18.90\ (3.83)_c$	19.65 (3.97)		
	$21.02\ (3.78)_a$	$19.76\ (3.97)_b$	$18.77\ (4.25)_b$	$17.64\ (3.91)_c$	19.21 (4.16)	8.65	.051
Sleep	$11.07\ (3.13)_a$	$10.73\ (2.97)_{a,b}$	$11.38\ (2.74)_a$	$10.85\ (3.36)_b$	11.07 (3.02)		
	$11.66\ (3.36)_a$	$10.88\ (2.97)_{a,b}$	$10.90\ (3.39)_a$	$9.96\ (3.53)_b$	10.82 (3.34)	3.21	.020
Low distractibility	$11.34\ (2.51)_a$	$11.57\ (2.82)_a$	$11.81\ (2.44)_a$	$12.41\ (2.58)_b$	11.83 (2.58)		
	$10.87\ (2.74)_a$	$10.89\ (2.65)_a$	$11.19\ (2.76)_a$	$11.86\ (2.96)_b$	11.19 (2.79)	5.48	.033
Persistence	$7.99\ (1.80)_{a,b}$	$8.11\ (1.74)_a$	$8.32\ (1.61)_{b,c}$	$8.71\ (1.62)_c$	8.32 (1.69)		
	$7.79\ (1.83)_{a,b}$	$7.47\ (1.84)_a$	$8.10\ (1.85)_{b,c}$	$8.36\ (1.60)_c$	7.92 (1.82)	7.03	.042
Mood quality	$23.34\ (4.27)_a$	$22.50\ (4.55)_{a,b}$	$23.80\ (3.34)_b$	$23.85\ (3.59)_b$	23.50 (3.84)		
	$23.35\ (4.19)_a$	$24.51\ (3.53)_{a,b}$	$24.75\ (3.50)_b$	$24.74\ (3.41)_b$	24.42 (3.65)	5.48	.033
Rigidity–flexibility	$14.27\ (2.20)_a$	$14.79\ (2.61)_b$	$14.95\ (2.21)_b$	$14.93\ (2.83)_b$	14.79 (2.45)		
	$14.10\ (2.67)_a$	$14.70\ (2.53)_b$	$15.36\ (2.16)_b$	$15.27\ (2.35)_b$	14.92 (2.45)	4.14	.025

Note: Of the two rows of pairs of means (standard deviations), the numbers on the top and the bottom indicate mean values for males and females, respectively. Values with the same subscript are not significantly different from each other based on Tukey B post hoc tests. Underlined total means indicate significant gender differences. Interaction effects between cluster membership and gender on rhythmicity sleeping and mood quality were statistically significant. All F statistic values are significant at $\alpha \leq .05$.

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

Table 4. *Young adult follow-up measures at Wave 5 (mean age = 24.2) by cluster groups (N = 762)*

	1. Multiprob. High Risk (n = 54 M, 67 F)	2. Smoking High Risk (n = 55 M, 134 F)	3. Normative (n = 131 M, 135 F)	4. Low Risk (n = 93 M, 93 F)	Total	F	η^2
Education (years)	14.35 (2.03)$_a$	14.53 (1.69)$_a$	15.44 (2.02)$_b$	15.32 (1.81)$_b$	15.08 (1.96)	6.01	.052
	15.22 (1.83)$_a$	15.13 (2.06)$_a$	15.72 (1.71)$_b$	15.88 (1.58)$_b$	15.49 (1.84)		
Physical health	3.68 (0.88)$_a$	3.71 (0.79)$_{a,b}$	3.81 (0.80)$_b$	3.77 (0.76)$_b$	3.76 (0.80)	2.56	.023
	3.57 (0.86)$_a$	3.76 (0.81)$_{a,b}$	3.93 (0.75)$_b$	4.00 (0.79)$_b$	3.84 (0.81)		
Mental health	3.89 (0.85)	3.89 (0.92)	4.04 (0.96)	4.03 (0.84)	3.99 (0.90)	1.56, *ns*	.014
	3.66 (0.99)	3.84 (0.95)	3.85 (0.90)	3.95 (0.90)	3.84 (0.93)		
Heavy drinking	1.29 (0.90)$_a$	1.19 (0.78)$_b$	0.81 (0.82)$_b$	0.42 (0.65)$_c$	0.84 (0.84)	23.73	.181
	0.57 (0.85)$_a$	0.40 (0.61)$_b$	0.35 (0.54)$_b$	0.25 (0.48)$_c$	0.38 (0.61)		
Cigarette smoking	1.46 (1.39)$_a$	1.69 (1.38)$_a$	0.59 (1.10)$_b$	0.29 (0.73)$_b$	0.83 (1.24)	19.31	.152
	1.28 (1.31)$_a$	1.23 (1.26)$_a$	0.52 (0.98)$_b$	0.35 (0.85)$_b$	0.82 (1.17)		
Marijuana use	1.06 (1.25)$_a$	0.58 (1.07)$_b$	0.39 (0.88)$_b$	0.08 (0.36)$_c$	0.44 (0.94)	14.20	.116
	0.57 (0.96)$_a$	0.16 (0.45)$_b$	0.18 (0.58)$_b$	0.11 (0.40)$_c$	0.22 (0.60)		
Illicit drug use ex. marijuana	0.20 (0.40)$_a$	0.14 (0.43)$_b$	0.11 (0.46)$_b$	0.01 (0.10)$_b$	0.10 (0.38)	4.94	.044
	0.23 (0.66)$_a$	0.02 (0.16)$_b$	0.04 (0.20)$_b$	0.03 (0.13)$_b$	0.06 (0.31)		
Depressive symptoms	2.12 (0.83)	2.35 (0.69)	2.19 (0.77)	2.05 (0.89)	2.17 (0.81)	1.20, *ns*	.011
	2.31 (0.88)	2.23 (0.92)	2.16 (0.78)	2.06 (0.96)	2.18 (0.88)		

Note: The values are means (standard deviations). The top and bottom values in each pair are for males and females, respectively. Reported values of heavy drinking, cigarette smoking, marijuana use, illicit drug use, and depressive symptoms are averages of log-transformed values using the natural logarithm after adding a constant of 1. Values with the same subscript are not significantly different from each other based on Tukey B post hoc tests. Underlined total means indicate significant gender differences. The interaction effects between cluster membership and gender on heavy drinking and marijuana use were statistically significant. All F statistic values are significant at $\alpha \leq .05$ unless noted otherwise.

and their interaction as explanatory terms in the model. The cluster membership derived at age 16 significantly predicted young adult outcome variables. The clusters had different mean locations in the years of education completed, overall physical health, heavy drinking, cigarette smoking, marijuana use, and illicit drug use (excluding marijuana use). In addition, there were gender differences for heavy drinking, marijuana use, and years of education completed, such that males reported higher levels of heavy drinking and marijuana use than females, and females reported completing more years of education. Interaction effects existed between gender and the cluster membership for heavy drinking and marijuana use.

Overall, the multiproblem high risk showed poor developmental profiles as young adults as indicated by highest levels of heavy drinking, marijuana use, and other illicit drug use. The multiproblem high risk also reported lower levels of educational attainment and physical health; the smoking high risk was statistically not distinguished from the multiproblem high risk on the domains of years of completed education, perceived physical health, and cigarette smoking, but significantly separated from the multiproblem high risk on heavy drinking, marijuana use, and illicit drug use. Compared to the normative who reported elevated levels of heavy drinking and marijuana use in young adulthood, the smoking high risk had lower levels of education and higher levels of cigarette smoking. The normative was, by and large, similar to the low-risk group with two exceptions of moderate levels of marijuana use and heavy drinking.

