5

Developmental Trajectories of Cigarette Smoking from Adolescence to Adulthood

Laurie Chassin, Patrick J. Curran, Clark C. Presson, Steven J. Sherman, and R. J. Wirth

Patterns of smoking behavior over time exhibit substantial variation, and these patterns, in turn, hold the potential to inform possible phenotypes of tobacco use and dependence. This chapter examines the literature concerning developmental trajectories of cigarette smoking between adolescence and adulthood. It also presents an empirical example that examines these trajectories by using data from the Indiana University Smoking Survey. Specific areas discussed include

- *Past studies describing smoking trajectories and their antecedents and correlates*
- *Empirically identified trajectories of smoking from adolescence to adulthood*
- *Statistical approaches for potentially identifying unique trajectory classes from empirical data*
- *Results from a dynamic cluster analysis of tobacco use trajectories from the ages of 10 to 42 years in a sample initially recruited from a midwestern school system*

The data discussed in this chapter provide a framework for a three-chapter section in this monograph exploring aspects of cigarette smoking trajectories and their potential to inform further genetic research. In particular, these data point to several key areas for further study, linking these trajectories of smoking behavior to possible dynamic or developmental phenotypes of nicotine dependence.

The analyses described herein were supported by National Institute of Health grant DA013555. The authors thank Jon Macy and the families of the Indiana University Smoking Survey for their assistance in data collection and Denise Kruszewski for assistance with literature reviews.

Introduction

In attempting to identify phenotypes of cigarette smoking, is it potentially informative to consider heterogeneity in trajectories of smoking from adolescence to adulthood? Moreover, could these developmental patterns be useful for genetic analyses of smoking behavior? This chapter considers developmental trajectories of cigarette smoking as part of a broader section within this monograph (chapters 5–7) that examines tobacco use trajectories and their role in informing an understanding of phenotypes of smoking behavior. This chapter reviews the literature on trajectories of cigarette smoking from adolescence to adulthood, raises methodological issues, and provides an empirical example of these trajectories in relation to aspects of adult smoking phenotypes.

A central premise of this monograph is that the adult smoking phenotype used in the field of behavioral genetics is a crude and heterogeneous phenotype that is not ideal for genetic study (or for studies of etiological mechanisms more broadly). Efforts to refine this phenotype include distinguishing among adult smokers on features such as amount smoked, the presence or absence of particular dependence symptoms, failed cessation, maximum length of abstinence, and other factors, as well as on the basis of candidate endophenotypes (chapters 8 and 9). However, developmental considerations about the initiation, acquisition, and course of cigarette smoking from adolescence to adulthood may also contribute to an understanding of smoking phenotypes. For example, different stages of the smoking acquisition process (e.g., initial onset versus progression) differ in their heritability,[1,2] suggesting that different points along smoking trajectories may be influenced by different etiological factors (see chapter 3). Moreover, other features of developmental trajectories of smoking—including age of onset, speed of acceleration in smoking rate, variability versus persistence in smoking over time, and trajectories of associated use of other substances (e.g., alcohol, marijuana)—may all be useful in defining more homogeneous phenotypes for genetic analysis.[3,4]

These research questions require data about adolescent origins to inform the identification of smoking phenotypes in adulthood. However, the need to understand adult outcomes is not the only reason for an interest in the adolescent origins of smoking trajectories. For example, an understanding of heterogeneity in adolescent smoking phenotypes is itself important for understanding the etiology of adolescent smoking and for the development of preventive intervention targeted at adolescent age groups. Thus, for multiple reasons, it is useful to examine developmental aspects of smoking trajectories from their initial onset through adulthood to identify multiple pathways and the mechanisms underlying these pathways.

Because this is a chapter on trajectories of smoking behavior, and not tobacco dependence, it is also important to note questions that this chapter will not address. It does not cover theoretical issues in the conceptualization of dependence. This discussion is provided in chapter 3. It does not discuss the issues that are raised in modeling genetically informative samples; these are covered in chapter 6. Finally, trajectories of cigarette smoking do not occur in isolation but are associated with other forms of substance use. A consideration of smoking trajectories in combination with other substance use is provided in chapter 7, along with an empirical example of smoking and alcohol-use trajectories.

A Developmental Psychopathology Perspective: Studying Multiple Trajectories over Time

Although a comprehensive treatment of a developmental psychopathology perspective is beyond the scope of this chapter, it is important to embed the study of smoking trajectories within this broader conceptual and empirical context. Developmental psychopathology has been defined as the study of "the origins and course of individual patterns of behavioral maladaptation."[5(p18)] This definition places maladaptive behavior within the context of normal development, as well as in relation to the interplay between an individual's internal and external contexts in which neurobiological development and psychosocial experience are proposed to influence each other in reciprocal fashion.[6]

A developmental psychopathology perspective recognizes that different influences may determine the initiation of behavior as opposed to the maintenance of that behavior (as also hypothesized by stage models of cigarette smoking—for example, Mayhew and colleagues[7]—and the "watershed" model in chapter 3). Moreover, from this perspective, it is hypothesized that multiple, differing etiological pathways may lead ultimately to the same outcome *(equifinality).* In addition, it is hypothesized that any given risk factor may produce a range of diverse outcomes *(multifinality).*[8]

A developmental psychopathology perspective thus leads to the study of multiple trajectories of behavior over time so as to be able to identify and explain these diverse patterns of development. Such pathways are probabilistic in nature, rather than reified "groups," and it is possible for an individual to change trajectories in response to some change in risk and protective factors. Methods used to study multiple trajectories include both traditional variable-centered approaches (in which predictor variables are related to some outcome) and person-centered approaches (in which relatively homogeneous subgroups are identified and studied).

For the purposes of this monograph, an important question is whether thinking about multiple developmental pathways or trajectories of smoking over the life span can be useful for refining phenotypes of smoking to be used in genetic research. In other words, does it make sense to consider dynamic or developmental phenotypes of smoking? As noted by Pickles and Hill,[9] the notion of "pathways" or trajectories is not only a rich metaphor but also one that raises many questions and challenges (including whether such pathways are actually empirically identifiable). For more information, the reader is referred to extended discussions in Pickles and Hill.[9] Other references include Gottlieb and Willoughby[10] for a study considering the comorbidity of attention deficit hyperactivity disorder (ADHD) and smoking and Bergman and colleagues[11] for an extended discussion of person-centered research methods.

Adolescent Cigarette Smoking and Tobacco Dependence

Adolescence is the developmental period during which smoking (and other substance use) is most commonly initiated. In 2005, 9.3% of 8th graders, 14.9% of 10th graders, and 23.2% of 12th graders reported smoking in the past 30 days of a survey.[12] As with most forms of substance use, smoking rises to a peak prevalence in the age period of 18–25 years,[13,14] but unlike other forms of substance use, which decline after the mid-20s, smoking is more persistent.[13,14]

Perhaps this is true because smoking is legal, addictive, and does not immediately impair performance.

Compared with information about the prevalence of adolescent cigarette smoking, less is known about tobacco dependence in adolescence because this has been a later focus of research attention. Two commonly used measurement methods involve modifications of the Fagerström Tolerance Questionnaire (FTQ)[15] or the *Diagnostic and Statistical Manual of Mental Disorders (DSM)* criteria.[16] These two approaches produce only modest concordance in classifying adolescents, except at high levels of smoking of at least 16 cigarettes per day.[17]

Moreover, as has been shown for other forms of substance-use disorders, some caution is warranted in applying adult dependence measures to adolescents.[18] Whether or not the construct of dependence is similar for adults and adolescents, the functioning of particular items and criteria may differ. Additional research is needed on the measurement equivalence of tobacco dependence criteria over the life span (see chapter 6 for an empirical example of studying measurement equivalence over age).

Reported prevalence rates of tobacco dependence among adolescent smokers have varied widely depending on sampling, definitions of dependence, and definitions of adolescent smoking. A review by Colby and colleagues[19] reports rates between 20% for a proxy *DSM* diagnosis among 12- to 17-year-olds who smoked in the past year and 68% for a diagnosis based on FTQ criteria among 13- to 17-year-olds who smoked one pack or more per day. Kandel and colleagues[17] found that the majority of adolescent daily smokers met criteria for dependence (87% according to *DSM* criteria and 63% according to Fagerström criteria). In addition, there were no race/ethnicity differences in prevalence when smoking intake was controlled. Using the criteria of the *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10),* O'Loughlin and colleagues[20] found that 65.9% of 7th grade daily smokers were dependent. Thus, the majority of adolescent daily smokers appear to show tobacco dependence.

Whether or not they meet full diagnostic criteria, adolescents commonly report individual dependence symptoms. Kandel and colleagues[21] found that tolerance, impaired control, and withdrawal were the most common *DSM* Fourth Edition *(DSM-IV)* symptoms reported in a multiethnic sample of 6th–10th graders. Colby and colleagues[19] reported that most adolescent smokers retrospectively recalled at least one withdrawal symptom, either as part of a quit attempt or during periods when they were restricted from smoking. As with adult smokers, withdrawal symptoms are common.[22] Craving is the most commonly reported symptom (see also Rojas and colleagues[23]). However, reports of withdrawal symptoms are influenced by expectancies.[24] Prokhorov and colleagues[25] note that withdrawal symptoms such as irritability, depression, insomnia, and trouble in concentrating can be characteristic of adolescents in general, rather than specific to tobacco withdrawal (see Hughes[22] for a similar point concerning the adult epidemiological literature).

There are some data concerning the amount of time and exposure required for adolescent smokers to develop dependence. Using retrospective data from the National Comorbidity Survey, Breslau and colleagues[26] found that the onset of *DSM* Third Edition Revised nicotine dependence typically occurred at least one year after the onset of daily smoking. Gervais and colleagues[27] used a prospective study of 7th graders and reported that a 25% cumulative probability of attaining an *ICD* tobacco dependence diagnosis occurred at 41 months after the first puff of a cigarette.

In a longitudinal study of a multiethnic sample of 6th–10th graders, Kandel and colleagues[21] found that a 25% cumulative probability of attaining *DSM-IV* nicotine dependence occurred 23 months after tobacco use onset.

In contrast to the onset of the full dependence diagnosis, there is a shorter time to the first symptom of dependence. Kandel and colleagues[21] found that 25% of adolescent tobacco users experienced *DSM-IV* symptoms within five months of use. DiFranza and colleagues[28] found that among the 40% of ever-smoking adolescents who reported dependence symptoms, the median latency from monthly smoking to the onset of symptoms was 21 days for girls and 183 days for boys.

Some researchers have suggested that adolescents experience dependence symptoms not only quickly but also at very low levels of consumption.[21] For example, O'Loughlin and colleagues[20] reported that, among 7th graders, 19.4% of weekly smokers met *ICD-10* dependence criteria. Dierker and colleagues[29] found a similar prevalence (22%) by using *DSM-IV* diagnostic criteria with college freshmen who had smoked in the past week. When considering the presence of *any* symptom (rather than the full diagnostic criteria), even higher percentages of adolescents report symptoms at low levels of consumption. For example, among a sample of adolescents who smoked in the past three months, some reported symptoms even though they had smoked only once or twice.[20] Similarly, in a sample of 7th graders, DiFranza and colleagues[28] found that the median frequency of smoking at the onset of symptoms was only two cigarettes, one day per week. These studies are noteworthy for their multiple measures and frequent assessments but also have sampling limitations in terms of somewhat low or unreported participation rates (a common problem in these kinds of studies).

Few studies have compared adolescents and adults in terms of the relation between consumption and dependence. Kandel and Chen[30] examined a proxy measure of *DSM* dependence in the National Household Survey on Drug Abuse data and found that adolescents met dependence criteria at lower levels of smoking intake than did adults. They concluded that adolescents are particularly vulnerable to becoming tobacco dependent. However, these differences between adolescents and adults might reflect cohort rather than age effects. In a 2001 study, Breslau and colleagues[26] suggest that members of more recent age cohorts who adopted smoking, despite widespread public knowledge about its negative effects, may be particularly deviant in personality and represent a subsample of the population who have a high probability of developing into committed (and dependent) smokers. Because age and cohort are confounded in these studies, it is not possible to separate these two interpretations, and both effects might be operative. Moreover, other researchers have suggested that, rather than a "hardening" of smoking, changes in tobacco control and prevention and in the social acceptability of smoking have produced a "softening" of smoking. A 2006 study indicated that in recent decades, smokers have decreased the number of cigarettes that they smoke,[31] suggesting a substantial prevalence of light smoking. Whether or not smoking has been "hardening" or "softening" or both (i.e., becoming more bimodal), significant changes have been occurring in the United States both in tobacco control and prevention activities and in social norms about smoking. Given the powerful role of cultural and social norms, tobacco policies (e.g., taxation, youth access laws) and changes in the overall prevalence of smoking in the United States, it is important for research on adolescent smoking trajectories to consider historical and cohort effects on the findings.[32]

In general, the adolescent literature suggests a relation between increased levels of consumption and the probability of tobacco dependence. However, this relation is far from perfect, and some adolescents report symptoms at low levels of smoking.[29] Reports of dependence symptoms at very low levels of intake may have multiple interpretations including problems of measurement (e.g., dependence symptoms being nonspecific), the possibility that tobacco dependence is a multidimensional construct (with only some dimensions related to smoking rate), and the possibility that there are heterogeneous subgroups of adolescents who are dependent on tobacco, some at quite low levels of consumption.[19]

In addition to these interpretations, adolescents may show dependence symptoms at low levels of intake because they are particularly sensitive (compared with adults) to developing tobacco dependence. This interpretation is consistent with data from rodent models that compare adolescent versus adult exposure to nicotine. Levin and colleagues[33] found that female rats that began self-administration in adolescence showed significantly higher levels of self-administration than those who began in adulthood. This difference in rate of self-administration lasted into adulthood. Similarly, Adriani and colleagues[34] found that early-adolescent mice exhibited a spontaneous drive to oral nicotine (compared to water), which was not demonstrated by middle or late adolescents. In addition, early-adolescent nicotine exposure led to significant place conditioning, which was not seen for either late-adolescent or adult exposure.

Although the mechanisms underlying these age differences are not well understood, the unique effects of adolescent exposure compared with adult exposure to nicotine self-administration may be mediated through differential sensitivity to nicotine effects. Levin and colleagues[33] found that adolescent rats showed more hypothermia to nicotine than did adults at a given dose, although adult rats showed more activity reduction. Belluzzi and colleagues[35] replicated these activity reduction effects with male rats and suggested that locomotor inhibition may be an aversive effect of nicotine that is more pronounced in adults than in adolescents. However, age differences in nicotine response may be further modified by sex differences[36] and by exposure to nicotine in combination with alcohol.[37]

There may also be developmental differences in nicotine withdrawal. For example, O'Dell and colleagues[38] found that adolescent rats showed decreased sensitivity to withdrawal after chronic nicotine administration. They also suggested that nicotine exposure in adolescence may produce maximal reinforcing effects and minimal aversive effects, thus promoting rapid acceleration of self-administration. Moreover, in rats, adolescent nicotine exposure even for a brief period (intermittent doses with twice-daily injections) at low dosage levels (producing plasma concentrations as little as 1/10 of regular smokers) has been reported to produce nicotinic acetylcholine receptor upregulation in brain regions associated with nicotine dependence. This may make the adolescent brain particularly sensitive to nicotine effects.[39] Taken together, these data suggest that adolescence may be a unique period of biological vulnerability, during which nicotine exposure produces particularly rapid escalation in trajectories of nicotine consumption that may persist over time as well as increased vulnerability for dependence because of differential sensitivity to nicotine effects.

Of course, significant caution is required in generalizing from animal models to human adolescents, given differences in methods of administration, dosages, and contextual factors such as restrictions on access to tobacco, the social and peer context of

self-administration, and self-selection into smoking for human adolescents. In other words, there are likely to be substantial genetic, environmental, and gene-environment correlation and interactions that affect whether human adolescents begin to smoke as well as the timing of their smoking onset. Finally, even in the rodent model, the empirical evidence concerning age differences in nicotine response is not always clear-cut. The evidence has shown variation with gender, and with the task or paradigm that is used, as well as interactions with exposure to other substances.[36,37,40] More needs to be learned concerning the mechanisms underlying these age-dependent effects as well as their magnitude and persistence over time.

This hypothesis of age-dependent vulnerabilities to the effects of tobacco has also been proposed for other forms of substance use (see Spear and Varlinskaya[41] for a review of alcohol data). Moreover, in addition to hypotheses concerning adolescent-specific vulnerabilities to substance-use effects, other models that are based in the study of neurobiological development suggest that adolescents are particularly vulnerable to risk-taking behaviors more broadly (which would include the use of tobacco and other substances). These models note that adolescents manifest a biologically driven disjunction between increased levels of novelty/sensation seeking and the lack of fully developed self-regulation mechanisms (see Steinberg[42] for a review). These models view adolescents' special vulnerability for substance use as a function of broader developmental characteristics rather than specific substance-use effects. These alternative models do not necessarily oppose each other in terms of explaining adolescence as a particularly vulnerable period for the initiation of cigarette smoking and other substance use. However, the notion of differential vulnerability to nicotine effects further predicts that adolescent initiation (compared with later onsets) will be more likely to be accompanied by higher consumption levels, steeper acceleration and greater persistence over time, and the development of dependence at lower levels of consumption.