Lifetime and 12-month substance use and MDD diagnoses

Lifetime and 12-month substance use disorder and MDD diagnoses were made for 313 males and 410 females of the 762 participants at Wave 5. Chi-square analysis showed that lifetime AUD and CUD were significantly associated with the cluster membership for males and females (see Table 5 and Figure 3). The multiproblem high risk, in particular, had 78.4 and 46.8% of males and females, respectively, who met criteria for a lifetime AUD diagnosis. The multiproblem high risk also showed 49.0 and 45.2% of males and females, respectively, who met criteria for a lifetime CUD diagnosis. The AUD and CUD diagnoses for the past 12 months were statistically associated with cluster group membership only for males. For males, the multiproblem high risk had 23.5 and 11.8% of the individuals meeting criteria for the 12-month AUD and CUD diagnoses, respectively. The AUD and CUD diagnoses for the past 12 months were not associated with cluster group membership for females. The lifetime and 12-month MDD diagnoses were not associated with cluster membership for males or females.

Discussion

Model-based cluster analysis

The current study sought to examine the heterogeneity of problem behaviors among adolescents using a mixture model-based cluster analysis approach. This new analytical procedure is a tool to empirically investigate population heterogeneity and accordingly to classify subpopulations using more rigorous statistical methods for the comparison of alternative models. Model-based cluster analysis utilizing finite mixture densities can be a valuable analytic tool for research on developmental psychopathology for a number of reasons. First, model-based cluster analysis can be used to generate a new set of hypotheses based on detected salient patterns of cases or individuals. Because developmental psychopathology is an integrative discipline that often faces unique challenges to synergize emerging cross-area information (e.g., genetics, neuroscience, psychology, sociology, epidemiology) as they become available to better understand the mechanisms through which psychopathology develops (Cicchetti & Rogosch, 2002), it often becomes necessary for researchers to explore data to generate new hypotheses for future directions. Second, this method can be used in comparing subgroups when measurement invariance across subpopulations cannot be established. Measurement invariance is defined as the measurement model that relates the observed variables to the underlying factors is identical

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

Table 5. *Young adult substance use disorder by cluster groups (N = 723)*

	1. Multiprob. High Risk	2. Smoking High Risk	3. Normative	4. Low Risk	Total
Male ($n = 313$)					
Lifetime AUD, χ^2 (3) = 31.80, $p < .05$	78.4% (40)	56.9% (29)	45.9% (56)	30.3% (27)	48.6% (152)
Lifetime CUD, χ^2 (3) = 38.14, $p < .05$	49.0% (25)	31.4% (16)	19.7% (24)	5.6% (5)	22.4% (70)
Lifetime MDD, χ^2 (3) = 1.10, *ns*	17.6% (9)	19.6% (10)	13.9% (17)	14.6% (13)	15.7% (49)
12-month AUD, χ^2 (3) = 12.27, $p < .05$	23.5% (12)	17.6% (9)	11.5% (14)	4.5% (4)	12.5% (39)
12-month CUD, χ^2 (3) = 8.12, $p < .05$	11.8% (6)	11.8% (6)	8.2% (10)	1.1% (1)	7.3% (23)
12-month MDD, χ^2 (3) = 2.64, *ns*	5.9% (3)	3.9% (2)	1.6% (2)	2.2% (9)	5.1% (16)
Female ($n = 410$)					
Lifetime AUD, χ^2 (3) = 27.09, $p < .05$	46.8% (29)	21.1% (27)	23.3% (30)	11.0% (10)	23.4% (96)
Lifetime CUD, χ^2 (3) = 56.76, $p < .05$	45.2% (28)	10.9% (14)	10.9% (14)	4.4% (4)	14.6% (60)
Lifetime MDD, χ^2 (3) = 3.38, *ns*	29.0% (18)	26.6% (34)	20.9% (27)	18.7% (17)	23.4% (96)
12-month AUD, χ^2 (3) = .48, *ns*	3.2% (2)	4.7% (6)	4.7% (6)	3.3% (3)	4.1% (17)
12-month CUD, χ^2 (3) = 2.79, *ns*	0.0% (0)	1.6% (2)	0.0% (0)	1.1% (1)	0.7% (3)
12-month MDD, χ^2 (3) = .53, *ns*	3.2% (2)	3.9% (5)	2.3% (3)	3.3% (3)	3.2% (13)

Note: AUD, alcohol use disorder; CUD, cannabis use disorder; MDD, major depressive disorder. Percentages are the proportions of cases meeting criteria for a diagnosis within each of the cluster groups. The numbers in parentheses indicate the actual number of case criteria for a diagnosis within each of the cluster groups.

across subpopulations and requires equality of factor loadings, intercepts, and residual variances across subpopulations (Meredith, 1993). Model-based cluster analysis does not require the assumption of measurement invariance across different populations, and may be used as an alternative method when the assumption of measurement invariance is either untenable or unreasonable. Third, a model-based cluster analysis can be used to address person-oriented questions that have arisen through variable-oriented research. For example, the literature on child maltreatment suggests that of the three types of often co-occurring maltreatment (i.e., emotional maltreatment, physical neglect and abuse, and sexual abuse), only emotional maltreatment predicted slower increases in self-esteem and slower decreases in depressive symptoms over time for 7- to 12-year-olds when initial levels of depressive symptoms, age, and the other types of maltreatment were controlled (Kim & Cicchetti, 2006). One important follow-up task may be to *empirically find* the group of children meeting the aggregate-level prediction and other groups of children who do not, and to then investigate their biological, psychological, and environmental characteristics that guided them toward divergent paths.

In the present study we identified four clusters of adolescents who varied with regard to their specific risk behaviors and levels of seven problem behaviors, and we found that the four clusters of adolescents differed significantly from one another on a number of adolescent temperamental characteristics and across young adult outcome measures. The results demonstrate the utility of a model-based cluster analysis as an improved alternative method to heuristic cluster analytic procedures when finding heterogeneous groups in a population. However, it is important to note that any classification should pass the tests of not only statistical validity but the *usefulness or meaningfulness* of the derived clusters (Everitt et al., 2001; Muthén, 2003; Rindskopf, 2003; von Eye & Bergman, 2003). Citing a point made by Needham (1965) regarding the classification of humans

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

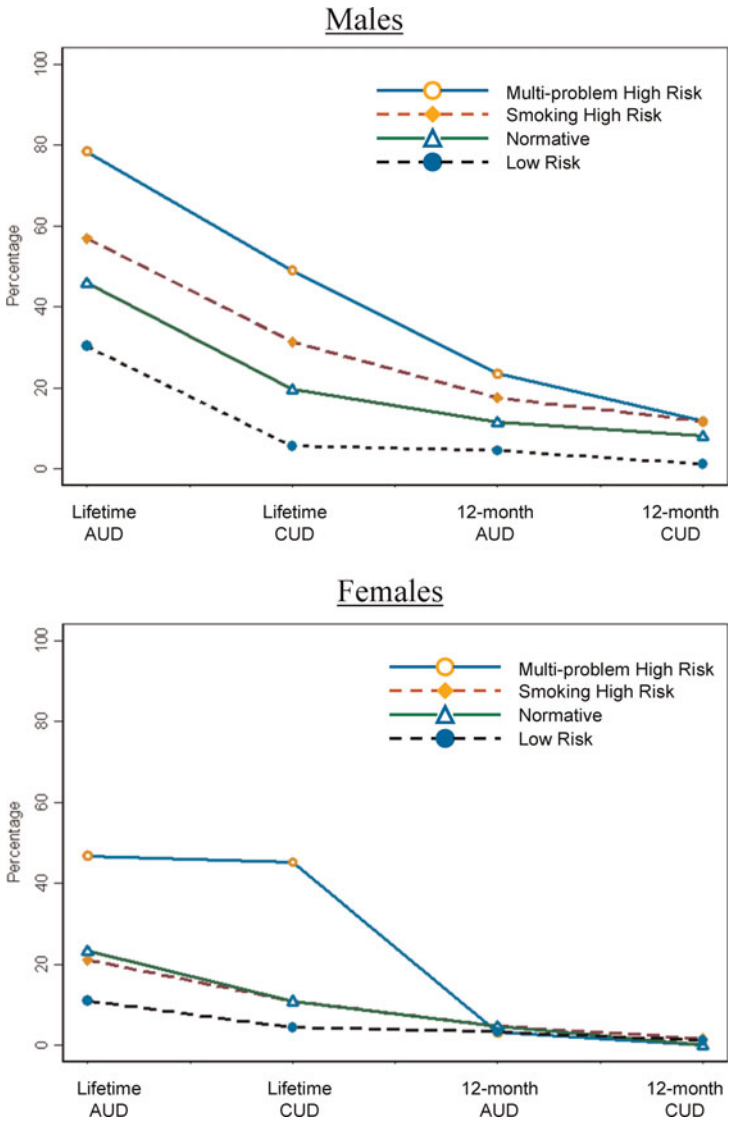


Figure 3. Lifetime and 12-month alcohol use disorder and cannabis use disorder diagnoses by cluster groups in young adulthood. [A color version of this figure can be viewed online at www.journals.cambridge.org]

into men and women and its usefulness as a classification because it carries more information than mere anatomical differences, Everitt and colleagues (2001, p. 4) illustrated this point using an analogy of classifying books:

A similar point can be made in respect of the classification of books based on subject matter and their classification based on the color of the book's binding. The former, with classes such as dictionaries, novels, biographies, etc., will be of far wider use than the latter with classes such as green, blue, red, etc. The reason why the first is more useful than the second is clear; the subject matter classification indicates more of a book's characteristics than the latter. So it should be remembered that in general a classification of a set of objects is not like a scientific theory and should be judged largely on *its usefulness*, *rather than in terms of whether it is "true" or "false"* [italics added].

Therefore, results from cluster analyses should be evaluated not only by the ways that objects are classified but also with regard to

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

the *generality or usefulness of the clusters* for subsequent use, which, in the present study, demonstrated predictive utility for young adult outcome measures.

Four clusters and their profiles

Substantively, we identified four clusters based on the problem behaviors of delinquent behaviors, sexual behavior, substance use, and depressive symptoms in adolescence. Examining clusters of individuals based on multiple problem behaviors enabled us to find distinctive cluster profiles of adolescents. As indicated by profile plots, illicit drug use at age 16 appears to be a good proxy of risk for a constellation of problem behaviors, including delinquent and violent behaviors for a normative community-based sample. The multiproblem high-risk group accounted for 18.3% of the total sample and exhibited the highest levels of other concurrent problem behaviors measured in this study, including delinquent behavior, other substance use (e.g., cigarette smoking, marijuana use, heavy drinking), and depressive symptoms, and reported lower school grades. Temperamentally, the multiproblem high-risk group was more irregular in terms of eating and sleeping habits, more rigid and distractible to changes in the environment, more active, less persistent in pursuing tasks, and had poor mood quality, compared to the low risk in adolescence. Individuals of this high-risk group had lower levels of usual school grades, educational expectations for themselves in adolescence, and subsequently completed fewer years of education by young adulthood. They also reported poor overall physical health, and continued involvement in heavy drinking, cigarette smoking, marijuana use, and other illicit drug use in young adulthood. The multiproblem high risk also featured elevated proportions of cases meeting diagnostic criteria for lifetime AUD and CUD for both males and females. High proportions of males in this cluster group also met criteria for the 12-month AUD and CUD diagnoses.

The smoking high-risk group accounted for 23.5% of the total sample. The smoking high risk manifested high levels of cigarette smoking, marijuana use, and depressive symptoms in adolescence. Temperamentally and academically, they closely mirrored the profiles of the highest risk group. Compared to the low risk, they were more irregular and hyperactive, more distractible, and less persistent in their self-reported temperament and, as expected, had lower levels of usual school grades, and educational expectations for themselves in adolescence and subsequently completed fewer years of education by young adulthood. Although these individuals also maintained involvement in heavy drinking, cigarette smoking, and marijuana use in young adulthood, the smoking high-risk groups reported lower levels of heavy drinking, marijuana use, and other illicit drug use in young adulthood than the multiproblem high risk. The normative and the low-risk groups accounted for 35.0 and 23.2% of the total sample, respectively. Although the normative engaged in moderate levels of delinquent and sexual behaviors, and heavy drinking, they were largely inseparable from the low risk in terms of their temperamental and educational attributes in adolescence. However, there was evidence that some individuals of the normative, compared to the low risk, showed involvement in marijuana use and heavy drinking episodes in young adulthood, which may be interpreted as occasional substance use in social settings for this subset of individuals.

The continuity and comorbidity path from middle adolescence to young adulthood

The transitional period from adolescence to emerging young adulthood is "a window of opportunity for changing the life course" in terms of having second-chance opportunities or experiencing a turning point (Masten et al., 2004). Adult roles can alter one's own expectations and circumstances, thus eliciting encouraging changes (i.e., "adult-role socialization"). It is known that marriage itself is associated with reduced alcohol use (Miller-Tutzauer, Leonard, & Windle, 1991) and with reduced crime (Sampson, Laub, & Wimer, 2006). However, the current study showed that there is also remarkable continuity of risk and dysfunction from adolescence to young adulthood. The current findings suggest that a majority of problem or risk behaviors are carried out by a small

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

group of adolescents (multiproblem high risk, 18.3%) who use hard drugs by age 16 and that these adolescents are more deeply entrenched of the two high risk groups toward the addictive life path as the gateway theory suggests (Kandel, 1975; Newcomb & Bentler, 1988), and *illicit drug use by age 16* appears to be a good proxy of risk for other problem behaviors.

The gateway theory of substance use suggests that there is an orderly stagelike progressive sequence of substance use that adolescents go through starting with legal substances such as tobacco or alcohol and proceeding to illicit hard drugs such as marijuana and cocaine, and that the orderly sequence is hard to reverse or to deviate from once initiated as if walking through a "tunnel" (Kandel, 2002). Although this is a rigid picture, it may be best understood in the context of recent studies that early entry into substance use such as heavy alcohol use, in and of itself, is shown to damage the prefrontal cortex, which is critical for executive functions such as impulse control and goal setting in animal models, and to lead to impaired cognitive functioning in human adolescents (Monti et al., 2005; Reyna & Farley, 2006). Therefore, it is plausible, for instance, that adolescents who initiate heavy drinking at an earlier age would be vulnerable for other risk behaviors because their decision-making ability to avoid them is compromised in ways that are stable across time and situations in addition to acute, erratic, risky decisions because of acute alcohol consumption.