Biologically based approaches offer a different (although not necessarily competing) interpretation from psychosocial models, which have also sought to explain why substance use is typically initiated in adolescence and then shows declines in adulthood. Psychosocial models often conceptualize adolescence as a high-risk period for substance-use initiation, because of adolescents' drives for independence, adult status, and peer acceptance, all of which can be seemingly facilitated by the adoption of "problem behaviors" such as cigarette smoking and other substance use.[43] The transition to adulthood (ages 18–25 years) has been viewed as a time of increasing diversity in trajectories (see Schulenberg and colleagues[44] for a review). During these years, the relative homogeneity and external control within the high school environment is replaced with less external structure and increased choices, including choices of entry into multiple roles (e.g., student, worker, spouse, and parent). The transition to adulthood is marked by decreased regulation by parents, which might lead to escalations in substance use, but also by increased responsibility for the performance of adult roles, which might lead to decreased substance use.[45] Thus, both psychosocial and biologically based models provide differing, but not mutually exclusive, interpretations of age-related trajectories of tobacco and other substance use.

Studies of Genetic Influences on Adolescent Smoking

Surprisingly, given the interest in genetic influences on tobacco use, few high-quality studies have focused on the family

aggregation of adolescent tobacco use.[46] In fact, even basic data on the relation between adolescent smoking and parental smoking have been conflicting, with some reviews suggesting only very weak relations between parental smoking and adolescent smoking onset.[47] However, methodological limitations prevent firm conclusions. Many studies do not directly measure parental smoking (relying instead on adolescent reports), do not differentiate biological parents from adoptive or foster parents, do not adequately define the phenotype of parental smoking (often failing to differentiate between parental nonsmoking and former smoking),[48] and do not consider possible effects of prenatal exposure (see Avenevoli and Merikangas[46] for a detailed methodological critique of this literature).

Similarly, studies have not carefully defined the phenotype of adolescent smoking. Parental smoking may be more strongly related to adolescents' smoking rate or early age of onset rather than adolescents' global smoking status. Finally, studies often dismiss the role of parental smoking if its effects are eliminated when other variables (such as peer smoking) are entered into predictive models. However, such a pattern is consistent with the mediation of parental smoking effects by peer smoking and does not by itself argue that the relation between parental smoking and adolescent smoking is an artifact or unimportant.

The literature on twin and adoption studies of adolescent substance use (including tobacco use) has been reviewed by Hopfer and colleagues,[49] who reported that both genetic and environmental influences were important. Heritability of tobacco use in their review ranged across studies from 36% to 60% and was stronger for tobacco than for alcohol or marijuana use. For example, McGue and colleagues,[50] using the Minnesota Twin Family Study (MTFS) sample, reported that approximately 50% of the variance in smoking initiation (studied in late adolescence) was attributable to genetic influences, with no significant effects of shared environment. Han and colleagues[51] studied lifetime tobacco use in the MTFS and found stronger evidence of shared environment and a pattern that was suggestive of (but not significant for) gender differences. Finally, Rhee and colleagues[52] found (both for lifetime tobacco use and for "problem use," as defined by presence of a dependence symptom) that female adolescents showed greater heritability and weaker shared environment effects than did male adolescents. They reported that this gender difference was inconsistent with the adult literature.[52] A later analysis of retrospective life calendar data in a sample of male twin pairs[53] suggests that nicotine use in adolescence shows strong family environment effects that decline in importance through young adulthood whereas genetic effects were weak in adolescence and increased with age.

As noted elsewhere in this monograph (chapters 2 and 6), heritability estimates for adolescent smoking vary depending on which adolescent smoking phenotypes are selected for study. Koopmans and colleagues[54] found stronger heritabilities for amount of smoking than for initiation. This is consistent with conclusions of other reviews (see chapter 2 and Rende and Waldman[55]) that environmental influences are more important in determining adolescents' initial tobacco exposure, whereas genetic influences are more important for determining reactions to that exposure. This pattern likely reflects the heterogeneity of adolescent smoking initiation. For example, some adolescent initiation of smoking will not progress past low levels of experimentation.[56] This developmentally limited experimentation may be only weakly related to genetic influence.

This supports the use of a developmental trajectory approach to identifying

phenotypes in which developmentally limited experimentation needs to be distinguished from other forms of adolescent smoking that may start earlier, escalate steeply over time, and persist over long periods. However, exceptions to these findings should also be noted. For example, McGue and colleagues[50] found similar results when studying smoking initiation and nicotine dependence in late adolescence. Maes and colleagues,[57] when studying an adult sample, found high heritabilities and an overlapping contribution of genetic factors for tobacco initiation, persistence, and nicotine dependence.

The adolescent literature (compared with the adult literature) also shows more shared environment effects on tobacco use.[51] White and colleagues,[58] in a longitudinal study using the Australian Twin Registry, found that common environmental influences were the most important factors influencing adolescent smoking, although for older adolescents and young adults, genetic factors were also important. Rende and colleagues[59] used National Longitudinal Study of Adolescent Health (Add Health) data to model current smoking, by using twins, full siblings, and half siblings, and found both significant heritability and significant shared environment effects. It is noteworthy that the authors found significant shared environment effects on high levels of smoking frequency, a more "severe" phenotype than smoking initiation.

Shared environment effects may reflect multiple influences including (but not limited to) general parenting behaviors, such as monitoring, support, and control, and parental socialization about smoking such as home smoking restrictions and smoke-free homes.[60] There have been some attempts to explain shared environment effects as due to the effects of parental smoking (which might reflect modeling mechanisms, greater access to cigarettes, greater exposure to secondhand smoke, greater exposure to tobacco promotional advertising, and/or more permissive attitudes of parents toward their adolescents' smoking). However, these findings are conflicting. Boomsma and colleagues[61] found that the association between parental smoking and adolescent smoking was due to genetic factors. They suggested that common environmental influences might reflect parenting behaviors and family environment, rather than parental smoking. For example, common environment effects might include parents' home smoking restrictions. However, White and colleagues[58] found that controlling for parental smoking reduced the common environmental effect. Differences between the findings of these two studies may be methodological. White and colleagues[58] used adolescent reports of parental smoking, whereas Boomsma and colleagues[61] used parent reports. Interestingly, White and colleagues[58] also found that peer smoking reduced the genetic effect. They suggested that genetic influences in adolescent smoking may act indirectly by influencing adolescents' choice of friends (although these findings weakened by late adolescence and young adulthood). These attempts to examine the relation of peer and parental smoking to the genetic and shared environment influences on adolescent smoking illustrate the importance of the gene-environment covariation in understanding the smoking acquisition process. That is, parental genotype and parental smoking likely covary with a wide variety of social environment factors including general parenting; parents' attitudes, values, and rules about their adolescents' smoking; adolescents' exposure to secondhand smoke (or conversely to smoke-free homes); and even adolescents' exposure to tobacco industry promotional items and advertising.

In addition to gene-environment covariation, there are likely to be important influences of gene-environment interactions

on adolescent smoking, although few studies have examined such interactions. Timberlake and colleagues[62] examined the moderating effect of religiosity on smoking among late adolescents/emerging adults (aged 18–27 years) from Add Health. They found that self-rated religiousness weakened the magnitude of genetic influence on smoking initiation (defined as having smoked an entire cigarette). Similar findings have been reported for alcohol-use initiation.[54] Surprisingly, however, organized religious activity had no moderating effect. Timberlake and colleagues[62] hypothesized that organized religious activity may reflect parental pressure, whereas self-rated religiousness might constitute a more genuine reflection of the adolescent's religious commitment. In any case, these results serve to highlight the potential importance of larger cultural and social environmental factors in moderating the magnitude of genetic effects. Given the lack of studies that test gene-environment interaction in adolescent smoking, this is an important area for future investigation.

Finally, several studies have investigated the genetic underpinnings of the association among different forms of adolescent substance use. Young and colleagues[63] studied a large sample of adolescents aged 12–18 years and defined the "problem" use of a substance by the presence of at least one symptom. They found that the correlation among substance-use behaviors was driven by both common genetic and common environment factors (as well as special twin environment factors) but that the more "severe" phenotypes (i.e., problem substance use as opposed to substance use) showed stronger genetic correlations. Similarly, McGue and colleagues[64] found a highly heritable factor that accounted for the association among multiple forms of disinhibitory psychopathology (including substance use) among 17-year-old twins from the MTFS. Interestingly, earlier problem behavior (retrospectively assessed) was only weakly heritable, but the link between early problem behavior and later disinhibitory psychopathology was genetically mediated. From the perspective of developmental trajectories of smoking, these findings suggest that a trajectory of stable, persistent smoking over time might show more genetic influence than adolescent smoking does at any one given point in time.

Although the literature is quite small, there have also been a few molecular genetic studies of adolescent smoking, several of which have focused on the dopamine system. Audrain-McGovern and colleagues[65] followed 615 adolescents from 9th to 11th grade. They found no effects of *SLC6A3* (dopamine transporter genetic variants) but found that *DRD2* genetic variants were related to smoking progression. Specifically, adolescents with previous smoking experience were more likely to increase their smoking as a function of increased *DRD2*A1* alleles, and this effect was stronger for adolescents with depressive symptoms. However, there were no effects for adolescent never smokers, suggesting that different stages of smoking progression may have different determinants. The authors suggest that *DRD2* genetic variants may index greater reward value from smoking and that depressed adolescents (who lack other sources of positive experiences) may be particularly susceptible to such increased reward value.

However, as noted in chapter 2, *DRD2* has been associated with numerous addictive and affective disorders, so these effects are not specific to tobacco. Findings from an Australian longitudinal adolescent study[66,67] reported a protective effect for the **K4* allele of the *TH* gene, which is involved in dopamine synthesis. The authors hypothesize that this protective effect may work by increasing endogenous dopamine levels or by reducing the perceived reward of nicotine. However, their findings of a

protective effect were limited to a strict definition of nicotine dependence, which included high frequency (more than six days per week), high quantity (more than 10 cigarettes per day), shorter periods of abstention (smoking within one hour of waking), and stability (present at two waves of longitudinal measurement). These findings illustrate the potential importance of carefully defined phenotypes of smoking.

Laucht and colleagues[68] also studied the dopamine pathway and focused on the *DRD4* exon III polymorphism associated with novelty seeking. They studied a sample of 15-year-olds from the Mannheim Study of Risk Children, which followed infants who were oversampled for obstetrical and psychosocial risk. They found that the *DRD4*7*-repeat allele was associated with greater smoking among males (including lifetime smoking, amount smoked, and earlier onset), but not among females. Moreover, novelty seeking mediated this relation, suggesting that novelty seeking is a potential endophenotype, at least among adolescent males. For females, however, there was an interaction between the *DRD4*7*-repeat allele and the long allele of *5-HTTLPR.* Females who lacked the *DRD4*7*-repeat allele and who were homozygous for the long allele of *5-HTTLPR* smoked the most.[69] This demonstrates both the potential importance of gene-gene interaction and of gender differences in the mechanisms underlying adolescent smoking. However, given the possibility of chance findings with multiple tests, these interactions require confirmation in multiple studies.

The short-short genotype of *5-HTTLPR* has also been associated with increased smoking among adolescents. Gerra and colleagues[70] found this genotype to be associated with smoking and with early onset (before 15 years of age), heavy smoking (more than 10 cigarettes per day), as well as with novelty seeking, irritability, and underachievement. Finally, several studies focused on *CYP2A6,* which inactivates nicotine to cotinine. Studying adolescents from the longitudinal McGill University Study on the National History of Nicotine Dependence, O'Loughlin and colleagues[71] defined nicotine dependence as more than three *ICD* symptoms. They found that those smokers who became dependent were more likely to have 1 or 2 copies of the inactive *CYP2A6*2* or **4* variant. In contrast, Audrain-McGovern and colleagues[72] found that slower metabolizers (those with *CYP2A6* variants) had a significantly slower growth in tobacco dependence symptoms from grades 9 to 12. Differences in study findings may be due to many methodological differences between the studies including differing ages of measurement and definitions of dependence (categorical *ICD* diagnoses in O'Loughlin and colleagues[71] and changes over time in Fagerström-type symptomatology in Audrain-McGovern and colleagues[72]). A meta-analysis of adult studies[73] failed to find any relation between the *CYP2A6* genotype and smoking status or amount smoked, but the authors also noted limitations in these conclusions due to the generality and heterogeneity of these smoking phenotypes. In addition, association studies have high false-positive rates.

In general, findings from genetic studies of adolescent smoking echo the conclusions of Lessov and colleagues[74] from the adult data: although heritable factors are important, there are also important common environmental influences (particularly on smoking initiation) as well as complex interactions both among multiple genes and between genetic and environmental factors. Moreover, inconsistent findings across studies reflect multiple factors, including variation in study designs and ascertainment, and high rates of false positives, but also wide variation in smoking phenotypes that are studied. Finally, low participation rates in molecular genetic

studies can jeopardize both the internal and external validity of the conclusions (e.g., several studies[66,71] had participation rates of 55% or less).

Developmental Trajectories

Age of Smoking Onset

To this point, age of onset has been discussed mostly as it relates to age-dependent nicotine effects in animal studies along with associated alterations in neural systems that, in turn, make it more likely that tobacco use will escalate and persist over time. These studies assign a causal role to adolescent exposure in producing steep acceleration in tobacco use. In other words, adolescence is thought to be a biological period of vulnerability during which exposure to tobacco increases risk for accelerating smoking by changing neural pathways.

However, age-dependent nicotine effects and associated changes in neural systems may not be the only reason that an early age of smoking onset is associated with acceleration and persistence. Rather, adolescents who begin tobacco use at particularly early ages may have unique characteristics, and their smoking may be maintained by different factors compared to those with late (after age 18) onset.[13,56] Perhaps early- and late-onset smoking represent two distinct subgroups (or what are called *transitional phenotypes* in chapter 3). For defining endophenotypes and phenotypes for the genetic study of tobacco dependence, this is an alternative model of adolescent exposure in which both early use and rapid acceleration are caused by a common underlying vulnerability (potentially one or more endophenotypes).

A similar hypothesis concerning age of onset has been advanced in the developmental psychopathology literature concerning antisocial behavior[75] in which childhood-onset, life-course-persistent delinquency is thought to be more strongly associated with inadequate parenting, neurocognitive problems, and violence. On the other hand, adolescent-onset delinquency is not characterized by these features; it is seen as more normative and more strongly linked to peer influence. Differences between childhood-onset, life-course-persistent delinquency and adolescent-onset delinquency are maintained into adulthood.[76] Moffitt's[75] developmental taxonomy of antisocial behavior illustrates the potential importance of developmental trajectories in understanding the heterogeneity that underlies adult phenotypes and when trying to create more homogeneous subgroups. Alternatively, differing ages of onset may simply reflect a continuum of severity with those having the highest levels of risk factors showing earliest entry,[43] or they may reflect differences in environmental opportunity and access to cigarettes. Thus, age of smoking onset is likely to have multiple determinants.

In terms of smoking trajectories, early smoking onset has been associated with steeper acceleration in smoking rate, greater persistence over time, and greater likelihood of developing dependence.[56,77,78] Note that this association does not mean that all early onset inevitably produces heavy smoking and dependence. As described later, some subgroups of early-onset smokers show experimental or developmentally limited patterns that do not persist over time.[56,79] Nevertheless, age of onset is related to greater risk for persistence and heavy use. Moreover, early-onset smoking has been reported to be more strongly related to parental smoking, whereas later onset (but still under 15 years of age) was related to peer (but not parental) smoking.[80] Similarly, a subgroup of smokers with early onset, steep acceleration, and persistence of heavy smoking over time also had the highest levels of smoking among biological parents.[56] These findings are consistent with reports that age of smoking onset

is heritable.[81,82] Moreover, Broms and colleagues[81] found that the same genetic influences on age of smoking onset did not account for the amount of smoking or smoking cessation, suggesting that age of onset may have distinct genetic underpinnings. Ling and colleagues[83] reported that a polymorphism of the dopamine transporter gene was associated with early onset of smoking and that it also magnified the relation between early onset and dependence.

Age of smoking onset has also been related to child psychopathology (see review by Upadhyaya and colleagues[84]). For example, ADHD has been associated with earlier initiation of regular smoking,[85,86] even after controlling for comorbidity.[87] A combination of conduct disorder and ADHD may be particularly predictive.[88] In contrast, anxiety disorders have been associated with delayed smoking onset,[89] although different forms of anxiety disorder may have different impacts. At least for alcohol use,[90] generalized anxiety disorder symptoms were associated with greater risk for initiation, whereas separation anxiety symptoms were associated with decreased risk.

Although some data suggest that early onset of smoking may constitute a unique phenotype in terms of showing different predictors than does late-onset smoking,[56,80] such a conclusion is still premature. First, few studies have contrasted early- and late-onset smoking, and much of the data concerning age of onset comes from retrospective studies, which might suffer from forward telescoping bias. For example, Johnson and Schultz,[91] using national interview data, found that older age at interview (within a given birth year) was less likely to produce a report of early smoking onset (see also Parra and colleagues[92]). Second, there is no specific age of onset that has been identified as "early," and this definition is likely to change within a social, cultural, and historical context. Third, it is unclear whether early-onset smoking is a unique phenotype distinct from other forms of early-onset substance use. For example, Yoon and colleagues[93] found that multiple forms of early substance use (including but not limited to smoking before 15 years of age) were linked to reduced P300 amplitude, which itself was highly heritable. They suggest that a failure in top-down control of behavior (as manifested by reduced P300 amplitude) may be one endophenotype that accounts for genetic influences on adolescent substance use more broadly (particularly for males). Thus, an early onset of smoking may also be associated with early onset of alcohol and other drug use, and these may be markers for an endophenotype associated with the "externalizing" spectrum broadly defined, not necessarily associated uniquely with tobacco use.