The current findings are consistent with other studies on the age of onset as an indicator of risk that have suggested that ages 12–15 as a sensitive time window in development for initiation of a range of problem behaviors for at-risk individuals (e.g., age 12 for any substance use; Kaplow, Curran, Dodge, & the Conduct Problems Research Group, 2002; and age 15 for sexual intercourse; Capaldi, Stoolmiller, Clark, & Owen, 2002; Raffaelli & Crockett, 2003). Current findings of illicit drug use at age 16 as a proxy of serious risks suggest that in early teen years coinciding with the advent of puberty, some adolescents initiate risk behaviors beginning with tobacco and alcohol use, then rapidly progress toward other problem behaviors of greater potential harms, and expand their inventory of problem behaviors by middle adolescence. In addition, once adolescents are on the risky path, they are more likely to have to tackle behavioral and social consequences of being arrested and questioned by police, or being suspended from school in adolescence.

In sum, the continuity and comorbidity path from middle adolescence to young adulthood may progress through a sequence of risk aggregation. First, some adolescents appear not to be on equal footing with others when facing risk behaviors as indicated by adolescent risky temperamental characteristics. For example, adolescents who are more active during sleep (low risk vs. all others) could be an indication of sleep problems that are increasingly linked to adjustment problems in childhood and adolescence, including early onset of alcohol and other drug use (Wong, Brower, Fitzgerald, & Zucker, 2004), and neurobehavioral functioning and behavior problems (Sadeh, Gruber, & Raviv, 2002), although the exact mechanisms of these associations remain unknown. Second, adolescents with early onset of multiple substance use and sustained use may become even more vulnerable by their own use. At a time period when the adolescent brain continues to develop well into the mid-20s, especially in the frontal lobes critical for executive functions such as impulse control, substance-abusing adolescents may be sabotaging their own chances in life by putting themselves at even higher risk for other risk behaviors that require better judgment and decision-making to avoid escalating, multiple problems. Thus, the maladaptive patterns of behavioral and emotional self-regulation, and reactivity to the demands of the surrounding environment, become more entrenched. Third, as at-risk adolescents develop, they may be more active in constructing their niche with others with similar attributes, values, and behaviors (e.g., selection of deviant peers), thereby the interlocking continuity gets reinforced as they transition into young adulthood. From a prevention science perspective, the current study underscores the importance of providing early intervention that is tailored for potentially variable needs of different adolescents. Empirically identifying adolescents at risk via sound methods would provide initial

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

steps for understanding and devising ways to intervene and gauging the intensity of the intervention needed.

Limitations and Future Directions

The present study has several limitations. First, the sample for this study was geographically and ethnically limited. Although the schools selected were public schools, the majority of the participants were Catholic, European American, and all participants resided in Western New York at the initial assessment. The sample in this study represents a normative, community-based adolescent sample that includes some high-risk but also low-risk adolescents. Thus, the cluster findings should be generalized to normative adolescent populations; however, they cannot be generalized to high-risk or at-risk populations such as juvenile delinquents. In addition, the current sample was drawn from a few primary sampling units (schools) within which adolescents were nested. This nested data structure typically results in biased standard errors. The exact degree of the bias is unknown, and the study results should be interpreted cautiously because of its homogeneous data characteristics. Second, we relied on self-reports by adolescents to assess their problem behaviors. Although self-reported measures of problem behaviors are widely used in the literature and the validity of those measures have been shown (e.g., smoking measure; Wills & Cleary, 1997, substance use; Winters et al., 1991), the extent to which these problem behaviors are correlated because of shared methods is unclear. Third, the present study did not exhaustively include all domains of interest in young adulthood and was limited to young adult outcomes of years of completed education, overall perceived mental and physical health, substance use, and substance use disorders. One obvious area not covered in this study was risky sexual behavior for all participants and reproductive health for females as one of the young adult outcomes, which has been suggested to be related to drinking and substance use in women (see Wilsnack, Plaud, Wilsnack, & Klassen, 1997). In addition, the sexual behavior measured at age 16 mostly likely reflected initiation of sexual activities rather than persistently higher risky sexual behavior (e.g., unprotected sex with multiple partners). For this reason, we suspect that sexual behavior at age 16 was not uniquely related to adolescent and young adult measures included in the present study. For example, the normative did not differ from the low risk in almost all measures. Consistent with data from a national sample of 11th graders, which indicated that 53.2% of adolescents have experienced sexual intercourse (Grunbaum et al., 2004), sexual initiation by age 16 may not be a good measure of generalized problem or risky behaviors. In addition, previous studies have shown that risk factors such as self-regulation may be associated more with risky sexual behavior rather than with the initiation of sexual activity (Raffaelli & Crockett, 2003). We also did not assess tobacco dependence at Wave 5, which would be of great interest, especially for the smoking high-risk group.

Despite these limitations, we believe that the current study has strengths and contributes to the literature. First, this study utilized a large sample of adolescents and found four distinctive clusters of adolescents that differed not only in their risk level but also in their risk *kind*. Many existing studies have focused primarily on the unique effects of one variable (e.g., cigarette smoking) above and beyond the effects of other predictors (e.g., heavy drinking). This approach does not address the issue of co-occurring problem behaviors at the individual level. In other recent studies that have used latent growth analysis, the analytic focus has most often been on one problem behavior at a time, thereby preventing the opportunity to understand the common and unique etiological processes of multiple behavior problems underlying their expression (Chassin et al., 2004). The approach taken in this study allowed us to identify other at-risk profiles that may be relevant to understanding and promoting successful transitions to adulthood. Moreover, unlike many other existing clustering procedures, the new mixture clustering method used in this study adopts finite mixture densities and the maximum likelihood estimation method that is ideal for finding groups unknown in advance but where patterns are suspected. The model-based cluster analysis

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

using MCLUST program (Fraley & Raftery, 2002a, 2002b, 2003), in particular, eliminates the subjective arbitrariness one faces when selecting a clustering procedure and deciding the number of clusters, which is especially important because a trade-off generally exists between model complexity and the number of clusters (e.g., Bauer & Curran, 2003, 2004).

Second, the current study not only identified a high-risk cluster of adolescents and their adolescent correlates at age 15.5–16 years old, but also prospectively followed them into young adulthood when they were approximately 24.2 years old. Results indicated that in addition to the highest risk group of adolescents identified in adolescence (i.e., the multiproblem high-risk group), there were other individuals sharing the unique risk profile of problem behaviors. The smoking high-risk group, in particular, continued to participate in heavy drinking, cigarette smoking, and marijuana use as young adults, although it is unclear from the current study whether their risk is reduced only in levels, compared to the multiproblem high risk, because of a lack of diverse young adult measures. The findings in this study direct us to contemplate the importance of investigating aggregated problem behaviors from a within-person perspective in the course of adolescence and young adulthood. Third, this study included depressive symptoms and sexual behaviors, which are highly prevalent among females. Thus, we were able to derive clusters that may be more typical of a normative sample for both male and female adolescents. It is highly important to learn more about different subgroups of individuals that differ not only in their level of risk but also in the nature (e.g., developmental changes) of their risk so that we can understand common and unique causal mechanisms associated with being a member of different clusters, and ultimately contribute to the development of effectively tailored prevention and intervention programs based on this knowledge. New studies may gain from the approach taken in this study to find a mixture of clusters for developmental psychopathology research and to conduct follow-up studies with more ethnically and racially diverse samples and with a broader range of risk and protective factors, as well as outcomes.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Arminger, G., & Stein, P. (1997). Finite mixtures of covariance structure models with regressors. *Sociological Methods and Research*, *26*, 148–182.

Arminger, G., Stein, P., & Wittenberg, J. (1999). Mixtures of conditional mean- and covariance-structure models. *Psychometrika*, *64*, 475–494.

Arnett, J. J. (1992). Reckless behavior in adolescence: A developmental perspective. *Developmental Review*, *12*, 339–373.

Arnett, J. J. (1999). Adolescent storm and stress, reconsidered. *American Psychologist*, *54*, 317–326.

Banfield, J. D., & Raftery, A. E. (1993). Model-based Gaussian and non-Gaussian clustering. *Biometrics*, *49*, 803–821.

Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods*, *8*, 338–363.

Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods*, *9*, 3–29.

Bergman, L. R., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology*, *9*, 291–319.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, *53*, 605–634.

Bongers, I. L., Koot, H. M., van der Ende, J., & Verhulst, F. C. (2004). Developmental trajectories of externalizing behaviors in childhood and adolescence. *Child Development*, *75*, 1523–1537.

Brame, B., Nagin, D. S., & Tremblay, R. E. (2001). Developmental trajectories of physical aggression from school entry to late adolescence. *Journal of Child Psychology and Psychiatry*, *42*, 503–512.

Capaldi, D. M., & Stoolmiller, M., Clark, S., & Owen, L. D. (2002). Heterosexual risk behaviors in at-risk young men from early adolescence to young adulthood: Prevalence, prediction, and association with STD contraction. *Developmental Psychology*, *38*, 394–406.

Celeux, G., & Govaert, G. (1992). Comparison of the mixture and the classification maximum likelihood in cluster analysis. *Journal of Statistical Computation and Simulation*, *47*, 127–146.

Chassin, L., Flora, D. B., & King, K. M. (2004). Trajectories of alcohol and drug use and dependence from adolescence to adulthood: The effects of familial alcoholism and personality. *Journal of Abnormal Psychology*, *113*, 483–498.

Chassin, L., Pitts, S. C., & Prost, J. (2002). Heavy drinking trajectories from adolescence to emerging adulthood in a high-risk sample: Predictors and substance abuse outcomes. *Journal of Consulting and Clinical Psychology*, *70*, 67–78.

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

Chassin, L., Presson, C. C., Pitt, S. C., & Sherman, S. J. (2000). The natural history of cigarette smoking from adolescence to adulthood in a Midwestern community sample: Multiple trajectories and their psychosocial correlates. *Health Psychology*, *19*, 223–231.

Chassin, L., Presson, C. C., Sherman, S. J., & Edwards, D. A. (1991). Four pathways to young-adult smoking status: Adolescent social-psychological antecedents in a midwestern community sample. *Health Psychology*, *10*, 409–418.

Cicchetti, D., & Rogosch, F. A. (2002). A developmental psychopathology perspective on adolescence. *Journal of Consulting and Clinical Psychology*, *70*, 6–20.

Cicchetti, D., & Toth, S. L. (1998). The development of depression in children. *American Psychologist*, *53*, 221–241.

Colder, C. R., Campbell, R. T., Ruel, E., Richardson, J. L., & Flay, B. R. (2002). A finite mixture model of growth trajectories of adolescent alcohol use: Predictors and consequences. *Journal of Consulting and Clinical Psychology*, *70*, 67–78.

Colder, C. R., Mehta, P., Balanda, K., Campbell, R. T., Mayhew, K. P., & Stanton, W. R., et al. (2001). Identifying trajectories of adolescent smoking: An application of latent growth mixture modeling. *Health Psychology*, *20*, 127–135.

Crockett, L. J., Moilanen, K. L., Raffaelli, M., & Randall, B. A. (2006). Psychological profiles and adolescent adjustment: A person-centered approach. *Development and Psychopathology*, *18*, 195–214.

Dahl, R. E. (2004). Adolescent brain development: A period of vulnerabilities and opportunities. *Annals New York Academy of Sciences*, *1021*, 1–22.

Dolan, C. V., Schmittmann, V. D., Lubke, G. H., & Neale, M. C. (2005). Regime switching in the latent growth curve mixture model. *Structural Equation Modeling*, *12*, 94–119.

Dolan, C. V., & van der Mass, H. L. J. (1998). Fitting multivariate normal finite mixtures subject to structural equation modeling. *Psychometrika*, *63*, 227–253.

Dornbusch, S. M., Mont-Reynaud, R., Ritter, P. L., Chen, Z., Steinberg, L. (1991). Stressful events and their correlates among adolescents of diverse backgrounds. In M. E. Colten & S. Gore (Eds.), *Adolescent stress: Causes and consequences* (pp. 111–130). New York: Aldine de Gruyter.

Dumenci, L., & Windle, M. (2001). Cluster analysis as a method of recovering types of intraindividual growth trajectories: A Monte Carlo study. *Multivariate Behavioral Research*, *36*, 501–522.

Eaton, D. K., Kann, L., Kinchen, S., Ross, J., Hawkins, J., Harris, W. A., et al. (2006). Youth risk behavior surveillance—United States, 2005. *Morbidity and Mortality Weekly Report*, *55*, 1–108.

Elliott, D. S., Huizinga, D., & Menard, S. (1989). *Multiple problem youth*. New York: Springer–Verlag.

Everitt, B. S. (2005). Finite mixture distributions. In B. S. Everitt & D. Howell (Eds.), *Encyclopedia of statistics in behavioral science*. London: Wiley.

Everitt, B. S., & Hand, D. J. (1981). *Finite mixture distributions*. London: Chapman & Hall CRC.

Everitt, B. S., Landau, S., & Leese, M. (2001). *Cluster analysis* (4th ed.). London: Arnold.

Flory, K., Lynam, D., Milich, R., Leukefeld, C., & Clayton, R. (2004). Early adolescent through young adulthood alcohol and marijuana use trajectories: Early predictors, young adult outcomes, and predictive utility. *Development and Psychopathology*, *16*, 193–213.

Fraley, C., & Raftery, A. E. (1998). How many clusters? Which clustering method? Answer via model-based cluster analysis. *The Computer Journal*, *41*, 578–588.

Fraley, C., & Raftery, A. E. (1999). MCLUST: Software for model-based cluster analysis. *Journal of Classification*, *16*, 297–206.

Fraley, C., & Raftery, A. E. (2002a). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, *97*, 611–631.

Fraley, C., & Raftery, A. E. (2002b). *MCLUST: Software for model-based clustering, density estimation, and discriminant analysis* (Technical Report No. 415). Seattle, WA: Department of Statistics, University of Washington.

Fraley, C., & Raftery, A. E. (2003). Enhanced software for model-based clustering, discriminant analysis, and density estimation: MCLUST. *Journal of Classification*, *20*, 263–286.

Gorman-Smith, D., Tolan, P. H., Loeber, R., Henry, D. B. (1998). Relation of family problems to patterns of delinquent involvement among urban youth. *Journal of Abnormal Child Psychology*, *26*, 319–333.