Rate of Acceleration from Initiation to Regular Smoking or Dependence

Another feature of developmental trajectories that might define a smoking phenotype is the speed at which an adolescent transitions from initial onset to regular smoking or dependence, or in other words, the slope of the growth curve of tobacco use.[94] Rate of acceleration itself may be a phenotype that reflects vulnerability to dependence on the basis of reactions to initial tobacco exposure as an endophenotype.[95] Retrospective data suggest that smoking is generally reported by study participants as an aversive experience initially; however, there is variability in how participants rate the experience. Participants who reported their initial experiences as relatively more positive (e.g., who report relaxation or a "buzz") and relatively less aversive were more likely to become smokers.[96–98] Thus, sensitivity to the pleasurable effects of nicotine and/or insensitivity to the negative effects are potential endophenotypes that might determine the speed of smoking acquisition (see chapter 8 for a detailed review of this

literature). However, the extant data are largely confined to retrospective self-reports of unknown dose amounts. More research is needed to determine whether individual differences in sensitivity to smoking's effects predict the degree of acceleration of smoking acquisition.

Importantly, a subsequent study empirically identified heterogeneity in trajectories of dependence symptoms over 36 months among novice smokers from a multiethnic sample of 6th–10th graders,[99] Using latent growth mixture modeling, the authors found that 47% of the sample developed no dependence symptoms. In contrast, 21% developed symptoms rapidly (within the first year), averaged more than two symptoms, and showed persistent symptoms, whereas 18% developed symptoms rapidly but averaged somewhat fewer symptoms and did not persist. Finally, 14% developed symptoms more slowly. Among those who developed symptoms rapidly, those who persisted showed significantly more parental tobacco dependence than did those who remitted. Moreover, compared to those whose symptoms developed more slowly, those who rapidly developed persisting symptoms showed more pleasant initial sensitivity and more conduct disorder symptoms. These data suggest that rapid acceleration and persistence of dependence symptoms may be informative phenotypes.

Empirically Identified Trajectories of Adolescent Smoking

Although adolescence is the typical age of smoking onset, there is substantial variability in age of onset, in steepness of acceleration, and in persistence over time. As noted above, this heterogeneity may reflect different phenotypes of smoking, which may be influenced by different underlying mechanisms and endophenotypes. Researchers have begun to empirically examine heterogeneity in the course of smoking over time in longitudinal studies. Studies are reviewed here that have examined tobacco use from early adolescence through either adolescence or adulthood (see table 5.1 for a summary of these studies). Only studies of tobacco use are in this review; studies of the joint trajectories of tobacco and alcohol or other drug use are reviewed in chapter 7.

For the youngest adolescent ages, Abroms and colleagues[100] examined trajectories of smoking from 6th to 9th grade by using growth mixture modeling with "stages" of smoking (ranging from no intention to smoke to smoking more than three cigarettes per month) as the outcome variable. Results showed five trajectory classes: never smokers, intenders, delayed escalators (who averaged monthly smoking by 9th grade), early experimenters, and early users (who averaged smoking three or more times per month by the end of 7th grade). Early users were an early-onset, sharply accelerating (albeit small) subgroup who surprisingly were distinguished from never smokers by decreased depression (unlike the other groups).

In another study, Vitaro and colleagues[80] distinguished among groups who started at 11, 12, and 13 years of age, and found that early onset was associated with antisocial behavior. Colder and colleagues[101] followed a similar age group (aged 12–16 years) by using growth mixture modeling. Similar to Abroms and colleagues,[100] they also identified an early-onset group (with onset between 12 and 13 years of age) that rapidly escalated to heavy smoking. A later-onset (after 14 years of age) group escalated less quickly and reached lower levels of smoking by 16 years of age. The remaining three groups were light smokers. Audrain-McGovern and colleagues[65] identified a never-smoker group, an experimenter

Table 5.1 Studies of Smoking Trajectories

Authors/ Year	Age	Gender	Ethnicity	Definition of smoking	Statistical analysis	Trajectory groups	Endophenotype findings	Comments
Chassin et al. 2000[56]	11–31 years	51% male	96% Non-Hispanic Caucasian	0 = not currently smoking 1 = up to monthly smoking 2 = up to weekly smoking 3 = weekly or more smoking, but only 10 or fewer cigarettes a day 4 = weekly or more smoking of 11–20 cigarettes per day 5 = weekly or more smoking of 20 or more cigarettes a day	Abstainer and erratic groups were defined a priori Latent class growth analysis mixture modeling	Abstainers (60%) Erratics (.02%) Early stables (12%) Late stables (16%) Quitters (5%) Experimenters (6%)	Abstainers reported lower levels of depression and lower levels of personality risk (including extraversion and conscientiousness). Early, stable group reported higher levels of depression than other groups.	
Colder et al. 2001[101]	12–16 years	52% female	79% Caucasian 18% African American 3% Asian Pacific/Asian Indian 0.3% Other	1 = used to smoke, but now I don't 2 = I've only tried a few puffs 3 = a few cigarettes per month or less 4 = less than a pack per week 5 = about a pack per week 6 = about one-half pack per day 7 = 1 pack per day or more	Piecewise latent growth mixture modeling	Early, rapid escalators Late, moderate escalators Late, slow escalators Stable, light escalators Stable puffers	None reported.	Percentages of sample in each trajectory group not reported.

Table 5.1 Studies of Smoking Trajectories *(continued)*

Authors/ Year	Age	Gender	Ethnicity	Definition of smoking	Statistical analysis	Trajectory groups	Endophenotype findings	Comments
Juon et al. 2002[102]	6–32 years	52.2% female	99% African American	Frequency and quantity of smoking	Multiple logistic regression	Nonsmokers (37%) Former smokers (12.9%) Current smokers/late adopters (25.6%) Current smokers/early adopters (24.1%)	Current smokers/early adopters were more likely to display antisocial behaviors in 1st grade and young adulthood than did nonsmokers and the other two smoker groups. Current smokers/early adopters were more likely to report both depression and drug problems than did nonsmokers.	
Soldz and Cui 2002[103]	6th–12th grade	55% female	79%–87% Caucasian 7%–9% African American 6%–9% Hispanic	0 = no cigarette use during the past month 1 = moderate use (≤40 cigarettes) during the past month 2 = heavy use (≥40 cigarettes) during past month	Generalized estimating equations approach	Nonsmokers (72.2%–93.5%) Light smokers (5.3%–8.9%) Heavy smokers (1.2%–20%)	Class truancy was related to smoking across grades.	
White et al. 2002[104]	12–31 years	50% female	92% Caucasian	Frequency of smoking in the past year and typical quantity per day	Latent growth mixture modeling Multinomial logistic regression analyses	Nonsmokers (39.6%) Occasional smokers (19%) Heavy smokers (1.2%–20%%)	Higher sensation seeking was linked with increased probability of belonging to a smoking trajectory group as well as heavier smoking over time. Delinquency and depression were not found to predict smoking group membership.	

Authors/ Year	Age	Gender	Ethnicity	Definition of smoking	Statistical analysis	Trajectory groups	Endophenotype findings	Comments
Audrain-McGovern et al. 2004[105]	14–18 years	52% female	63% Caucasian 12% Hispanic 11% Asian 8% African American 6% other	0 = never smoker 1 = puffer (never having smoked a whole cigarette) 2 = experimenter (<100 cigarettes ever) 3 = current smoker (smoked <20 days in last 30 days and >100 in lifetime) 4 = frequent (smoked ≥20 days in last 30 days and >100 in lifetime)	Latent class growth modeling	Early/fast adopters (8%) Late/slow adopters (24%) Experimenters (23%) Never smokers (45%)	Adolescents higher in novelty seeking and depressive symptoms were more likely to be early/ fast adopters and late/ slow adopters than never smokers or experimenters.	

Table 5.1 Studies of Smoking Trajectories *(continued)*

Authors/ Year	Age	Gender	Ethnicity	Definition of smoking	Statistical analysis	Trajectory groups	Endophenotype findings	Comments
Orlando et al. 2004[106]	13–23 years	52% male	67% Caucasian 10% African American 11% Hispanic 8% Asian 4% Other	0 = nonsmoker in past year 1 = <3 times in past year and <3 times in past month 2 = 3–10 times in past year and <3 times in past month 3 = 11± times in past year and <3 times in past month OR 3–5 times in past month 4 = 6± days in past month and <3 cigarettes per day 5 = 6± days in past month and about one-half pack per day 6 = 6± days in past month and about one-half pack per day 7 = 6± days in past month and 1 pack or more per day	Latent growth mixture modeling Multinomial logistic regression analyses	Nonsmokers (28%) Stable highs (6%) Early increasers (10%) Late increasers (10%) Decreasers (6%) Triers (40%)	At 13 and 15 years of age, nonsmokers reported fewer deviant behaviors and internalizing symptoms than did other smoking groups. At age 23, late increasers were more likely than triers and nonsmokers to have engaged in deviant behavior in the past year. Also, at 23 years of age, nonsmokers reported fewer internalizing symptoms than early increasers and late increasers.	

Authors/ Year	Age	Gender	Ethnicity	Definition of smoking	Statistical analysis	Trajectory groups	Endophenotype findings	Comments
Stanton et al. 2004[107]	9–18 years	Not reported	96% Caucasian 4% Maori/ Polynesian	Count of number of cigarettes smoked in past month	Latent growth mixture modeling Multinomial logistic regression analyses	Early, rapid escalators (11.4%) Late, rapid escalators (38.8%) Late, moderate escalators (14.3%) Late, slow escalators (11.4%) Stable puffers (12.7%) Late, slow escalators-puffers (11.4%)	Attention deficit disorder predicted early, rapid escalators, as well as late, slow escalators-puffers. Conduct disorder was an early predictor of smoking. Depression and behavior problems were a predictor of midadolescent smoking.	Sample comprised only New Zealanders.
Vitaro et al. 2004[80]	10–15 years	50.7% female	>90% Caucasian and French speaking	Number of cigarettes smoked during the week and during the day before data collection	Latent growth mixture modeling Logistic regression analyses	Never smokers (75.4%) 11–12-year-old starters (5.7%) 12–13-year-old starters (11.1%) 13–14-year-old starters (7.9%)	Membership in the 11–12-year-old starter group was associated with increased antisocial behaviors.	Antisocial behavior was analyzed as part of a composite general maladjustment score.
White et al. 2004[108]	10–25 years	100% male	42% Caucasian (C) 56% African American (AA) 2% Other/ Mixed	At screening, if ever tried tobacco, even a puff, and if so, what age (age of onset) At subsequent assessments, lifetime use, past year use, and number of cigarettes smoked per day	Latent class growth modeling Hierarchical logistic regression	Nonsmokers: C: 44.3% AA: 55.9% Occasional smokers: C: 23.7% AA: 27.3% Heavy smokers: C: 32% AA: 16.7%	None reported.	Separate trajectories were examined for Caucasians and African Americans.

Table 5.1 Studies of Smoking Trajectories *(continued)*

Authors/ Year	Age	Gender	Ethnicity	Definition of smoking	Statistical analysis	Trajectory groups	Endophenotype findings	Comments
Abroms et al. 2005[100]	6th–9th grade	Not reported	Not reported	0 = did not smoke in past 30 days or past 12 months and had no intention of smoking in high school 1 = did not smoke in past 30 days or 12 months but intended to smoke at least 1 or 2 times in high school 2 = smoked in the past 12 months but not in past 30 days 3 = smoked 1 to 2 times in the past 30 days 4 = smoked 3 or more times in the past 30 days	Latent growth mixture modeling Logistic regression to examine risk factors	Never smokers (41.2%) Intenders (33.5%) Delayed escalators (8.9%) Early experimenters (13.9%) Early users (2.5%)	Higher levels of deviance acceptance were associated with being an intender, an early experimenter, and an early user compared to a never smoker. Higher levels of depression decreased the likelihood of being an early user rather than a never smoker.	Ages not reported.
Karp et al. 2005[109]	12–17 years	64.8% female	100% Canadian	For 3-month intervals, number of days smoked each month and average number of cigarettes smoked per day each month	Individual growth curve modeling Latent class growth modeling	Low initial use, gradual increase (72.4%) Low initial use, rapid increase (11.1%) Low initial use, then increase in use, then decrease in use (10.8%) High-intensity initial use, then decrease in use (5.7%)	Depression, novelty seeking, and impulsivity did not seem to predict class membership.	All participants were novice smokers on entering the study.

Authors/ Year	Age	Gender	Ethnicity	Definition of smoking	Statistical analysis	Trajectory groups	Endophenotype findings	Comments
Brook et al. 2006[110]	14–26 years	51% female	51% African American (AA) 49% Puerto Rican (PR)	1 = none 2 = a few cigarettes or less per week 3 = 1–5 cigarettes per day 4 = about one-half pack per day 5 = about 1 pack per day 6 = more than 1 pack per day	Latent growth mixture modeling	Nonsmokers: AA: 56% PR: 36.5% Maturing out: AA: 6.9% PR: 12.9% Late starting: AA: 19.2% PR: 18.4% Early starting: AA: 20.3% PR: 25%	None reported.	
Riggs et al. 2007[78]	12–24 years	44% female	84% Non-Hispanic Caucasian	Amount smoked per week	Latent class growth analysis	Abstainers (47%) Low users (24%) Late, heavy users (16%) Early, heavy users (12%)	None reported.	Late is defined as after 15 years of age.
Maggi et al. 2007[111] Maggi 2008[112]	10–21 years	49.3% female		Separate models for probability of trying a cigarette and smoking frequency	Latent class growth analysis	Stable nonsmokers (48.4%) Late experimenters-nonsmokers (17.2% Experimenters-daily smokers (5.8%) Late experimenters-daily smokers (4.1%) Early experimenters-occasional smokers (10.5%) Late experimenters (13.9%)	None reported.	Late is defined as after ages 14–15 years. Sample from Canadian National Longitudinal Survey of Children and Youth.

Table 5.1 Studies of Smoking Trajectories *(continued)*

Authors/ Year	Age	Gender	Ethnicity	Definition of smoking	Statistical analysis	Trajectory groups	Endophenotype findings	Comments
Bernat et al. 2008[79]	12–19 years	49% female	85% Non-Hispanic Caucasian	Frequency of smoking (from never user to smoked most days)	Latent class growth analysis	Nonsmokers (55%) Triers (17%) Occasional users (10%) Early, established smokers (7%) Late, established smokers (7%) Decliners (4%)	None reported.	Community sampling using random digit dialing; 58.5% response rate.
Lessov-Schlagger et al. 2008[113]	13–24 years	49% female	92% Non-Hispanic Caucasian 2% Hispanic 3% Black 2% Native American	Quantity smoked in the past week	Latent class growth analysis	Experimenters (48.5%) Late increasers (16.3%) Early increasers (15.5%) Quitters (9.2%) Persistent (10.5%)	None reported.	Nonadopters excluded. Late is defined as after 18 years of age.

group, a late/slow adopter group, and a small (8%) early/fast adopter group. The two adopter groups showed elevated novelty seeking.

Wills and colleagues[114] followed students from 6th to 10th grade. Using cluster analysis, they distinguished stable nonsmokers and experimenters from three smoking groups that varied by age of onset: early (by 6th grade), intermediate (by 9th grade), and late (by 10th grade). Looking at psychosocial risk factors, their results suggested a continuum of risk in which early, intermediate, and late onset were ordered from highest to lowest risk. Moreover, they noted that scores on the risk factors changed over time, leading to an increased risk just before smoking onset.

Thus, rather than unique phenotypes, these findings are most consistent with a continuum model of time-varying risk with smoking onset resulting from an increase in psychosocial and/or genetic risk. However, because of the short developmental span of assessment, all of the onset groups that were identified in this study might still be considered "early" in developmental terms.

Finally, Karp and colleagues[109] focused on only a subsample of participants who had already begun to smoke and followed them from an average age of 13 to 17 years. Because all participants were smokers, no differentiation of age of onset can be made in this study. Over this short time span, most of the participants (72%) remained at low levels of smoking. However, the other 28% escalated their smoking, divided among rapid, low, and moderate groups, with the rapidly accelerating group representing 6% of the sample. Escalating youth were more likely to show symptoms of nicotine dependence, but other predictors did not differentiate among the groups (perhaps because all of these participants were already smokers and most remained at low levels of smoking during the study).

Several studies traced trajectories to slightly older ages of 18–21 years. Stanton and colleagues[107] modeled monthly smoking in the Dunedin study and found six trajectory classes. Again, there was an early, rapidly escalating group, but also a later (after 13 years of age), escalating group, and both ended at 18 years of age with high levels of smoking. Note that the definition of "late" escalation was still "early" in adolescence (13 years of age), which might account for both these groups' steep acceleration and high final smoking rates. Early, rapid escalators had higher conduct problems at 13 years of age than did all of the other groups and higher depression at 15 years of age than all but the late, moderate escalators. Early, rapid escalators also had higher attention deficit scores than the late, moderate escalators (but did not differ from the other groups).

Soldz and Cui[103] followed a large sample of students through 12th grade. Using cluster analysis, they identified nonsmokers, quitters, experimenters, early escalators, late escalators, and stable smokers. These findings are noteworthy for identifying a small quitting group that was absent in the other studies. Psychosocial protective factors (e.g., church attendance, time with fathers) were highest in nonsmokers, compared with continuous smokers, who showed an elevated risk profile. A similarly small group of decliners (4% of the sample) was identified by Bernat and colleagues[79] in a study of adolescents (aged 12–19 years). These "decliners" are intriguing because (unlike experimenter groups that are often identified) they show high levels of smoking frequency. Moreover, unlike the quitters in the Soldz and Cui study, the decliners showed elevated baseline risk profiles (compared to nonsmokers). However, the findings warrant replication because no measure of smoking quantity was considered and, unlike other studies, there was no "early stable" smoking group, so that very early onset was not present. Given the participation rate

(58.5%), perhaps the highest risk adolescents were absent from the sample. Finally, Maggi and colleagues[111] and Maggi[112] assessed trajectories for participants in the Canadian National Longitudinal Survey of Children and Youth. They found that a group of occasional smokers were not distinguishable from daily smokers until the age of 21 years.

Although there have been few studies, most researchers agree on the existence of a group with early onset, steep acceleration (or stably high levels of smoking from an early age), and high final levels of smoking. This group was also elevated, in general, in profiles of psychosocial risk. There was also substantial agreement about the existence of light-smoking groups, often associated with later onset. However, it is difficult to draw conclusions about a life course trajectory of smoking from studies that track the behavior only through high school. Age-related patterns of substance use more broadly (i.e., alcohol and illegal drugs) typically show adolescent initiation, peaks in emerging adulthood (ages 18–25 years), and later declines. In other words, some forms of substance use are developmentally limited and some persist.[115] Thus, to adequately map heterogeneity in smoking trajectories requires studies that span the early years of onset to adulthood to differentiate both early versus late onset and developmentally limited versus persistent use.

Very few studies have tracked smoking over such long age periods. White and colleagues[104] studied adolescents who were recruited through random telephone sampling and examined a quantity-frequency measure of cigarette use. They found three groups: a heavy smoking group that showed steep acceleration and heavy smoking, an occasional smoking group that "matured out" after 18 years of age, and a nonsmoking group. Female gender was associated with maturing out, and higher disinhibition was associated with regular smoking. However, adolescent risk factors did not differentiate the occasional and heavy groups. This lack of differentiation is likely due to the fact that the two smoking groups did not significantly differ in age of onset and because this study had a relatively small sample size for trajectory group differentiation.

Orlando and colleagues[106] tracked a large school-based sample from 13 to 23 years of age. They identified nonsmokers, stable high smokers, early increasers (increases between 13 and 14 years of age), late increasers (after 18 years of age), decreasers, and triers. Importantly, they found substantial onset after high school that has not often been recognized and cannot be found by studies that track participants only through high school. However, by 23 years of age, the different trajectory groups merged into two groups: low- and high-frequency smokers. An identical finding was reported by Lessov-Schlaggar and colleagues,[113] who studied a smaller sample from adolescence to 24 years of age. Moreover, these authors[113] found that smoking more than a few cigarettes per week in adolescence resulted in similar levels of nicotine dependence (as least as measured by the Fagerström Test for Nicotine Dependence and the Nicotine Dependence Syndrome Scale. These findings suggest that early onset is not very informative about adult smoking outcomes. In contrast, Riggs and colleagues[78] found that an early-onset-trajectory group showed more frequent weekly smoking and greater reported dependence than a later-onset-trajectory group at 24 years of age. Differences between these studies may reflect the fact that the "late" groups in the studies by Orlando and colleagues[106] and Lessov-Schlaggar and colleagues[113] increased smoking after 18 years of age whereas the "late" group in the study by Riggs and colleagues[78] increased smoking after 15 years of age. However, the question of whether heterogeneity in adolescent age and smoking course predicts adult smoking levels and adult nicotine dependence may be difficult to resolve when participants

are followed only until the ages of 23 or 24 years. It is possible that developmentally limited smoking had not yet declined and that further divergence would occur after the age range of the mid-20s.

Chassin and colleagues[56] studied a large school-based sample of ages 11 to 31 years. They removed two a priori groups (abstainers and a small group of erratic smokers who showed periods of relapse and remission) and empirically identified four groups described here. Early, stable smokers showed middle-school onset (ages 12–13 years) and averaged daily smoking by 15 years of age. They attained a high level of smoking (averaging more than one-half pack per day by 18 years of age) and stayed stable over the study. Late-onset smokers did not transition to weekly smoking until after 18 years of age and averaged less than one-half pack per day at their peak. An experimenter group never progressed past weekly smoking, and a quitter group declined after 21 years of age (similar to other forms of substance-use behavior). The early, stable and the erratic groups showed the riskiest profile on psychosocial factors. They were the least socially conventional, and their parents and peers were most likely to smoke. Interestingly, although both the early, stable and the experimenter groups showed early smoking onset, the experimenters were less likely to have parents who smoked (perhaps reflecting the heritability of smoking persistence). The late, stable group showed low levels of early risk factors and higher levels of college attendance. Their late onset might reflect transition out of the supervision provided in the parental home as well as some college environment factors. Surprisingly, a "chipper," or very light smoker, group did not emerge, perhaps reflecting its low prevalence in the population (or its correlation with late-onset smoking).

It is worth noting the similarities between the two studies. Both Orlando and colleagues[106] and Chassin and colleagues[56] identified an early-onset, rapidly accelerating group, and both identified substantial late (after high school) onset. Both studies also identified an experimental group that does not progress to regular smoking. However, the Chassin study did not find that the trajectories merged into two outcomes (high smoking and low smoking). Differences between the two studies might reflect the ethnic distribution of the samples. The Chassin sample was almost entirely non-Hispanic Caucasian. There were also possible cohort effects in that the two studies were five years apart in their baseline data collections. Finally, because the Chassin study tracked participants longer into adulthood, greater differentiation of smoking trajectories as a function of developmentally limited smoking might have been achieved.

One important issue in describing trajectories of smoking is potential ethnic differences. Four studies have focused on African American samples. In one study, Juon and colleagues[102] examined data from the Woodlawn study that followed inner city (predominantly low socioeconomic status) 1st graders to the age of 33 years. Rather than empirically identifying trajectories, they divided participants into nonsmokers, former smokers, late-onset smokers (after 18 years of age), and early-onset smokers. Compared to late-onset smokers, early-onset smokers were more likely to have been rated as aggressive by their 1st grade teachers, to have moved less often, to have had more lax parental supervision, and to have had more drug problems. Thus, early onset was generally associated with a profile of greater psychosocial risk in this study as it was with the other studies of ethnically diverse or Caucasian samples described above.

In another study, White and colleagues[108] modeled trajectories of the number of cigarettes smoked each day separately for African Americans and whites in the

Pittsburgh Youth Study, a prospective, longitudinal study of males that oversampled high-risk children and tracked them from ages 10 to 25 years. For each race, the three identified groups were nonsmokers, light smokers, and heavy smokers. However, there were differences in prevalence such that whites began smoking earlier and reached higher quantities of smoking than did African Americans. Similarly, Blitstein and colleagues[116] found that African Americans were more likely to show slow progression than rapid progression in their smoking. Brook and colleagues[110] modeled trajectories for African American and Puerto Rican adolescents aged 14–26 years and identified a group of nonsmokers, maturing-out smokers, late-starting smokers, and early-starting smokers. Although few studies have examined ethnicity, these findings converge in suggesting ethnic differences in onset and speed of progression that should be considered in describing smoking trajectories.

In short, among the few studies that have examined smoking trajectories from adolescence to adulthood, there is some convergence in terms of identifying an early-onset group that is either stably high or rapidly escalating, a later-onset group, and light-smoking groups that do not progress to regular smoking. However, there are important discrepancies, such as whether multiple trajectories do and do not diverge in their "final" endpoints and whether conceptually expected (but possibly low prevalence) groups such as stable light smokers (i.e., chippers) can be empirically identified. Studies have not empirically identified relapsing and remitting groups because (in all likelihood) these forms of growth are very complex to model. Most important, other than the fact that early-onset, sharply accelerating, and persistent groups are usually the most "at-risk" groups in terms of familial smoking, psychosocial risk, and measures of externalizing and internalizing problems, there is little evidence that any one trajectory group constitutes a unique phenotype characterized by specific endophenotypes. Indeed, some studies have not been able to empirically predict different smoking trajectories (other than differentiating between smoking and nonsmoking groups). However, very little work has been done to link the hypothesized preexposure endophenotypes (see table 5.1 and chapter 8) to trajectories that might constitute dynamic phenotypes of smoking. This is an important direction for future research. Finally, these studies have been limited to cigarette smoking and have not considered trajectories of dependence symptoms or of other forms of tobacco use.

Statistical Models for Evaluating Alternative Developmental Phenotypes of Smoking Behavior

To this point, multiple possible developmental phenotypes have been described that can be hypothesized to underlie the onset and maintenance of smoking behavior. For example, one hypothesized phenotype might be defined by age of onset, such that children who begin smoking at a very young age might constitute a unique subgroup. Another phenotype might be hypothesized to be reflected in the latency from onset to regular or heavy smoking. Yet another phenotype might be evidenced in smoking persistence over age, such that individuals who repeatedly fail to quit (or never attempt to quit) form an important homogeneous phenotypic subgroup. Finally, phenotypes might be represented by combinations of these features that are captured by developmental trajectories that show particular ages of onset, slopes of acceleration, peaks of smoking rates,

and persistence of smoking over time. These are just a few examples of many different observable behaviors or patterns of behavior over time that might indicate developmental phenotypes of smoking.

Given the heterogeneity in hypothesized phenotypes of interest, there is correspondingly no single statistical model that will allow for the optimal empirical evaluation of the viability of all of these phenotypes. Instead, a statistical model must be selected that most closely corresponds to the theoretical model of the phenotype.[117,118] For example, if a hypothesized phenotype postulates age of onset as a critical component, then survival analysis might be the ideal strategy. Alternatively, if a hypothesized phenotype is focused on time from onset to a transition to a category or stage of dependence, then latent transition analysis might be most appropriate. Because this chapter focuses on developmental trajectories (i.e., patterns of smoking behavior over development) as potential phenotypes, statistical methods for clustering developmental trajectories are emphasized. However, this is in no way meant to suggest that these methods are the only (or even necessarily the optimal) way to empirically evaluate smoking phenotypes. Rather, the goal is to provide a brief review of analytic methods for clustering growth trajectories, to summarize the relative strengths and weaknesses of these approaches, and to provide an empirical example of these techniques through the analysis of a large longitudinal study of smoking behavior.

Growth Curve Modeling

A core premise of the discussion so far is that the identification of phenotypes of smoking behavior might be well served by considering developmental trajectories of smoking-related behaviors and outcomes. This requires the collection and analysis of repeated-measures data, for which there are a plethora of analytic options. Traditional methods such as repeated-measures analysis of variance (ANOVA) and multivariate analysis of variance have long been known to be limited by strict underlying assumptions that are rarely met in practice.[119] Examples include the requirements of complete case data, equally spaced assessments, and normally distributed repeated measures. However, over the past two decades, a number of significant improvements in statistical models for repeated measures data have been introduced that overcome nearly all of these prior limitations. Because variations of these models arose within multiple literatures, there are various terms to which these are commonly referred. Examples include growth curve models, random coefficient models, latent curve models, and latent trajectory models. The evolution of growth curve models can be broadly traced to two modeling traditions: mixed (or multilevel) models and structural equation models (SEMs). Although there are a small number of important differences between models estimated within the mixed and SEM traditions, it is well known that these two approaches are isomorphic under a broad set of conditions.[120–124]

The standard growth curve model is based on the principle that a set of observed repeated measures drawn from a sample of individuals can be adequately reproduced as a function of an underlying, unobserved developmental trajectory.[119,125,126] The functional form of the trajectory might be linear, curvilinear, or characterized by some constant level that does not systematically change over time. Fitting a growth model to the observed statistics results in sample estimates of the parameters that define the underlying trajectory.[127–129] To accomplish this, one or more latent factors are used to define the functional form of the trajectory. The means of the latent factors are the *fixed effects* of the model and represent the estimated trajectory pooling across all individuals

in the sample. The variances of the latent factors are the *random effects* of the model and represent the degree of individual variability around the fixed effects. Finally, one or more covariates can be included to test for systematic differences in the parameters that govern the trajectories as a function of the covariates (e.g., to identify characteristics of individuals who begin at a higher level and increase at a steeper rate over time). There are a number of powerful extensions to the standard growth curve model, and these have been widely used in many studies of development and change (see Bollen and Curran[130] for a comprehensive review of these models).

Growth modeling methods might provide one important approach to the study of developmental phenotypes for cigarette smoking. For example, a particular genotype or putative endophenotype score could be used to predict the slope of a growth curve, such that the genotype or endophenotype predicted steeper acceleration in smoking trajectories (see Audrain-McGovern and colleagues[72] for an empirical example). Thus, if variation in acceleration of trajectories is a phenotype of interest, growth modeling provides a useful method for testing such phenotypes. Similarly, if early, heavy smoking is a phenotype of interest, then a particular genotype or endophenotype could be used to predict the intercept of the growth curve at a specific age of interest. Finally, the intercept and slope components of a developmental trajectory of smoking could themselves be used as predictors of a distal outcome such as nicotine dependence or some other measured characteristic. Taken together, growth curve models can be used to evaluate a number of important questions related to individual differences in developmental trajectories and their potential relation to smoking phenotypes and endophenotypes.

Despite the significant advantages growth curve models offer, they have limitations. Of greatest importance for this chapter, a strong underlying assumption of the standard growth model is that all individuals are sampled from a single population. This, in turn, implies that the population is governed by a single multivariate distribution of trajectory parameters from which all individuals in the sample are randomly drawn. For example, a population might be characterized by an overall trajectory that is defined by some mean intercept and linear slope. Further, there is a bivariate distribution of individual intercepts and individual slopes around these mean trajectory values. Particular individuals might by characterized by intercept values that are larger or smaller than others or by linear slopes that are steeper or less steep than others. Importantly, though, all individual trajectories are assumed to be drawn from this single bivariate population distribution. In this conceptualization, phenotypes are viewed as arrayed on a continuum of severity rather than as qualitatively different categories that represent discontinuous, separate populations.

The assumption that the sample is drawn from a single (or *homogeneous*) population is perfectly reasonable in many research applications. However, this assumption presents a substantial limitation if there is theoretical reason to believe that individuals within a single sample may have been drawn from one of several populations. Although the standard growth model can be expanded to explicitly incorporate multiple populations,[127,131] this is only possible if the grouping variable has been directly observed. Examples of observed grouping variables include gender, ethnicity, treatment condition, or observed genotype. However, significant challenges arise if the grouping variable has not been, or might never be, directly observed. A salient example of this is in the study of phenotypes. A single sample might consist of individuals drawn from one of several populations

(or phenotypic groups), yet these groups are inextricably mixed when using a standard single sample growth model. Further, the multiple group growth model is not a viable strategy because the phenotypic group membership is not directly observed. The challenge then becomes estimating the existence of these discrete groups on the basis of patterns of observed responses drawn from a single sample of individuals.[132,133] Fortunately, a broad class of analytic methods exists that allows for the clustering of trajectories into two or more discrete groups.

Clustering Trajectories

A long and rich history drawn from fields including statistics, biostatistics, psychometrics, econometrics, and criminology has focused on the complex task of seeking empirical evidence for the existence of unobserved groups. A wide array of techniques have been developed including cluster analysis, latent class analysis, latent profile analysis, finite mixture modeling, and growth mixture modeling. A comprehensive exploration of these techniques is, however, beyond the scope of this chapter; see Bauer and Curran,[134,135] Muthén,[136] and Nagin[137] for reviews.

The shared foundation of these analytic approaches is that an apparently homogenous sample of individuals is in actuality drawn from two or more discrete populations. Failure to properly model the mixing of multiple populations (or *population heterogeneity*) can lead to biased or invalid conclusions about the structural relations that exist within any of the multiple populations.[138] That is, fitting a model to the aggregation of multiple populations will likely not accurately reflect any one population, much less the full set. However, here lies the challenge: because population membership was not directly observed in the sample, the existence of these groups must be inferred on the basis of other measured characteristics of the sample. Key analytic tasks include the identification of the optimal number of groups, the proper specification of structural relations of the observed variables within each group, and the probabilistic assignment of each individual as a member of each of the multiple groups.

Traditional clustering techniques make assignments of individuals to groups based on ad hoc measures such as the sum of distances or the sum of squared Euclidean distances from the mean (or *centroid*) of each cluster. A prominent example of this is the classic method of *k*-means clustering.[139] This is an iterative approach in which variability is maximized between groups and minimized within groups. The *k*-means approach typically begins with the placement of *k*-points into the data space, where *k* represents the number of clusters. This set of points defines the initial group centroids. Next, each individual observation is assigned membership to the group that is defined by the closest centroid. Once all of the observations have been assigned to a cluster, a new set of centroids are computed on the basis of the individuals assigned to that group. This process is then repeated until the change in centroids from one iteration to the next is negligible. Although this is a straightforward and sometimes useful clustering procedure, the strong assumption of perfect reliability of measures, the sensitivity to outliers, and the ad hoc nature of class assignment has limited the use of this approach in practice.

Subsequent clustering methods incorporate likelihood-based approaches to estimation in which class extraction, class membership probabilities, and covariate relations are estimated simultaneously. Two closely related yet distinct approaches are increasingly used for clustering trajectories. The first approach does not incorporate random effects associated with the growth

process within class and is sometimes referred to as latent class growth analysis (LCGA) because of the shared similarities of this approach with traditional latent class analysis.[140,141] The second approach allows for the estimation of random effects within each class and is sometimes referred to as growth mixture modeling/models (GMM) because of the shared similarities of this approach with finite mixture modeling.[142] The historical lines of development that ultimately led to these methods span more than a century.[141–149] Drawing on these prior developments, two individuals can be predominantly credited with the latest methods of LCGA and GMM: Daniel Nagin and Bengt Muthén.

LCGA has primarily been developed by Daniel Nagin and his colleagues.[133,137,150,151] Like latent class analysis,[152] LCGA assumes conditional independence within class. As such, the within-class trajectory model is defined only by fixed effects. A mean trajectory is thus estimated for all individuals within a class, but there is no individual variability around these class-specific mean values. Posterior probabilities are estimated that reflect the probability that each individual belongs to each of the total number of classes. The effects of covariates can be included in LCGA, but these influences are limited to either predicting the set of class membership probabilities, or predicting class membership itself if each individual has been assigned to a single class on the basis of the posterior probabilities. LCGA has a variety of strengths, including the ability to directly model continuous, truncated continuous, and discrete repeated measures, and the expansion of the model to forming classes on the basis of simultaneous trajectories of two constructs over time.