Grunbaum, J. A., Kann, L., Kinchen, S., Ross, J., Hawkins, J., Lowry, R., et al. (2004). Youth risk behavior surveillance—United States, 2003. *Morbidity and Mortality Weekly Report*, *53*, 1–95.

Guo, J., Chung, I. J., Hill, K. G., Hawkins, J. D., Catalano, R. F., & Abbott, R. D. (2002). Developmental relationships between adolescent substance use and risky sexual behavior in young adulthood. *Journal of Adolescent Health*, *31*, 354–362.

Hand, D. J., & Bolton, R. J. (2004). Pattern discovery and detection: A unified statistical methodology. *Journal of Applied Statistics*, *31*, 885–924.

Hardin, J., & Rocke, D. M. (2002). Outlier detection in the multiple cluster setting using the minimum covariance determinant estimator. *Computational Statistics and Data Analysis*, *44*, 625–638.

Haselager, G. J. T., Cillessen, A. H. N., Van Lieshout, C. F. M., Riksen-Walraven, J. M. A., & Hartup, W. W. (2002). Heterogeneity among peer-rejected boys across middle childhood: Developmental pathways of social behavior. *Developmental Psychology*, *38*, 446–456.

Hill, K. G., White, H. R., Chung, I.-J., Hawkins, D., & Catalano, R. F. (2000). Early adult outcomes of adolescent heavy drinking: Person- and variable-centered analyses of heavy drinking trajectories. *Alcoholism: Clinical and Experimental Research*, *24*, 892–901.

Hipp, J. R., & Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychological Methods*, *11*, 36–53.

Hix-Small, H., Duncan, T. E., Duncan, S. C., & Okut, H. (2004). A multivariate associative finite growth mixture modeling approach examining adolescent alcohol and marijuana use. *Journal of Psychopathology and Behavioral Assessment*, *26*, 255–270.

Insightful Corporation (1988–2006). S-PLUS 7 for Windows. Seattle, WA: Author.

Jedidi, K., Jagpal, H. S., & DeSarbo, W. S. (1997). STEMM: A general finite mixture structural equation model. *Journal of Classification*, *14*, 23–50.

Jessor, R., van den Bos, J., Vanderryn, J., Costa, F. M., & Turbin, M. S. (1995). Protective factors in adolescent problem behavior: Moderator effects and developmental change. *Developmental Psychology*, *31*, 923–933.

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

Jester, J. M., Nigg, J. T., Adams, K., Fitzgerald, H. E., Puttler, L. I., Wong, M. M., et al. (2005). Inattention/hyperactivity and aggression from early childhood to adolescence: Heterogeneity of trajectories and differential influence of family environment characteristics. *Development and Psychopathology*, *17*, 99–125.

Johnston, L. D., O'Malley, P. M., & Bachman, J. G. (2006). *Monitoring the Future national survey results on drug use, 1975–2005*. Bethesda, MD: National Institute on Drug Abuse.

Kandel, D. (1975). Stages in adolescent involvement in drug use. *Science*, *190*, 912–914.

Kandel, D. B. (2002). Examining the gateway hypothesis: Stages and pathways of drug involvement (pp. 3–15). In D. B. Kandel (Ed.), *Stages and pathways of drug involvement.* Cambridge: Cambridge University Press.

Kaplow, J. B., Curran, P. J., Dodge, K. A., & the Conduct Problems Prevention Research Group (2002). Child, parent, and peer predictors of early-onset substance use: A multisite longitudinal study. *Journal of Abnormal Child Psychology*, *30*, 199–216.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., et al. (1994). Lifetime and 12-month prevalence of *DSM-III-R* psychiatric disorders in the United States. *Archives of General Psychiatry*, *51*, 8–19.

Kim, H. K., Capaldi, D. M., & Stoolmiller, M. (2003). Depressive symptoms across adolescence and young adulthood in men: Predictions from parental and contextual risk factors. *Development and Psychopathology*, *15*, 469–495.

Kim, J., & Cicchetti, D. (2006). Longitudinal trajectories of self-system processes and depressive symptoms among maltreated and nonmaltreated children. *Child Development*, *77*, 624–639.

Leisch, F. (2004). FlexMix: A general framework for finite mixture models and latent class regression in R. *Journal of Statistical Software*, *11*, 1–18.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of American Statistical Association*, *83*, 1198–1202.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.

Loeber, R., & Southamer-Loeber, M. (1998). Development of juvenile aggression and violence. *American Psychologist*, *53*, 242–259.

Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, *10*, 21–39.

Magnusson, D. (2000). The individual as the organizing principle in psychological inquiry: A holistic approach. In L. R. Bergman, R. B. Cairns, L.-G. Nilsson, & L. Nystedt (Eds.), *Developmental science and the holistic approach* (pp. 33–48). Mahwah, NJ: Erlbaum.

Masten, A. S., Burt, K. B., Roisman, G. I., Obradovic, J., Long, J. D., & Tellegen, A. (2004). Resources and resilience in the transition to adulthood: Continuity and change. *Development and Psychopathology*, *16*, 1071–1094.

McLachlan, G. J., & Chang, S. U. (2004). Mixture modeling for cluster analysis. *Statistical Methods in Medical Research*, *13*, 347–361.

McLachlan, G. J., Do, K., & Ambroise, C. (2004). *Analyzing microarray gene expression data.* New York: Wiley.

McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: Wiley.

Meilă, M. (2003). Comparing clusterings by the variation of information. In B. Schölkopf & M. K. Warmuth (Eds.), *Learning theory and kernel machines: Proceedings of the 16th Annual Conference on Computational Learning Theory* (pp. 173–187). New York: Springer.

Meilă, M. (2007). Comparing clusterings—An information based distance. *Journal of Multivariate Analysis*, *98*, 873–895.

Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*, *58*, 525–543.

Miller, P., & Plant, M. (2002). Heavy cannabis use among UK teenagers: An exploration. *Drug and Alcohol Dependence*, *65*, 235–242.

Miller-Tutzauer, C., Leonard, K. E., & Windle, M. (1991). Marriage and alcohol use: A longitudinal study of "maturing out." *Journal of Studies on Alcohol*, *52*, 434–440.

Moffitt, T. E. (1993). Adolescence-limited and life-course persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674–701.

Monti, P. M., Miranda, R., Jr., Nixon, K., Sher, K. J., Swartzwelder, S., Tapert, S. F., et al. (2005). Adolescence: Booze, brain, and behavior. *Alcoholism: Clinical and Experimental Research*, *29*, 207–220.

Muthén, B. O. (2003). Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran (2003). *Psychological Methods*, *8*, 369–377.

Muthén, B. O. (2006). *Latent variable hybrids: Overview of old and new models*. Presented at the University of Maryland Center for Integrated Latent Variable Research conference, "Mixture Models in Latent Variable Research," May 18–19, 2006.

Muthén, L. K., & Muthén, B. O. (1998–2006). *Mplus user's guide* (4th ed.). Los Angeles: Author.

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

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

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

Nagin, D. S., Farrington, D. P., & Moffitt, T. E. (1995). Life-course trajectories of different types of offenders. *Criminology*, *33*, 111–139.

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

Needham, R. M. (1965). Computer methods for classification and grouping. In I. Hymes (Ed.), *The use of computers in anthropology* (Vol. 11, pp. 91–110). The Hague: Mouton.