GMM has primarily been developed by Bengt Muthén and his colleagues.[132,136,153,154] Like finite mixture modeling,[142] GMM assumes a multivariate normal distribution of the observations within each class. This, in turn, allows for the estimation of a growth model within each class that is characterized by both fixed and random effects. Like LCGA, an overall mean growth function is estimated within each class; however, there is also the ability to incorporate individual variability around these mean values. This, in turn, allows for the inclusion of covariates as predictors of class membership, of the growth process within class, or both. One of several close ties between the two techniques is that restricting the within-class variability to zero in GMM is an equivalent parameterization to LCGA. Further, if only a single class is extracted in GMM, this is equivalent to the standard single group latent curve model. And if class membership is directly observed, then GMM is equivalent to the standard multiple group latent curve model (e.g., Bollen and Curran,[130] Chapter 6). More standard growth models can thus be viewed as restricted parameterizations of the more general GMM.[154] Growth mixture models also offer a variety of strengths including the incorporation of continuous, nonnormal, and discrete outcomes, as well as a number of advantages provided by the general SEMs (e.g., multiple indicator latent factors and formal tests of mediation).

The LCGA and GMM approaches to estimating population heterogeneity in longitudinal trajectories represent an exciting advance that has salient implications for the study of phenotypes in that these analytic techniques offer a close correspondence to the theoretical model of different developmental trajectories as phenotypes of smoking behavior. However, because these methods are new, there is still much to learn about their performance under a variety of research conditions. Further, as with any advanced statistical technique, a number of challenges are encountered when fitting these models to longitudinal data in practice,[134,135,137,155–157] and both the advantages and challenges must be understood.

This chapter has already articulated many of the potential strengths associated with alternative techniques for clustering developmental trajectories over time. However, important issues must be considered when using these techniques. Because of space constraints, it is not possible to present a comprehensive discussion of all of these issues; see Bauer and Curran,[134,135,155] Muthén,[136,154,156] Nagin,[133,137] and Nagin and Tremblay[157] for more detailed explorations. Instead, several specific issues are explored that are particularly salient in the empirical study of smoking phenotypes.

Theoretical Distinctions between Discrete and Continuous Phenomena

Possibly one of the most challenging issues immediately encountered when considering the use of clustering methods to empirically study smoking phenotypes is fundamentally philosophical. There has been a centuries-old conflict over the very nature of taxa and continua and the intersections between the two. The primary issue at hand is whether phenotypes are characterized as discrete, continuous, or some intersection of the two.[158–163]

For example, consider two distinct phenotypes that are based upon a set of repeated observations taken on a sample of individuals over time. The first might be defined by individuals characterized by an early onset (i.e., *intercept*) and steep acceleration (i.e., *slope*) of use, and a second by individuals characterized by a later onset and less steep acceleration of use. Any given individual uniquely belongs to one phenotype or the other. In contrast, consider the same set of individuals observed on the same set of repeated observations. However, instead of distinctly belonging to one of two groups, the developmental growth process is characterized by a continuous bivariate distribution of intercepts and slopes; some individuals begin earlier and others later, and some increase more steeply and others less so. This latter situation is the same as that described earlier for the single-group growth curve model. Although from this single-group model some arbitrary cutoff might be defined that forms two groups, in reality the developmental process operates across a smooth (but not necessarily normal) continuum. This is not to say that some cutoff would not be of potential use (e.g., as in clinical diagnoses), but it would be theoretically invalid to conclude that two distinct groups exist in the population.

Although some argue that a potentially flawed model might still provide a useful summary of a set of observed data,[164] the misattribution of discrete versus continuous processes is particularly challenging in the search for potential phenotypes.[155] That is, if the goal is to identify observable characteristics that might identify an underlying genotype, the arbitrary creation of discrete groups in the presence of true continua would be of limited use. Great care must be taken from both theoretical and empirical perspectives in the accumulation of evidence for or against the existence of discrete phenotypes. As explored further below, several conditions might lead to the spurious identification of multiple classes when in actuality none exist. Appreciation of these potential alternative explanations for the identification of multiple classes will aid in building a cumulative science.

Static Versus Dynamic Clustering

Once it has been determined that there is a theoretical foundation for the positing of multiple classes, the next challenge is to determine whether groups will be based on static or dynamic methods of clustering. Whereas static clusters are typically based on data derived from a single cross-sectional assessment, dynamic clusters are based on longitudinal data assessed repeatedly over time. From a static perspective, only the

characteristics of an individual at a given fixed point of development are informative with regard to their association with a particular phenotype. A salient example is a simple assessment of the presence or absence of nicotine dependence, although other static phenotype groups have been posited. Importantly, the characteristics of the trajectory that an individual may have traversed to arrive at that particular point in development are not of interest (or, at a minimum, the information has no predictive utility).

Given the notion of equifinality (as seen in "watershed" or stage models of smoking; e.g., in chapter 3) it is possible that regardless of the multiple pathways that led initially into variation in smoking onset and acceleration, the only relevant phenotypes of nicotine dependence are the "mature" phenotypes that define variation in ultimate nicotine dependence and inability to abstain. If so, then dynamic phenotypes such as trajectories of smoking behavior are of potential but time-limited interest and are ultimately replaced by other static phenotypes. In contrast, a dynamic perspective not only considers the characteristics of an individual at a particular point in development but also explicitly considers the path that individual followed through the years leading up to the particular fixed point. For example, although three individuals might all report nicotine dependence at 25 years of age, one may have reached that point with an initial onset in early adolescence, one with an onset in late adolescence, and one with an onset in early adulthood. These very different ages of onset may aid in the identification of developmentally informed phenotypes of smoking behavior, the distinction of which would have been wholly occluded if only considering nicotine dependence at 25 years of age. An example of the importance of dynamic phenotypes was proposed by Shaw and colleagues,[165] who found that it was the trajectory of *change* in the thickness of the cerebral cortex rather than simply the thickness itself that was related to intelligence. From a statistical standpoint, the implementation of static and dynamic clustering techniques can lead to fundamentally different groupings of the same sample of individuals. As such, the selection of the optimal analytic approach has significant implications for both the inferences drawn from a given research study and for the development of an integrated understanding of empirical findings across existing literatures. The ultimate empirical reconciliation would primarily depend on the identification of meaningful genotypic differences as a function of static versus dynamic clustering.

Estimation of Within-Class Variability

Assuming an interest in the estimation of dynamic clusters, the next challenge is to determine which analytic approaches best correspond to the theoretical model under study. Whereas LCGA does not incorporate random variability in the growth process within class, GMM can include or omit these within-class random effects. The estimation of random effects for the growth process offers several advantages, including the incorporation of one or more predictors of the random growth parameters within class.[136] However, the omission of random effects also offers several advantages, including greater stability of estimation and correspondence to the hypothesized homogeneity within class.[137] This is an important decision because the inclusion or exclusion of within-class random effects for the growth parameters can directly influence the number of classes that are extracted from the same sample data.[135] One obvious strategy would be to fit models with and without within-class random effects and compare the correspondence between the two. If any observed differences are minimal, then the choice between the two parameterizations is less salient. However, in reality it is likely that the

solutions will differ, possibly substantially so. Moreover, the available theoretical models of smoking behavior are not well enough developed to determine whether random effects should be estimated within each class. As such, it falls on the applied researcher to consider the advantages and disadvantages of these approaches and make an informed and justifiable decision.

Importance of Nonnormality

As with the finite mixture models upon which it is based, GMM makes a strong assumption of within-class multivariate normality for both the repeated measures and the random trajectories. This assumption, in turn, dictates that the marginal distributions of the repeated measures (that is, the distribution of the measures for the fully aggregated sample) be nonnormally distributed. This is a straightforward result of mixing two or more normal distributions; under all but a small number of atypical conditions, the distribution of the mixture of two or more normal distributions must itself be, by definition, nonnormally distributed.[148] This assumption is what allows for the very extraction of multiple classes from a single sample. The complex nonnormal distribution for the aggregated sample can be approximated by the extraction of two or more normal distributions defined by different means and variances. Indeed, this is the most direct tie between GMM and the classic finite mixture model.[135]

Yet, this assumption poses a vexing problem, particularly when applying GMM within many areas of substance-use research. Namely, it has been shown that not only is marginal nonnormality a necessary condition for multiple class extraction but also is a *sufficient* condition.[134] Computer simulation studies have shown that when modestly nonnormal data are generated from a single homogeneous population, GMM identifies multiple groups 100% of the time.[134] Of course, this is precisely what the model is intended to do; multiple normal within-class distributions are estimated to approximate the more complex nonnormal aggregate distribution. A fundamental error, however, would be to conclude that multiple groups exist within the population when the optimal fitting multiple class model resulted solely as a function of the nonnormal aggregate distribution.[155]

This issue poses a key challenge when applying these techniques to the study of smoking phenotypes. Given the very nature of the construct under study, many observable measures of smoking behavior are not going to follow a normal distribution. But how does one know if this nonnormality is due to the inappropriate aggregation of data drawn from multiple classes, or instead, is simply an accurate reflection of the distribution of the construct? No analytic method was found that will distinguish which of these two conditions most likely accounts for the observed nonnormality of the measures under study. Further, it is unclear how GMM might best be used to empirically test for population heterogeneity when it is highly likely (if not nearly certain) that multiple classes will be extracted on the basis of the marginal distribution alone. More specifically, how is a research hypothesis subjected to potential falsification when the outcome is known before the test is conducted?

As with prior challenges, it is commonly recommended that theory be used as a guide under such circumstances. However, it is not always clear how this might actually be accomplished. In some sense, this presents a basic Aristotelian syllogism: if the aggregate data are nonnormally distributed, multiple classes will be extracted regardless of population heterogeneity; smoking-related measures are nonnormally distributed; thus, multiple classes will be extracted when fitted to smoking-related measures, regardless of population heterogeneity. As addressed

later, an important strategy is to avoid the reification of class extraction from a single sample of data. A triangulation of findings from multiple studies using multiple outcomes may offer the best strategy when searching for evidence of smoking-related phenotypes.

Proper Model Specification

Just as multiple classes can be extracted to approximate a nonnormal aggregate distribution, multiple classes can also be extracted to "absorb" the bias introduced by the estimation of a misspecified model.[135] The model misspecification might arise from the incorrect parameterization of the functional form of the trajectory, from the exclusion of one or more structural parameters, or from the omission of nonlinear relations among two or more constructs. Regardless of source, it has been shown that an incorrect model fitted to data drawn from a single population can result in the identification of multiple classes when none truly exist.[135] Given the ubiquity of misspecified models in applied research,[166] the potential for spurious class extraction related to model misspecification poses another key challenge when applying these methods to the study of smoking phenotypes.

Alternative Spans of Study

Given that the focus of this chapter has been on identifying potential smoking phenotypes from the estimation of developmental trajectories, an obvious challenge is the impact of the developmental span under study. Of course, alternative developmental spans of study do not typically pose a challenge within a given study. That is, most applications will use all of the repeated observations that are available for analysis. However, this poses a much greater challenge when attempting to identify consistent findings from the existing literature. Consider two hypothetical developmental phenotypes of smoking: one that consists of a late onset, modest acceleration, stable plateau, and a rapid decline to a low level of use, and the other consisting of precisely the same pattern except for a rapid decline to complete cessation. If one study were to follow a sample up to the point of decline, whereas another were to follow a sample past the point of decline, the resulting classes would likely be quite different between the two studies.[167,168] The fact that different groups are obtained for different spans of measurement does not negate the validity of the obtained groups. As expected theoretically, individuals can change their trajectories with a change in risk or protective factors, and certain trajectory groups would not be hypothesized to appear until certain ages. For example, developmentally limited or late-onset forms of substance use cannot be distinguished until well into adulthood. However, changes in identified trajectory groups with changes in developmental span pose a salient challenge when one attempts to draw a broader understanding about the characteristics of the underlying population from multiple studies covering multiple developmental stages. A similar problem may occur in terms of alternative frequencies of repeated assessments. That is, assessments that occur frequently (e.g., daily, weekly, or monthly) can capture complex fluctuations in the outcome variable that will be lost when assessments occur more rarely over longer intervals (yearly or less than yearly).

Inclusion Versus Exclusion of Abstainers

A long-standing issue that arises in almost any study of substance use is how to best handle stable abstainers (see chapter 6 for a more detailed discussion). One (often unsatisfactory) option is to simply delete these from the analysis. This is clearly not ideal given both the discarding of valuable data and the introduction of a biased sample

relative to the population from which it was drawn. A second option is to treat abstainers as a unique class before the execution of the clustering analysis. Thus, abstainers are an "observed" class whose existence need not be estimated; they are then added to the other classes that are identified via the clustering techniques. Although preferable to omitting these data entirely, the possibility remains that some abstainers truly are using but either denied this use or misrecorded their responses, or that a true abstainer is quite close to becoming a first-time user. Treating all of these as complete abstainers does not allow for the possibility of these other issues. Finally, newer techniques have been developed that allow for a hybrid-type modeling approach in which one model is fitted to a 0/1 dichotomy of no use versus use, and another model is fitted simultaneously to those who are reporting use.[169] These techniques are quite promising but need to be more fully explored.

Summary

There are several important hypothesized developmental phenotypes of smoking behavior, each of which can be empirically evaluated by using one of a number of analytical methodologies. Techniques such as survival analysis, latent transition analysis, and growth curve modeling can be used with varying degrees of success to empirically evaluate these predictions. This chapter has focused on one such approach—namely, methods for clustering developmental trajectories. As described earlier, these methods offer multiple potential advantages for the study of phenotypes of smoking behavior. However, it is also important to closely consider challenges that arise when using these methods in practice. This chapter considers some of the particularly important issues that arise when studying smoking phenotypes on the basis of the clustering of developmental trajectories over time. The goal is to highlight these challenges so that future applications of trajectory clustering techniques may be cognizant of these issues and proceed in a thoughtful and careful manner.

An Empirical Example: Trajectories in the Indiana University Smoking Survey

To provide an empirical example of these issues raised by modeling smoking trajectories, a series of models are presented using data from the Indiana University Smoking Survey. The focus is on a comparison of two time windows (ages 10 to 32 years and 10 to 42 years), relating static indicators of adult nicotine dependence to clusters that are dynamic (developmental smoking trajectories estimated via LCGA). These analyses allow for three important comparisons. First, the same sample can be used to directly compare the optimal number of groups based on a 22-year developmental window with the optimal number of groups based on a 32-year developmental window. This will demonstrate whether different phenotypic groups might be identified if the same sample were followed over a longer developmental period. Second, the stability of group membership can be examined. That is, the extent to which an individual assigned to a particular phenotype in the shorter window remains a member of that same phenotype in the longer window can be assessed. Finally, the relation between adult nicotine dependence and trajectory group membership examines whether these trajectory groups have any systematic implications for nicotine dependence. Taken together, these analyses will offer a concrete empirical demonstration of both several key advantages and several specific challenges that are encountered when using these techniques in practice.

The Indiana University Smoking Survey is an ongoing, cohort-sequential study of the natural history of cigarette smoking that began in 1980.[13,56,170,171] Between 1980 and 1983, all consenting 6th–12th graders in a county school system in the Midwest completed annual surveys (total *N* who were assessed at least once = 8,487). The sample included 10 cohorts that correspond to the graduating classes of 1980–1989. Follow-ups were conducted in 1987 (73% retention, *N* = 6,234, ages 15–25 years), 1993 (73% retention, *N* = 6,223, ages 21–32 years), and 1999 (71% retention, *N* = 6,068, ages 27–37 years), as well as 2005 (70% retention, *N* = 5,931, ages 32–42 years). Because the sample was 96% non-Hispanic Caucasian, ethnic differences were not considered.

Sample representativeness has been described in detail elsewhere.[13,171] Demographically, the sample is similar to the community from which it was drawn: 64% marriage rates in this sample compared with 66% among adults of similar ages in the Midwest,[172] and 97% high school graduation rates in the sample compared with 92% among adults of similar ages in the Midwest.[173] At the last completed follow-up (1999), the smoking rate in the sample was 26%, the same rate found statewide.[174] Thus, the sample is representative of its community; that is, predominantly white and well educated.

Attrition biases have been discussed in detail elsewhere.[171,175] For each follow-up, those who were lost were compared with those who were retained in terms of their earlier data. Dropouts were more likely to be smokers and have more positive attitudes and beliefs about smoking. They also had parents and friends who were more likely to smoke (effect sizes ranging from r^2 of .01 to .02). Because of the consistent pattern of findings, some caution is warranted when making generalizations.

For these analyses, two subsamples of participants were selected. First, trajectories were modeled using the first six waves of measurement, selecting participants who had been measured at least once (ages ranged from 10 to 32 years, and 51% were males). Next, these results were compared with trajectories obtained from considering the entire available data (waves 1–8, age range = 10–42 years, 51% male).

Procedures

Adolescent data were collected with group-administered questionnaires in school. In 1987, these procedures were followed for cohorts who were still in high school. For cohorts who had graduated from high school (and for all participants in 1993 and after), a survey was sent by mail and followed up by telephone interviews if questionnaires were not returned. Participants were paid $15 for mailed surveys, and in 1999, they also were entered into a lottery for prizes of $200. At the 2005 follow-up, mailed surveys with telephone interview follow-ups were again used with participant fees ($30) and lottery incentives.

Measures

Smoking Level

Smoking level was determined by two items. Participants reported their current smoking status as "never smoked, not even a single puff"; "smoked once or twice 'just to try' but not in the last month"; "do not smoke, but in the past I was a regular smoker"; "smoke regularly but no more than once a month"; "smoke regularly, but no more than once a week"; or "smoke regularly and more than once a week." Participants also reported the number of cigarettes they typically smoked each day (from 0 to 20 or more). To improve the validity of self-reported smoking status in adolescence, a bogus pipeline was used from 1981 to 1983. As reported elsewhere,[171] a study using an unannounced bioassay with a subsample

of the participants supported the validity of the self-reports.