Newcomb, M. D., & Bentler, P. M. (1988). *Consequences of adolescent drug use: Impact on the lives of young adults.* Newbury Park, CA: Sage.

Orlando, M., Tucker, J. S., Ellickson, P. L., & Klein, D. J. (2004). Developmental trajectories of cigarette smoking and their correlates from early adolescence to young adulthood. *Journal of Consulting and Clinical Psychology*, *72*, 400–410.

Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology*, *12*, 427–441.

Potter, C. C., & Jenson, J. M. (2003). Cluster profiles of multiple problem youth. *Criminal Justice and Behavior*, *30*, 230–250.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401.

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

Raffaelli, M., & Crockett, L. J. (2003). Sexual risk taking in adolescence: The role of self-regulation and attraction to risk. *Developmental Psychology*, *39*, 1036–1046.

Reyna, V. F., & Farley, F. (2006). Risk and rationality in adolescent decision making. *Psychological Science in the Public Interest*, *7*, 1–44.

Rindskopf, D. (2003). Mixture or homogeneous? Comment on Bauer and Curran (2003). *Psychological Methods*, *8*, 364–368.

Sadeh, A., Gruber, R., & Raviv, A. (2002). Sleep, neurobehavioral functioning, and behavior problems in school-age children. *Child Development*, *73*, 405–417.

Sampson, R. J., Laub, J. H., & Wimer, C. (2006). Does marriage reduce crime? A counterfactual approach to within-individual causal effects. *Criminology*, *44*, 465–508.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. New York: Chapman & Hall.

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

Schulenberg, J., Sameroff, A. J., & Cicchetti, D. (2004). The transition to adulthood as a critical juncture in the course of psychopathology and mental health. *Development and Psychopathology*, *16*, 799–806.

Shiffman, S., Kassel, J. Paty, J., Gnys, P., & Zettler-Segal, M. (1994). Smoking typology profiles of chippers and regular smokers. *Journal of Substance Abuse*, *6*, 21–35.

SPSS Inc. (1989–2004). SPSS 13.0 for Windows. Chicago: Author.

Steinley, D. (2003). Local optima in K-means clustering: What you don't know may hurt you. *Psychological Methods*, *8*, 294–304.

Stoolmiller, M., Kim, H. K., & Capaldi, D. M. (2005). The course of depressive symptoms in men from early adolescence to young adulthood: Identifying latent trajectories and early predictors. *Journal of Abnormal Psychology*, *114*, 331–345.

Tarter, R. E., & Vanyukov, M. (1994). Alcoholism: A developmental disorder. *Journal of Consulting and Clinical Psychology*, *62*, 1096–1107.

Tonidandel, S., & Overall, J. E. (2004). Determining the number of clusters by sampling with replacement. *Psychological Methods*, *9*, 238–249.

Tubman, J. G., Vicary, J. R., von Eye, A., & Lerner, J. V. (1990). Longitudinal substance use and adult adjustment. *Journal of Substance Abuse*, *2*, 317–334.

Tubman, J. G., & Windle, M. (1995). Continuity of difficult temperament in adolescence: Relations with depression, life events, family support, and substance use across a one-year period. *Journal of Youth and Adolescence*, *24*, 133–153.

Tucker, J. S., Ellickson, P. L., & Klein, D. J. (2002). Smoking cessation during the transition from adolescence to young adulthood. *Nicotine and Tobacco Research*, *4*, 321–332.

Tucker, J. S., Orlando, M., & Ellickson, P. L. (2003). Patterns and correlates of binge drinking trajectories from early adolescence to young adulthood. *Health Psychology*, *22*, 79–87.

van der Kloot, W. A., Spaans, A. M. J., & Heiser, W. J. (2005). Instability of hierarchical cluster analysis due to input order of the data: The PermuCLUSTER solution. *Psychological Methods*, *10*, 468–476.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). New York: Springer.

von Eye, A. (2002). *Configural frequency analysis: Methods, models, and applications*. Mahwah, NJ: Erlbaum.

von Eye, A., & Bergman, L. R. (2003). Research strategies in developmental psychopathology: Dimensional identity and the person-oriented approach. *Development and Psychopathology*, *15*, 553–580.

von Eye, A. & Mun, E. Y. (2005). *Analyzing rater agreement*. Mahwah, NJ: Erlbaum.

von Eye, A., Mun, E. Y., & Indurkhya, A. (2004). Typifying developmental trajectories—A decision making perspective. *Psychology Science*, *1*, 65–98.

White, H. R., Pandina, R. J., & Chen, P-H. (2002). Developmental trajectories of cigarette use from early adolescence into young adulthood. *Drug and Alcohol Dependence*, *65*, 167–178.

Wiesner, M., & Capaldi, D. (2003). Relations of childhood and adolescent factors to offending trajectories of young men. *Journal of Research in Crime and Delinquency*, *40*, 231–262.

Wiesner, M., & Kim, H. K., & Capaldi, D. (2005). Developmental trajectories of offending: Validation and prediction to young adult alcohol use, drug use, and depressive symptoms. *Development and Psychopathology*, *17*, 1–20.

Wills, T. A., & Cleary, S. D. (1997). The validity of self-reports of smoking: Analyses by race/ethnicity in a school sample of urban adolescents. *American Journal of Public Health*, *87*, 56–61.

Wilsnack, S. C., Plaud, J. J., Wilsnack, R. W., & Klassen, A. D. (1997). Sexuality, gender, and alcohol use. In R. W. Wilsnack & S. C. Wilsnack (Eds.), *Gender and alcohol* (pp. 250–287). New Brunswick, NJ: Rutgers University, Center of Alcohol Studies.

Windle, M. (1989). Predicting temperament-mental health relationships: A covariance structure latent variable analysis. *Journal of Research in Personality*, *23*, 118–144.

Windle, M. (1991). The difficult temperament in adolescence: Associations with substance use, family support, and problem behaviors. *Journal of Clinical Psychology*, *47*, 310–315.

Windle, M. (1992). The Revised Dimensions of Temperament Survey (DOTS-R): Simultaneous group confirmatory factor analysis for adolescent gender groups. *Psychological Assessment*, *4*, 228–234.

Windle, M. (1996). An alcohol involvement typology for adolescents: Convergent validity and longitudinal stability. *Journal of Studies on Alcohol*, *57*, 627–637.

Windle, M., & Lerner, R. M. (1986). Reassessing the dimensions of temperamental individuality across the lifespan: The Revised Dimensions of Temperament Survey (DOTS-R). *Journal of Adolescent Research*, *1*, 213–230.

Windle, M., & Mun, E. Y. (2006). *The course of depressive symptoms from ages 14 through 30: Growth mixture modeling and young adulthood outcomes.* Unpublished manuscript.

Windle, M., Mun, E. Y., & Windle, R. C. (2005). Adolescent-to-young adulthood heavy drinking trajectories and their prospective predictors. *Journal of Studies on Alcohol*, *66*, 313–322.

Windle, M., & Wiesner, M. (2004). Trajectories of marijuana use from adolescence to young adulthood: Predictors and outcomes. *Development and Psychopathology*, *16*, 1007–1027.

Winters, K. C., Stinchfield, R. D., Henly, G. A., & Schwartz, R. H. (1991). Validity of adolescent self-reports of alcohol and other drug involvement. *The International Journal of Addictions*, *25*, 1379–1395.