For these analyses, responses were combined into a six-level variable to reflect current smoking at each measurement wave, reflecting both frequency and quantity of smoking as follows: 0 = not currently smoking (nonsmokers, ex-smokers, those who had smoked once or twice, but not in the past month); 1 = up to monthly smoking; 2 = up to weekly smoking; 3 = weekly or more smoking, but only 10 or fewer cigarettes per day; 4 = weekly or more smoking of 11–20 cigarettes per day; and 5 = weekly or more smoking of 20 or more cigarettes each day. In some cases, responses to these items were ambiguous, and additional items were consulted concerning the number of cigarettes smoked yesterday and the time since the last cigarette was smoked.

Adult Tobacco Dependence

Tobacco dependence was measured with the Fagerström measure as well as by examining two individual items: number of cigarettes smoked in a typical day and time to first cigarette in the morning.

Family History of Smoking

Family history of smoking was based on participants' reports of lifetime smoking among their biological parents at waves 5 and 6. Family history was scored positive if at least one biological parent was reported to be a smoker.

Data Analysis

Individual, time-specific, smoking behavior scores were modeled over ages 10–32 years (using waves 1–6) and ages 10–42 years (using waves 1–8) with LCGA. However, several groups were constructed a priori. Individuals who reported never having smoked a single puff or having smoked once or twice "just to try" but never progressing beyond this category at any measurement were defined a priori as continuous abstainers (N = 4,642 in waves 1–6, and N = 4,298 in waves 1–8). Individuals who were never measured as smokers but only as "ex-smokers" were defined a priori as stable quitters (N = 672 in waves 1–6, and N = 669 in waves 1–8). Individuals who reported periods of smoking, quitting, and then smoking again were defined a priori as a "relapsing/remitting" group (N = 535 in waves 1–6, and N = 874 in waves 1–8). All other participants were clustered empirically.

All LCGA model estimation was performed using Proc Traj[133,176] as available for Statistical Analysis Software (SAS). A series of latent class models were estimated ranging from one to seven classes assuming a censored normal (0, 5) response distribution. No user-supplied start values were used in the initial models. A number of criteria were used to select the "best" model for the two age ranges. The criteria included an overall reduction in the Bayesian Information Criterion[177] and Akaike Information Criterion[178,179] to be less than 100 for the next largest class as well as for the model to remain stable (with the addition of and changes to subsequent start values) and consistent with substantive theory. Model selection for waves 1–6 and 1–8 was done independently of each other.

Once the appropriate models were chosen, the modal probability of class membership was used to place individuals into a particular class (the class for which their membership probability was highest). These classes serve as the independent variables in all further analyses. Class or "group" membership was used to predict several outcomes including Fagerström dependence, number of cigarettes smoked in a typical day, time to first morning cigarette, and family history of smoking. Each outcome was evaluated independently.

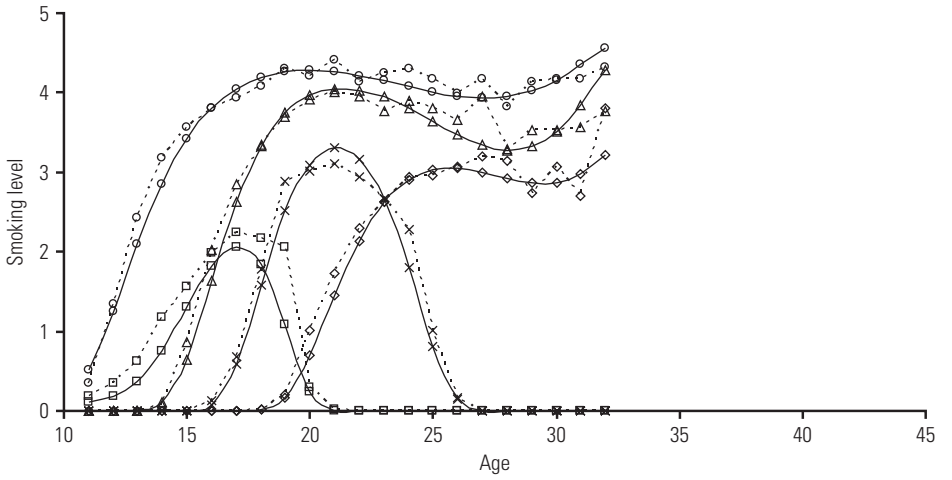
For dichotomous outcomes, logistic regression was employed and parameter estimates were obtained using SAS's Logistic procedure. Each model consisted of dummy-coded class membership variables (0 for not in the class, 1 for in the class). Abstainers were zero on all membership variables. Linear contrasts were then used to test for significant differences between each class's ability to predict the outcome. Continuous outcomes were predicted with ANOVAs.

Results

LCGA Results

A five-class solution was found to be optimal for waves 1–6. As seen in figure 5.1, the five classes were experimenters (4.5%); developmentally limited smokers (3.4%); early-onset, persistent smokers (7.4%); high-school-onset, persistent smokers (9.9%); and late-onset, persistent smokers (6.4%). As noted earlier, a priori groups were stable abstainers (54%), stable quitters (8%), and relapsing/remitters (6%). Experimenters began smoking early (around 11 years of age), but never smoked more than occasionally, and generally quit smoking by 21 years of age. Developmentally limited smokers started smoking around the age of 16 years, smoked regularly but averaged around 10 cigarettes per day at their peak, and gave up smoking by 27 years of age. Early-onset, persistent smokers typically started around the age of 11 years and increased their smoking quickly to a peak of more than one-half pack per day, which they maintained over development. High-school-onset, persistent smokers started around 14 years of age, quickly increased their smoking, and although never quite as heavy as the early group, continued to smoke heavily. Late-onset, persistent smokers started around the age of 18 years and quickly became moderate smokers.

Figure 5.1 Five-Class Solution for Waves 1–6



Note. Smoking level was determined by participants' reported frequency and quantity of smoking, expressed numerically as follows: 0 = not currently smoking (nonsmokers, ex-smokers, those who had smoked once or twice, but not in the past month); 1 = up to monthly smoking; 2 = up to weekly smoking; 3 = weekly or more smoking, but only 10 or fewer cigarettes per day; 4 = weekly or more smoking of 11–20 cigarettes per day; and 5 = weekly or more smoking of 20 or more cigarettes each day. Dotted lines denote observed trajectories. Solid lines denote model-implied trajectories. Experimenters are denoted by squares; developmentally limited smokers by "x"s; early-onset, persistent smokers by circles; high-school-onset, persistent smokers by triangles; and late-onset, persistent smokers by diamonds.

As seen in table 5.2, the gender makeup of the groups was generally evenly split, except for the late-onset, persistent group, which had significantly more males compared to every other group except for the relapsing/remitting groups. However, there was a stronger relation between group membership and education. When measured at wave 6, the sample of participants reporting education data had approximately 36% of individuals with BA degrees or higher. The least educated group comprised the early-onset, persistent smokers (3.4% of whom had completed a college degree). They significantly differed from all other groups. Other groups with relatively low levels of educational attainment were the high-school-onset group (14.4% with BAs or higher) and the relapsing/remitting group (13.5% with BAs or higher), who significantly differed from all other groups but not from each other. The abstainers were the most highly educated group (47.9% with BAs or higher), and they significantly differed from all other groups. Late-onset and developmentally limited groups were also relatively well educated (approximately 38% with BAs or higher) and did not significantly differ from each other.

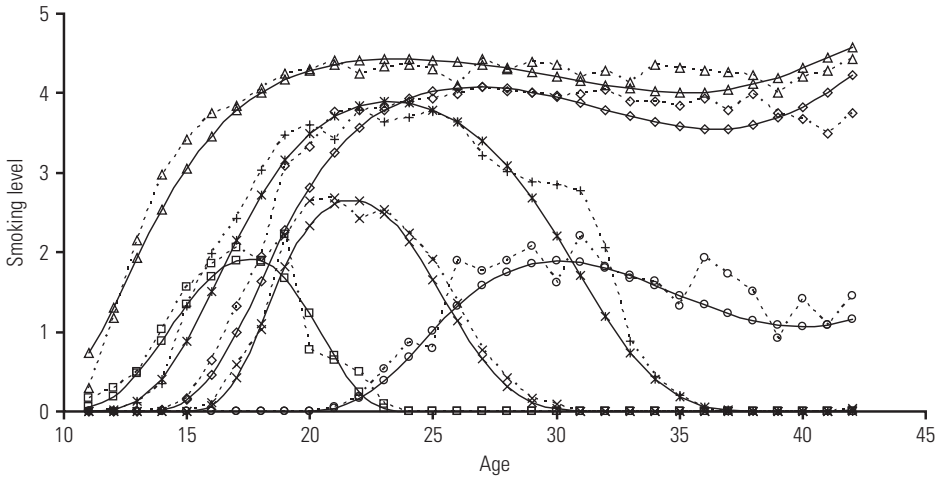
A six-class solution was found to be optimal for waves 1–8. As seen in figure 5.2, the six classes were experimenters (4.2%); developmentally limited smokers (4.7%); successful quitters (2.4%); early-onset, persistent smokers (8.4%); high-school-onset, persistent smokers (8.7%); and late-onset, persistent smokers (3.1%). Compared to the waves 1–6 model, this model produced one additional class: successful quitters. As noted earlier, the a priori groups were stable abstainers (50.4%), stable quitters (7.8%), and relapsing/remitters (10.24%). Similar to the experimenters who were identified in waves 1–6, this experimenter group began smoking around the age of 11 years, never smoked more than occasionally, and, in general, quit smoking by the age of 22 years. Similarly, the developmentally limited group generally started around the age of 16 years, smoked more on average than did the experimenters, (but less than one-half pack at their peak), and gave up smoking by the age of 30 years. As in the waves 1–6 model, the early-onset, persistent group typically started smoking around the age of 11 years, increased their smoking behavior quickly, and maintained the highest level of smoking over the course of

Table 5.2 Trajectory Group Sizes and Relationship with Gender and Educational Attainment

	N		%		Gender (% female)		College degree or higher (%)	
Waves	**1–6**	**1–8**	**1–6**	**1–8**	**1–6**	**1–8**	**1–6**	**1–8**
Abstainers	4,642	4,298	54.39	50.36	49.59^{a}	50.58^{ab}	47.85	56.50^{a}
Stable quitters	672	669	7.87	7.84	48.21^{a}	46.64^{ab}	35.00^{ac}	47.73^{b}
Developmentally limited	288	398	3.37	4.66	45.83^{ab}	43.97^{ab}	38.37^{a}	59.31^{a}
Experimenters	381	362	4.46	4.24	51.44^{a}	55.25^{a}	28.01^{c}	41.60^{b}
Early onset	628	715	7.36	8.38	48.57^{a}	48.04^{ab}	3.44	6.90
Successful quitters	—	208	—	2.44	—	57.21^{ab}	—	38.71^{b}
High school onset	844	744	9.89	8.72	50.89^{a}	44.62^{b}	14.42^{b}	24.32^{c}
Late onset	544	266	6.37	3.12	38.60^{bc}	30.83	37.78^{a}	45.89^{b}
Relapse/remit	535	874	6.27	10.24	45.05^{ac}	46.00^{ab}	13.45^{b}	26.17^{c}
Total	8,534		100		48.59		36.19	45.75

Note. Groups that share subscripts do not significantly differ from one another ($p < .05$).

Figure 5.2 Six-Class Solution for Waves 1–8



Note. Smoking level was determined by participants' reported frequency and quantity of smoking, expressed numerically as follows: 0 = not currently smoking (nonsmokers, ex-smokers, those who had smoked once or twice, but not in the past month); 1 = up to monthly smoking; 2 = up to weekly smoking; 3 = weekly or more smoking, but only 10 or fewer cigarettes per day; 4 = weekly or more smoking of 11–20 cigarettes per day; and 5 = weekly or more smoking of 20 or more cigarettes each day. Dotted lines denote observed trajectories. Solid lines denote model-implied trajectories. Experimenters are denoted by squares; developmentally limited smokers by "x"s; quitters by pluses (+); early-onset, persistent smokers by circles; high-school-onset, persistent smokers by triangles; and late-onset, persistent smokers by diamonds.

the study (more than one-half pack per day). High-school-onset, persistent smokers (as described in the waves 1–6 model) began to smoke around the age of 16 years and smoked fairly heavily over development (almost as much as the early-onset group). The late-onset, persistent group identified in the waves 1–8 model tended to start a little later (21 years of age) than in the waves 1–6 model (18 years of age). Moreover, this group generally increased more slowly and maintained a relatively low level of smoking behavior. Successful quitters (the group that did not emerge in the waves 1–6 model) started slightly later, around the age of 12 years, smoked fairly heavily for many years, and eventually quit in adulthood (by 37 years of age).

As shown in table 5.2, the pattern of gender and education differences was quite similar to the waves 1–6 model. The groups were generally evenly split in gender, with the percentage of females in a group ranging from 30.83% (late onsetters, who significantly differed from all other groups) to 57.21% (successful quitters who did not differ from any other group except for late onsetters).

As with the waves 1–6 model, a strong relation was found with educational attainment. When measured at wave 8, 45.8% of the sample who reported their educational attainment had completed a BA degree or higher. The trajectory group with the highest educational attainment was the developmentally limited smokers (59.3% with BAs or higher), and they did not significantly differ from the abstainers (56.5%). Other groups with relatively high levels of educational attainment were the stable quitters (47.7% with BAs or higher), experimenters (41.6% with BAs or higher), successful quitters (38.7% with BAs or higher), and late-onset smokers (45.9% with

BAs or higher), and these groups did not significantly differ from each other. As in the waves 1–6 model, the lowest level of educational attainment was found for the early-onset, persistent group (only 6.9% with a BA degree or higher), and this group significantly differed from all others. The high-school-onset group (24.3% with a BA or higher) and the relapsing/remitting group (26.2% with a BA or higher) were also relatively less educated and significantly differed from all other groups, but not from each other.

Stability of Classification across the Models

In general, group membership remained fairly consistent across analyses. As can be seen in table 5.3, a large proportion of the sample (81.9% of the total sample, 64.6% of the empirically classified subsample) was classified in the same group across the different age span models. Moreover, differences in classification across models were theoretically reasonable. Of those who were stable abstainers in waves 1–6, 93% were still abstainers 10 years later, and those who changed categories were most likely to become late-onset, persistent smokers or successful quitters. Of those who were stable quitters in waves 1–6, 82% were stable quitters 10 years later, and those who were not were categorized as relapsing/remitters. Of those who were developmentally limited smokers in waves 1–6, 62% were classified in the same group 10 years later, and those who were not were most likely to be classified as relapsing/remitters. Of those who were experimenters in waves 1–6, 77% were classified in the same group 10 years later, and those who were not were most likely to be classified as relapsing/remitters. Of those who were early-onset, persistent smokers in waves 1–6, 90% were classified in the same group 10 years later, and those who were not were most often classified as successful quitters or experimenters.

Table 5.3 Stability of Classification across the Waves 1–6 and Waves 1–8 Models

	Waves 1–8 trajectory group								
Waves 1–6	Abstainers	Stable quitters	Developmentally limited	Experimenters	Early onset	Successful quitters	High school onset	Late onset	Relapse/ remit	Total
Abstainers	4,298	115	3	3	0	3	76	128	16	4,642
Stable quitters	0	554	0	0	0	0	0	0	118	672
Developmentally limited	0	0	178	16	0	5	28	1	60	288
Experimenters	0	0	5	292	6	1	4	0	73	381
Early onset	0	0	0	24	563	24	0	0	17	628
Successful quitters	0	0	0	0	0	0	0	0	0	0
High school onset	0	0	65	27	146	137	433	0	36	844
Late onset	0	0	147	0	0	38	203	137	19	544
Relapse/remit	0	0	0	0	0	0	0	0	535	535
Total	4,298	669	398	362	715	208	744	266	874	8,534

The largest difference between the two models is seen in the successful quitting group, which did not emerge in the earlier years. The waves 1–8 successful quitting group was drawn primarily from the high-school-onset and late-onset smokers in waves 1–6. This finding indicates that a later onset of smoking is associated with greater likelihood of successful cessation.

In a parallel finding, there was somewhat less stability in the waves 1–6 high-school-onset smokers, with 51% classified in the same group 10 years later. Those who changed groups were likely to be classified as either early onset (reflecting the similarities between middle school and high school onset) or successful quitters (reflecting those who succeeded in cessation).

Finally, there was low stability in the waves 1–6 late-onset smokers, with only 25% being classified as late onset in the waves 1–8 model. Those who were late-onset smokers in the waves 1–6 model were most likely to be classified as high school onset in the waves 1–8 model (37%), reflecting the ambiguity of their onset at the age of 18 years as somewhat like adolescents and somewhat like adults. Moreover, a substantial number of the waves 1–6 late-onset smokers (27%) were classified as developmentally limited smokers in the waves 1–8 model, reflecting the fact that they stopped smoking by adulthood. Conversely, the late-onset group that emerged in the waves 1–8 model was drawn mostly from the waves 1–6 late-onset smokers and abstainers and showed a later age of onset than did the first model (22 years of age rather than 18 years).

Family History Analysis

Logistic regression models that related trajectory group membership to family history of smoking produced similar results for the waves 1–6 and waves 1–8 groups. In both models, the lowest likelihood of having a smoking parent was for late-onset smokers, abstainers, and developmentally limited smokers, who did not significantly differ from each other. In the waves 1–6 model, the highest likelihood of having a smoking parent was for early-onset, high-school-onset, and relapsing/remitting groups, who differed from the other groups, but not from each other. In the waves 1–8 model, the highest likelihood of having a smoking parent was for the early-onset and the successful quitter groups, who significantly differed from abstainers and late-onset smokers, but not from each other (table 5.4).

Indicators of Adult Nicotine Dependence: Amount Smoked, Time to First Cigarette, and Fagerström Dependence Diagnoses

For both the waves 1–6 and 1–8 models, the groups who smoked at the end of the trajectory (early onset, high school onset, late onset, and relapsing/remitting) were compared on indicators of nicotine dependence measured at wave 8 (ANOVAs and logistic regressions in table 5.5). (Note that this analysis examines current dependence at wave 8, but does not identify the timing of onset of dependence). The findings for the two models were quite similar. For both waves 1–6 and 1–8 groupings, the early-onset, high-school-onset, and late-onset groups significantly differed from each other in both models on all indicators. The early-onset group showed the highest percentage of tobacco dependence (more than one-half of the group); this percentage was strikingly higher than the late-onset group (12%–19% in the second model). Similarly, the early-onset group smoked at high levels (averaging one pack per day), whereas the late-onset group was closer to "chippers" (averaging seven cigarettes per day for the waves 1–8 late-onset group). Also paralleling these findings, the early-onset group was more

Table 5.4 Relationship of Trajectory Group Membership to Family History of Smoking

	Waves 1–6		Waves 1–8	
	N	With at least 1 smoking parent (%)	*N*	With at least 1 smoking parent (%)
Abstainers	3,359	67.40[a]	3,154	66.93[a]
Stable quitters	452	71.68[ab]	451	71.40[ac]
Developmentally limited	269	68.77[ab]	353	69.12[ac]
Experimenters	310	73.23[bc]	290	74.48[abc]
Early onset	327	85.02[d]	390	83.85[bc]
Successful quitters	—	—	164	86.59[bd]
High school onset	622	81.19[cd]	519	78.23[cd]
Late onset	460	67.61[ab]	217	63.59[a]
Relapse/remit	432	80.32[bcd]	693	77.20[cd]
Total % with a smoking parent		71.27		
Total *N*		6,231		

Note. Groups that share superscripts do not significantly differ from one another ($p < .05$).

Table 5.5 Relationship of Trajectory Group Membership to Smoking Dependence Indices in Wave 8

	Waves 1–6			Waves 1–8		
	Dependent using FTND ≥ 6 (%)	Number of cigarettes smoked in 1 day	Smoke first cigarette of day immediately (<5 min.) (%)	Dependent, using FTND ≥ 6 (%)	Number of cigarettes smoked in 1 day	Smoke first cigarette of day immediately (<5 min.) (%)
Early onset	59.0[a]	22.1[a]	41.4[a]	54.4[a]	22.0[a]	39.4[a]
High school onset	34.0[b]	17.1[b]	22.0[b]	31.7[b]	16.3[b]	19.1[b]
Late onset	18.8[c]	12.0[c]	10.7[c]	11.9[c]	6.7[c]	6.9[c]
Relapse/remit	34.7[b]	16.5[b]	19.5[b]	25.3[d]	13.6[d]	13.5[d]
Overall	38.8	17.6	24.9	34.6	16.1	21.4

Note. Groups that share superscripts do not significantly differ from one another ($p < .05$). FTND = Fagerström Test for Nicotine Dependence.

likely to smoke the first cigarette of the day immediately upon awakening (39%–41% of these groups in the two models), whereas less than 10% of the late-onset groups did this.

Differences between the waves 1–6 and waves 1–8 models involved the relapsing/remitting group. For the waves 1–6 model, this group resembled the high-school-onset group on all indicators, whereas it significantly differed from the early- and late-onset group on these same indicators. In the waves 1–8 model, the relapsing/remitting group was distinctly and significantly different from all of the other groups and was the second lowest (to the late-onset group) in indicators of nicotine dependence.

Discussion

These findings demonstrate both the potential utility and some of the challenges involved with empirically identifying multiple developmental trajectories of smoking. Analyses revealed meaningful heterogeneity in trajectories that would be

relevant for genetic studies. Results also showed that static assessment at any one time point has limitations. For example, early-onset smoking is correlated with steep acceleration and high persistence. Examining only a single age point of onset, however, would not reveal this finding because some of these early onsetters will merely experiment and not progress to regular smoking. However, a developmental trajectory that combines early onset, steep acceleration, and high persistence produces the highest risk for adult dependence and shows high levels of family history of smoking. This suggests a phenotype of interest for genetic analysis. The very low educational attainment of this group also suggests the presence of other risk factors (as identified in other studies, table 5.1). Thus, the low educational attainment may reflect the effects of endophenotypes that undermine educational success, such as conduct problems, impulsivity, behavioral undercontrol, and attention deficit. It is also important to assess the possible role of low socioeconomic status (which also constrains educational attainment).

Although the early-onset, persistent group is clearly at highest risk, the differences between the early-onset and high-school-onset groups appear to be quantitative and dimensional, rather than demonstrating a qualitatively distinct etiological pathway of smoking acquisition. These trajectories show both steep acceleration and persistence at high levels of smoking, although the early-onset group was somewhat elevated in indicators of adult dependence and somewhat lower in educational attainment. Thus, the available data suggest that there is simply a difference in severity between the early-onset and high-school-onset groups (although a consideration of predictors of trajectories might reveal other patterns).

Moreover, even among adolescents whose onset of smoking is early in adolescence, some became successful quitters (albeit a small prevalence). These successful quitters had family histories of smoking equivalent to the early-onset group, suggesting that a simple genetic explanation of cessation is likely to be insufficient. The similarities between the early-onset and successful quitting groups in family history is consistent with a stage model demonstrating multifinality, so that similar factors may have led the quitters and persistent groups to initiate smoking, but different factors (possibly arising in adulthood) ultimately determine successful smoking cessation. Interestingly, the groups differed in educational attainment, so perhaps social contextual and intrapersonal factors related to success in higher education may provide some way to distinguish successful quitters from early-onset, persistent smokers.

Although the early-onset and high-school-onset groups appear to represent continuous distributions of risk, the late-onset group appears more qualitatively distinct. In fact, in many ways, late-onset group members resembled abstainers. They have low levels of familial smoking and high levels of educational attainment. They also show low levels of adult smoking and dependence. In these ways, they resemble tobacco "chippers," who may be relatively less vulnerable to tobacco dependence. Thus, whatever mechanisms produced their late-onset smoking, these mechanisms may be different from those that underlie early and high school smoking acquisition. This late-onset group has been underrecognized and understudied, unless they are considered to be the same as "chippers" (light smokers).

The findings also demonstrate several challenges that arise in attempting to empirically identify multiple trajectories of smoking. For example, although the groups obtained from waves 1–6 and 1–8 were substantially similar in terms of classification, the solutions were not identical. Changes in group classification occur with changes in the ages under

study, sample characteristics, variations in smoking measures and other factors, and these changes make it challenging to examine the robustness of trajectories derived across different studies. However, it is also important to note that some of the differences between the waves 1–6 and 1–8 models reflect meaningful developmental changes, and that trajectories measured at different stages of the life course would be expected to differ. For example, trajectories of developmentally limited or late-onset smoking cannot emerge until a sufficient age span is measured. Moreover, individuals can change trajectory groups with a meaningful change in their smoking behavior (i.e., individuals may smoke at high levels but then become successful long-term quitters). Future research might explore such meaningful developmental changes by using methods developed to identify "regime switching" (i.e., switching between trajectory groups[180]).

Moreover, the findings were produced by a combination of approaches in which some groups were defined a priori and others were derived empirically. For example, a wholly empirical approach could not differentiate individuals who were always measured as abstainers from those who were always measured as ex-smokers since both groups would need to have identical scores to enter into the model. Also, these analyses did not examine modeling solutions that incorporate within-class variability, and all of these modeling decisions will affect the trajectory groups produced.

Finally, although not necessarily a limitation of this approach, these analyses did not address many questions that go beyond the scope of a single chapter. For example, tobacco use was not examined in forms other than cigarette smoking (which may affect patterns and trajectories of smoking). In addition, prospective predictors of these trajectories (other than family history of smoking) were not examined nor were trajectory group memberships related to hypothesized endophenotypes.

Future Research Directions

The literature review and empirical example provided in this chapter point to several directions for future research. The first task is a better specification of the relation between trajectories of smoking behavior and the development of nicotine dependence, as well as the relation between adolescent and adult trajectories. Given stage models (such as the "watershed" model in chapter 3), developmental trajectories of smoking acquisition may better be considered as "transitional" phenotypes whose etiological determinants differ from those that underlie the phenotypes of nicotine dependence. Thus, developmental phenotypic information may be very important at particular stages in the smoking trajectory to mark diverse etiological mechanisms underlying acquisition, but these diverse pathways may become relatively less important in the presence of tobacco dependence. In addition, the measurement equivalence of tobacco dependence symptoms over the life span requires further study to determine the similarities and differences between tobacco dependence in adolescence and adulthood. More research is needed on the heterogeneity in time course and predictors of the transitions from initial exposure to dependence. Moreover, further research is required to understand the age-specific effects of initial nicotine exposure, which have shown a significant relation between an early age of onset and steeper acceleration over time. The mechanisms underlying this relation require further exploration in both animal and human models.

Another important question is whether a particular individual feature of a trajectory (e.g., age of onset or steepness of

acceleration) is the important phenotype or whether it is more useful to consider an entire trajectory group. The probabilistic nature of empirically identified trajectories (which change with different measures and age spans and from which individuals may enter and exit over time) creates a conflict with the goal of defining "true" and unchanging groups for genetic analysis. Moreover, different research approaches are needed to determine whether phenotypes are best considered as categorical "groups" or as representations of an underlying continuous dimension. For future research on multiple trajectories of smoking, an accumulating literature on the subject will help to determine whether empirically identified trajectory groups are reliable across different samples, and whether there are important ethnic differences in these groups. Moreover, very few studies relate these trajectories to measured genotypes or that examine trajectories in genetically informative samples (see chapter 7 for an example). More studies are needed to understand these areas. It may also be useful to consider tobacco within a context of other substance use (chapter 7) or broader "externalizing" disorders because the underlying genotype may reflect a broader tendency to disinhibition rather than a specific risk for tobacco dependence. Along with this consideration, studies are needed that relate smoking trajectories to indicators of hypothesized endophenotypes and to indicators of tobacco dependence.

Finally, practical issues must be considered in this type of research. Unless reliable and valid methods can be developed to retrospectively reconstruct trajectories, these must be derived from costly longitudinal studies. Moreover, within longitudinal studies, there are trade-offs between intensity and frequency of measurement intervals. For example, more frequent and more intense measurements provide greater resolution of transition points but require greater participant commitment and greater financial resources. These studies also require large sample sizes if heterogeneity in trajectories is of interest (especially given the high prevalence of nonsmoking). A useful approach may be to target important ages or high-risk groups or to use accelerated longitudinal designs. One potentially helpful strategy may be to take advantage of existing longitudinal data sets by adding measures of endophenotypes (as well as genetic data) that are not time varying or age dependent. These represent some of a number of open questions that remain to be addressed through future research.

Summary

As demonstrated in this chapter, a consideration of developmental trajectories of cigarette smoking has potential for refining phenotypes of smoking for genetic analysis and for illuminating etiological mechanisms. Research has demonstrated meaningful heterogeneity in age of onset, slope of acceleration, and peaks and persistence of use. These features have been found to be significantly related to some hypothesized endophenotypes as well as to indicators of tobacco dependence. In a very small number of studies, these trajectory features have also been related to genetic variability. However, this literature is still in a very early stage. It is premature to conclude that smoking trajectories (or even individual features of trajectories) will constitute important phenotypes of smoking.

At the same time, the study presented here demonstrates that developmental aspects of trajectories of smoking acquisition may be useful in refining phenotypes of smoking. There is evidence (including the evidence provided in the empirical example) that heterogeneity in trajectories is related to indicators of nicotine dependence in adulthood. Identifying these trajectories may

help to illuminate the multiple etiological pathways that underlie the development of tobacco dependence. The next chapters in this section extend this work to genetically informative designs (chapter 6) and to the consideration of dual trajectories of tobacco and alcohol use in a genetically informative design (chapter 7).

Conclusions

1. Previous studies (and the empirical example presented in the chapter) have identified multiple developmental trajectories of tobacco use from adolescence to adulthood. These trajectory groups, which vary in age of onset, rate of acceleration, and persistence of smoking over time also vary in their antecedents and correlated risk factors. These trajectories may be informative as developmental phenotypes for genetic studies of tobacco use.
2. Statistical approaches such as latent class growth analysis and growth mixture modeling can be useful in evaluating developmental trajectories of smoking behavior. However, challenges in using these approaches include the handling of within-class random effects, the impact of a nonnormal aggregate distribution on the classes extracted, the need for proper model specification and parameterization, the span of evaluated data, and the impact of abstainers on the model.
3. Analysis of a 25-year cohort-sequential study of smoking behavior identified six distinct trajectories of smokers across eight waves of data collection. These trajectory groups were experimenters; developmentally limited smokers; early-onset, persistent smokers; high-school-onset, persistent smokers; late-onset, persistent smokers; and successful quitters, with a priori groups of stable abstainers, stable quitters, and relapsing/remitters. Trajectory group membership was related to educational attainment, family history of smoking, and indicators of nicotine dependence.

References

1. Kendler, K. S., M. C. Neale, P. Sullivan, L. A. Corey, C. O. Gardner, and C. A. Prescott. 1999. A population-based twin study in women of smoking initiation and nicotine dependence. *Psychological Medicine* 29 (2): 299–308.
2. Madden, P. A., A. C. Heath, N. L. Pedersen, J. Kaprio, M. J. Koskenvuo, and N. G. Martin. 1999. The genetics of smoking persistence in men and women: A multicultural study. *Behavior Genetics* 29 (6): 423–31.
3. Lerman, C., and W. Berrettini. 2003. Elucidating the role of genetic factors in smoking behavior and nicotine dependence. *American Journal of Medical Genetics Part B, Neuropsychiatric Genetics* 118 (1): 48–54.
4. Swan, G. E., K. S. Hudmon, L. M. Jack, K. Hemberger, D. Carmelli, T. V. Khroyan, H. Z. Ring, et al. 2003. Environmental and genetic determinants of tobacco use: Methodology for a multidisciplinary, longitudinal family-based investigation. *Cancer Epidemiology, Biomarkers & Prevention* 12 (10): 994–1005.
5. Sroufe, L. A., and M. Rutter. 1984. The domain of developmental psychopathology. *Child Development* 55 (1): 17–29.
6. Cicchetti, D. 2006. Development and psychopathology. In *Developmental psychopathology, vol. 1, theory and method*, 2nd ed., ed. D. Cicchetti and D. J. Cohen, 1–24. Hoboken, NJ: Wiley.
7. Mayhew, K. P., B. R. Flay, and J. A. Mott. 2000. Stages in the development of adolescent smoking. *Drug and Alcohol Dependence* 59 Suppl. 1: S61–S81.
8. Sroufe, L. A. 1997. Psychopathology as an outcome of development. *Development and Psychopathology* 9 (2): 251–68.
9. Pickles, A., and J. Hill. 2006. Developmental pathways. In *Developmental psychopathology, vol. 1, theory and method,* 2nd ed., ed. D. Cicchetti and D. J. Cohen, 211–43. Hoboken, NJ: Wiley.
10. Gottlieb, G., and M. T. Willoughby. 2006. Probabilistic epigenesis of psychopathology. In *Developmental psychopathology, vol. 1, theory and method*, 2nd ed., ed. D. Cicchetti and D. J. Cohen, 673–700. Hoboken, NJ: Wiley.
11. Bergman, L. R., A. von Eye, and D. Magnusson. 2006. Person-oriented research strategies in developmental psychopathology. In *Developmental psychopathology, vol. 1, theory and method*, 2nd ed., ed. D. Cicchetti and D. J. Cohen, 850–88. Hoboken, NJ: Wiley.
12. Johnston, L. D., P. M. O'Malley, J. G. Bachman, and J. E. Schulenberg. 2006. *Monitoring the Future: National results on adolescent drug use; Overview of key findings, 2005* (NIH publication no. 06-5882). Bethesda, MD: US Department of Health and Human Services, National Institutes of Health, National Institute on Drug Abuse. http://www.monitoringthefuture.org/pubs/monographs/overview2005.pdf.
13. Chassin, L., C. C. Presson, J. S. Rose, and S. J. Sherman. 1996. The natural history of cigarette smoking from adolescence to adulthood: Demographic predictors of continuity and change. *Health Psychology* 15 (6): 478–84.
14. Chen, K., and D. B. Kandel. 1995. The natural history of drug use from adolescence to the mid-thirties in a general population sample. *American Journal of Public Health* 85 (1): 41–47.
15. Fagerström, K. O. 1978. Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addictive Behaviors* 3 (3–4): 235–41.
16. American Psychiatric Association. 1994. *Diagnostic and statistical manual of mental disorders: DSM-IV.* 4th ed. Washington, DC: American Psychiatric Association.
17. Kandel, D., C. Schaffran, P. Griesler, J. Samuolis, M. Davies, and R. Galanti. 2005. On the measurement of nicotine dependence in adolescence: Comparisons of the mFTQ and a DSM-IV-based scale. *Journal of Pediatric Psychology* 30 (4): 319–32.
18. Chung, T., C. S. Martin, and K. C. Winters. 2005. Diagnosis, course, and assessment of alcohol abuse and dependence in adolescents. *Recent Developments in Alcoholism* 17:5–27.
19. Colby, S. M., S. T. Tiffany, S. Shiffman, and R. S. Niaura. 2000. Are adolescent smokers dependent on nicotine? A review of the evidence. *Drug and Alcohol Dependence* 59 Suppl. 1: S83–S95.
20. O'Loughlin, J., J. DiFranza, R. F. Tyndale, G. Meshefedjian, E. McMillan-Davey, P. B. Clarke, J. Hanley, and G. Paradis. 2003. Nicotine-dependence symptoms

are associated with smoking frequency in adolescents. *American Journal of Preventive Medicine* 25 (3): 219–25.