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

Wong, M. M., Brower, K. J., Fitzgerald, H. E., & Zucker, R. A. (2004). Sleep problems in early childhood and early onset of alcohol and other drug use in adolescence. *Alcoholism: Clinical and Experimental Research*, *28*, 578–587.
World Health Organisation. (1997). *Composite International Diagnostic Interview, Version 2.0*. Geneva: Author.
Yeung, K. Y., Fraley, C., Murua, A., Raftery, A. E., & Ruzzo, W. L. (2001). Model-based clustering and data transformations for gene expression data. *Bioinformatics*, *17*, 977–987.
Zucker, R. A., Ellis, D. A., Bingham, C. R., & Fitzgerald, H. E. (1996). The development of alcoholic subtypes. *Alcohol Health & Research World*, *20*, 46–54.
Zucker, R. A., Fitzgerald, H. E., & Moses, H. D. (1995). Emergence of alcohol problems and the severe alcoholisms: A developmental perspective on etiological theory and life course trajectory. In D. Cicchetti & D. Cohen (Eds.), *Manual of developmental psychopathology* (Vol., 2, pp. 677–711). New York: Wiley.

Appendix A

Appendix B: Geometric Properties of Covariance Matrix

Model-based clustering utilizes finite mixture models with the underlying idea that the observed data set is originated from a mixture of several subpopulations, and that the group membership is a priori unknown. Although the parametric densities can be various, multivariate normal densities are often used to model the components. The observed covariance matrix is decomposed into the following general form

$$\sum_k = \lambda_k D_k A_k D_k^T,$$

where λ_k parameterizes the *volume* of the kth cluster, D_k indicates the *orientation* of that cluster, and A_k represents the *shape* of that cluster (Banfield & Raftery, 1993, Fraley & Raftery, 1998, 1999, 2002a, 2002b, 2003). The subscript k indicates that each component can have different volume, shape, and orientation. The 10 models tested simultaneously using the MCLUST software package reflect different spectral decompositions: A, spherical, equal volume, equal shape; B, spherical, unequal volume, equal shape; C, diagonal, equal volume, equal shape; D, diagonal, unequal volume, equal shape; E, diagonal, equal volume, varying shape; F, diagonal, unequal volume, varying shape; G, ellipsoidal, equal volume, equal shape, equal orientation; H, ellipsoidal, equal volume, equal shape, varying orientation; I, ellipsoidal, unequal volume, equal shape, varying orientation; J, ellipsoidal, unequal volume, varying shape, varying orientation.

Model A is the most simplistic model and Model J is the most complex model. Model A, in

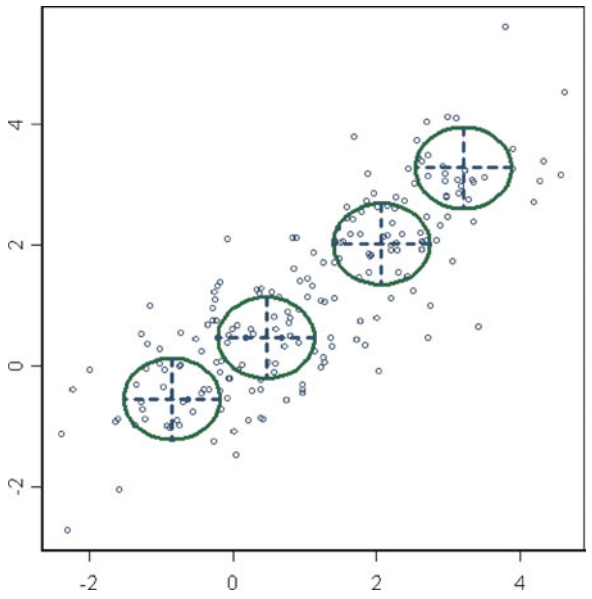


Figure A.1. A graphic illustration of volume, shape, and orientation in two-dimensional space. An example of Model A, spherical, equal volume, and equal shape in two-dimensional space. It was drawn using two normally distributed random variables ($N = 200$). The solid lines indicate spherical clusters and encircled dotted lines indicate orientation. [A color version of this figure can be viewed online at www.journals.cambridge.org]

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

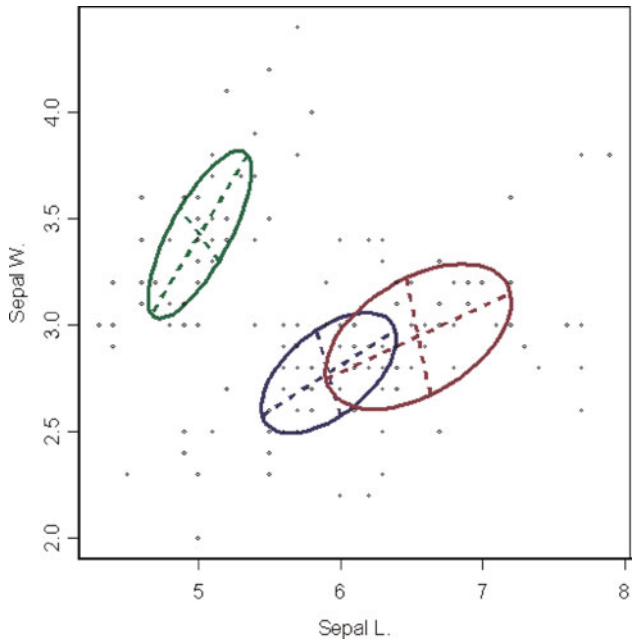


Figure A.2. An example of Model I is ellipsoidal, unequal volume, equal shape, and varying orientation in two-dimensional space. These data came from the classic Fisher's Iris data set featured in Everitt and Hand (1981) and Fraley and Raftery (1998, 1999, 2002a, 2002b, 2003). Two ellipsoids are different in their respective volume and orientation. [A color version of this figure can be viewed online at www.journals.cambridge.org]

particular, has been shown to be very closely related to the *k*-means clustering algorithm (Celeux & Govaert, 1992; Everitt et al., 2001; Yeung et al., 2001). Note that a trade-off generally exists between the number of clusters and the number of estimated parameters in the sense that the more simplistic models may require a larger number of clusters to adequately fit the data, whereas a small number of clusters may be sufficient with increasingly more complex models to provide adequate fit to the data (Yeung et al., 2001; for an empirical demonstration applied to, for instance, latent class growth trajectory analysis, see Bauer & Curran, 2003, 2004).

Appendix C: Partial MCLUST Input File

```
library (mclust)

# model 1—the reported results #
cls <- EMclust (set1) # set1 is the name of the data file to be clustered #
cls
plot (cls)
clssum <- summary (cls, set1, warnSingular = T) # default values used #
clssum

clssum$bic
clssum$n
clssum$pro
clssum$mu
```

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press

```
clssum$sigma
clssum$decomp
class <- clssum$classification
uncer <- clssum$uncertainty

# model 2 - the relaxed tolerance level was used #
cls2 <- EMclust (set1, eps = 1.e-8, tol = 1.e-2, warnSingular = T)
cls2
plot (cls2)
clssum2 <- summary (cls2, set1)
clssum2

clssum2$bic
clssum2$n
clssum2$pro
clssum2$mu
clssum2$sigma
clssum2$decomp
class2 <- clssum2$classification
uncer2 <- clssum2$uncertainty

# two clustering results compared using the VI criterion #

compareClass(class, class2)
```

https://doi.org/10.1017/S095457940800014X Published online by Cambridge University Press