21. Kandel, D. B., M. C. Hu, P. C. Griesler, and C. Schaffran. 2007. On the development of nicotine dependence in adolescence. *Drug and Alcohol Dependence* 91 (1): 26–39.
22. Hughes, J. R. 2007. Effects of abstinence from tobacco: etiology, animal models, epidemiology, and significance: a subjective review. *Nicotine & Tobacco Research* 9 (3): 329–39.
23. Rojas, N. L., J. D. Killen, K. F. Haydel, and T. N. Robinson. 1998. Nicotine dependence among adolescent smokers. *Archives of Pediatrics & Adolescent Medicine* 152 (2): 151–56.
24. Killen, J. D., S. Ammerman, N. Rojas, J. Varady, F. Haydel, and T. N. Robinson. 2001. Do adolescent smokers experience withdrawal effects when deprived of nicotine? *Experimental and Clinical Psychopharmacology* 9 (2): 176–82.
25. Prokhorov, A. V., K. S. Hudmon, P. M. Cinciripini, and S. Marani. 2005. "Withdrawal symptoms" in adolescents: A comparison of former smokers and never-smokers. *Nicotine & Tobacco Research* 7 (6): 909–13.
26. Breslau, N., E. O. Johnson, E. Hiripi, and R. Kessler. 2001. Nicotine dependence in the United States: Prevalence, trends, and smoking persistence. *Archives of General Psychiatry* 58 (9): 810–16.
27. Gervais, A., J. O'Loughlin, G. Meshefedjian, C. Bancej, and M. Tremblay. 2006. Milestones in the natural course of onset of cigarette use among adolescents. *Canadian Medical Association Journal* 175 (3): 255–61.
28. DiFranza, J. R., J. A. Savageau, N. A. Rigotti, K. Fletcher, J. K. Ockene, A. D. McNeill, M. Coleman, and C. Wood. 2002. Development of symptoms of tobacco dependence in youths: 30 month follow up data from the DANDY study. *Tobacco Control* 11 (3): 228–35.
29. Dierker, L. C., E. Donny, S. Tiffany, S. M. Colby, N. Perrine, and R. R. Clayton. 2007. The association between cigarette smoking and DSM-IV nicotine dependence among first year college students. *Drug and Alcohol Dependence* 86 (2–3): 106–14.
30. Kandel, D. B., and K. Chen. 2000. Extent of smoking and nicotine dependence in the United States: 1991–1993. *Nicotine & Tobacco Research* 2 (3): 263–74.
31. O'Connor, R. J., G. A. Giovino, L. T. Kozlowski, S. Shiffman, A. Hyland, J. T. Bernert, R. S. Caraballo, and K. M. Cummings. 2006. Changes in nicotine intake and cigarette use over time in two nationally representative cross-sectional samples of smokers. *American Journal of Epidemiology* 164 (8): 750–5.
32. Chassin, L., C. C. Presson, S. J. Sherman, and K. Kim. 2003. Historical changes in cigarette smoking and smoking-related beliefs after 2 decades in a midwestern community. *Health Psychology* 22 (4): 347–53.
33. Levin, E. D., A. H. Rezvani, D. Montoya, J. E. Rose, and H. S. Swartzwelder. 2003. Adolescent-onset nicotine self-administration modeled in female rats. *Psychopharmacology (Berl)* 169 (2): 141–49.
34. Adriani, W., S. Macri, R. Pacifici, and G. Laviola. 2002. Peculiar vulnerability to nicotine oral self-administration in mice during early adolescence. *Neuropsychopharmacology* 27 (2): 212–14.
35. Belluzzi, J. D., A. G. Lee, H. S. Oliff, and F. M. Leslie. 2004. Age-dependent effects of nicotine on locomotor activity and conditioned place preference in rats. *Psychopharmacology (Berl)* 174 (3): 389–95.
36. Faraday, M. M., B. M. Elliott, and N. E. Grunberg. 2001. Adult vs. adolescent rats differ in biobehavioral responses to chronic nicotine administration. *Pharmacology, Biochemistry, and Behavior* 70 (4): 475–89.
37. Rezvani, A. H., and E. D. Levin. 2004. Adolescent and adult rats respond differently to nicotine and alcohol: Motor activity and body temperature. *International Journal of Developmental Neuroscience* 22 (5–6): 349–54.
38. O'Dell, L. E. , A. W. Bruijnzeel, S. Ghozland, A. Markou, and G. F. Koob. 2004. Nicotine withdrawal in adolescent and adult rats. *Annals of the New York Academy of Sciences* 1021:167–74.
39. Abreu-Villaca, Y., F. J. Seidler, D. Qiao, C. A. Tate, M. M. Cousins, I. Thillai, and T. A. Slotkin. 2003. Short-term adolescent nicotine exposure has immediate and persistent effects on cholinergic systems: Critical periods, patterns of exposure, dose thresholds. *Neuropsychopharmacology* 28 (11): 1935–49.

40. Olmstead, M. C. 2006. Animal models of drug addiction: Where do we go from here? *Quarterly Journal of Experimental Psychology (Colchester)* 59 (4): 625–53.

41. Spear, L. P., and E. I. Varlinskaya. 2005. Adolescence. Alcohol sensitivity, tolerance, and intake. *Recent Developments in Alcoholism* 17:143–59.

42. Steinberg, L. 2004. Risk taking in adolescence: What changes, and why? *Annals of the New York Academy of Sciences* 1021:51–58.

43. Jessor, R., and S. L. Jessor. 1977. *Problem behavior and psychosocial development: A longitudinal study of youth.* New York: Academic Press.

44. Schulenberg, J. E., A. J. Sameroff, and D. Cicchetti. 2004. The transition to adulthood as a critical juncture in the course of psychopathology and mental health. *Development and Psychopathology* 16 (4): 799–806.

45. Bachman, J. G., K. N. Wadsworth, P. M. O'Malley, L. D. Johnston, and J. E. Schulenberg. 1997. *Smoking, drinking, and drug use in young adulthood: The impacts of new freedoms and new responsibilities.* Mahwah, NJ: Lawrence Erlbaum.

46. Avenevoli, S., and K. R. Merikangas. 2003. Familial influences on adolescent smoking. *Addiction* 98 Suppl. 1: 1–20.

47. Conrad, K. M., B. R. Flay, and D. Hill. 1992. Why children start smoking cigarettes: Predictors of onset. *British Journal of Addiction* 87 (12): 1711–24.

48. Bauman, K. E., V. A. Foshee, M. A. Linzer, and G. G. Koch. 1990. Effect of parental smoking classification on the association between parental and adolescent smoking. *Addictive Behaviors* 15 (5): 413–22.

49. Hopfer, C. J., T. J. Crowley, and J. K. Hewitt. 2003. Review of twin and adoption studies of adolescent substance use. *Journal of the American Academy of Child & Adolescent Psychiatry* 42 (6): 710–19.

50. McGue, M., I. Elkins, and W. G. Iacono. 2000. Genetic and environmental influences on adolescent substance use and abuse. *American Journal of Medical Genetics* 96 (5): 671–77.

51. Han, C., M. K. McGue, and W. G. Iacono. 1999. Lifetime tobacco, alcohol and other substance use in adolescent Minnesota twins: Univariate and multivariate behavioral genetic analyses. *Addiction* 94 (7): 981–93.

52. Rhee, S. H., J. K. Hewitt, S. E. Young, R. P. Corley, T. J. Crowley, and M. C. Stallings. 2003. Genetic and environmental influences on substance initiation, use, and problem use in adolescents. *Archives of General Psychiatry* 60 (12): 1256–64.

53. Kendler, K. S., E. Schmitt, S. H. Aggen, and C. A. Prescott. 2008. Genetic and environmental influences on alcohol, caffeine, cannabis, and nicotine use from early adolescence to middle adulthood. *Archives of General Psychiatry* 65 (6): 642–82.

54. Koopmans, J. R., W. S. Slutske, G. C. van Baal, and D. I. Boomsma. 1999. The influence of religion on alcohol use initiation: Evidence for genotype X environment interaction. *Behavior Genetics* 29 (6): 445–53.

55. Rende, R., and I. Waldman. 2006. Behavioral and molecular genetics and developmental psychopathology. In *Developmental Psychopathology, vol. 1, theory and method,* 2nd ed., ed. D. Cicchetti and D. J. Cohen, 427–64. Hoboken, NJ: Wiley.

56. Chassin, L., C. C. Presson, S. C. Pitts, and S. J. Sherman. 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 (3): 223–31.

57. Maes, H. H., P. F. Sullivan, C. M. Bulik, M. C. Neale, C. A. Prescott, L. J. Eaves, and K. S. Kendler. 2004. A twin study of genetic and environmental influences on tobacco initiation, regular tobacco use and nicotine dependence. *Psychological Medicine* 34 (7): 1251–61.

58. White, V. M., J. L. Hopper, A. J. Wearing, and D. J. Hill. 2003. The role of genes in tobacco smoking during adolescence and young adulthood: A multivariate behaviour genetic investigation. *Addiction* 98 (8): 1087–1100.

59. Rende, R., C. Slomkowski, J. McCaffery, E. E. Lloyd-Richardson, and R. Niaura. 2005. A twin-sibling study of tobacco use in adolescence: Etiology of individual differences and extreme scores. *Nicotine & Tobacco Research* 7 (3): 413–19.

60. Farkas, A. J., E. A. Gilpin, M. M. White, and J. P. Pierce. 2000. Association between household and workplace smoking restrictions and adolescent smoking. *JAMA:*

The Journal of the American Medical Association 284 (6): 7171–22.

61. Boomsma, D. I., J. R. Koopmans, L. J. Van Doornen, and J. F. Orlebeke. 1994. Genetic and social influences on starting to smoke: A study of Dutch adolescent twins and their parents. *Addiction* 89 (2): 219–26.
62. Timberlake, D. S., S. H. Rhee, B. C. Haberstick, C. Hopfer , M. Ehringer, J. M. Lessem, A. Smolen, and J. K. Hewitt. 2006. The moderating effects of religiosity on the genetic and environmental determinants of smoking initiation. *Nicotine & Tobacco Research* 8 (1): 123–33.
63. Young, S. E., S. H. Rhee, M. C. Stallings, R. P. Corley, and J. K. Hewitt. 2006. Genetic and environmental vulnerabilities underlying adolescent substance use and problem use: General or specific? *Behavior Genetics* 36 (4): 603–15.
64. McGue, M., W. G. Iacono, and R. Krueger. 2006. The association of early adolescent problem behavior and adult psychopathology: A multivariate behavioral genetic perspective. *Behavior Genetics* 36 (4): 591–602.
65. Audrain-McGovern, J., C. Lerman, E. P. Wileyto, D. Rodriguez, and P. G. Shields. 2004. Interacting effects of genetic predisposition and depression on adolescent smoking progression. *American Journal of Psychiatry* 161 (7): 1224–30.
66. Anney, R. J., C. A. Olsson, M. Lotfi-Miri, G. C. Patton, and R. Williamson. 2004. Nicotine dependence in a prospective population-based study of adolescents: The protective role of a functional tyrosine hydroxylase polymorphism. *Pharmacogenetics* 14 (2): 73–81.
67. Olsson, C., R. Anney, S. Forrest, G. Patton, C. Coffey, T. Cameron, A. Hassett, and R. Williamson. 2004. Association between dependent smoking and a polymorphism in the tyrosine hydroxylase gene in a prospective population-based study of adolescent health. *Behavior Genetics* 34 (1): 85–91.
68. Laucht, M., K. Becker, M. El-Faddagh, E. Hohm, and M. H. Schmidt. 2005. Association of the DRD4 exon III polymorphism with smoking in fifteen-year-olds: A mediating role for novelty seeking? *Journal of the American Academy of Child & Adolescent Psychiatry* 44 (5): 477–84.
69. Skowronek, M. H., M. Laucht, E. Hohm, K. Becker, and M. H. Schmidt. 2006. Interaction between the dopamine D4 receptor and the serotonin transporter promoter polymorphisms in alcohol and tobacco use among 15-year-olds. *Neurogenetics* 7 (4): 239–46.
70. Gerra, G., L. Garofano, A. Zaimovic, G. Moi, B. Branchi, M. Bussandri, F. Brambilla, and C. Donnini. 2005. Association of the serotonin transporter promoter polymorphism with smoking behavior among adolescents. *American Journal of Medical Genetics Part B, Neuropsychiatric Genetics* 135 (1): 73–78.
71. O'Loughlin, J., G. Paradis, W. Kim, J. DiFranza, G. Meshefedjian, E. McMillan-Davey, S. Wong, J. Hanley, and R. F. Tyndale. 2004. Genetically decreased CYP2A6 and the risk of tobacco dependence: A prospective study of novice smokers. *Tobacco Control* 13 (4): 422–28.
72. Audrain-McGovern, J., N. Al Koudsi, D. Rodriguez, E. P. Wileyto, P. G. Shields, and R. F. Tyndale. 2007. The role of CYP2A6 in the emergence of nicotine dependence in adolescents. *Pediatrics* 119 (1): e264–e274.
73. Carter, B., T. Long, and P. Cinciripini. 2004. A meta-analytic review of the CYP2A6 genotype and smoking behavior. *Nicotine & Tobacco Research* 6 (2): 221–27.
74. Lessov, C. N., G. E. Swan, H. Z. Ring, T. V. Khroyan, and C. Lerman. 2004. Genetics and drug use as a complex phenotype. *Substance Use and Misuse* 39 (10–12): 1515–69.
75. Moffitt, T. E. 1993. Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review* 100 (4): 674–701.
76. Moffitt, T. E., A. Caspi, H. Harrington, and B. J. Milne. 2002. Males on the life-course-persistent and adolescence-limited antisocial pathways: Follow-up at age 26 years. *Development and Psychopathology* 14 (1): 179–207.
77. Breslau, N., and E. L. Peterson. 1996. Smoking cessation in young adults: Age at initiation of cigarette smoking and other suspected influences. *American Journal of Public Health* 86 (2): 214–20.
78. Riggs, N. R., C. P. Chou, C. Li, and M. A. Pentz. 2007. Adolescent to emerging adulthood smoking trajectories: When do smoking trajectories diverge, and do they predict early adulthood nicotine dependence? *Nicotine & Tobacco Research* 9 (11): 1147–54.

79. Bernat, D. H., D. J. Erickson, R. Widome, C. L. Perry, and J. L. Forster. 2008. Adolescent smoking trajectories: Results from a population-based cohort study. *Journal of Adolescent Health* 43 (4): 334–40.

80. Vitaro, F., B. Wanner, M. Brendgen, C. Gosselin, and P. L. Gendreau. 2004. Differential contribution of parents and friends to smoking trajectories during adolescence. *Addictive Behaviors* 29 (4): 831–35.

81. Broms, U., K. Silventoinen, P. A. Madden, A. C. Heath, and J. Kaprio. 2006. Genetic architecture of smoking behavior: A study of Finnish adult twins. *Twin Research and Human Genetics* 9 (1): 64–72.

82. Heath, A. C., K. M. Kirk, J. M. Meyer, and N. G. Martin. 1999. Genetic and social determinants of initiation and age at onset of smoking in Australian twins. *Behavior Genetics* 29 (6): 395–407.

83. Ling, D., T. Niu, Y. Feng, H. Xing, and X. Xu. 2004. Association between polymorphism of the dopamine transporter gene and early smoking onset: An interaction risk on nicotine dependence. *Journal of Human Genetics* 49 (1): 35–39.

84. Upadhyaya, H. P., D. Deas, K. T. Brady, and M. Kruesi. 2002. Cigarette smoking and psychiatric comorbidity in children and adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry* 41 (11): 1294–1305.

85. Lambert, N. M., and C. S. Hartsough. 1998. Prospective study of tobacco smoking and substance dependencies among samples of ADHD and non-ADHD participants. *Journal of Learning Disabilities* 31 (6): 533–44.

86. Molina, B. S., and W. E. Pelham Jr. 2003. Childhood predictors of adolescent substance use in a longitudinal study of children with ADHD. *Journal of Abnormal Psychology* 112 (3): 497–507.

87. Milberger, S., J. Biederman, S. V. Faraone, L. Chen, and J. Jones. 1997. ADHD is associated with early initiation of cigarette smoking in children and adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry* 36 (1): 37–44.

88. Flory, K., and D. R. Lynam. 2003. The relation between attention deficit hyperactivity disorder and substance abuse: What role does conduct disorder play? *Clinical Child and Family Psychology Review* 6 (1): 1–16.

89. Costello, E. J., A. Erkanli, E. Federman, and A. Angold. 1999. Development of psychiatric comorbidity with substance abuse in adolescents: Effects of timing and sex. *Journal of Clinical Child Psychology* 28 (3): 298–311.

90. Kaplow, J. B., P. J. Curran, A. Angold, and E. J. Costello. 2001. The prospective relation between dimensions of anxiety and the initiation of adolescent alcohol use. *Journal of Clinical Child Psychology* 30 (3): 316–26.

91. Johnson, E. O., and L. Schultz. 2005. Forward telescoping bias in reported age of onset: An example from cigarette smoking. *International Journal of Methods in Psychiatric Research* 14 (3): 119–29.

92. Parra, G. R., S. E. O'Neill, and K. J. Sher. 2003. Reliability of self-reported age of substance involvement onset. *Psychology of Addictive Behaviors* 17 (3): 211–18.

93. Yoon, H. H., W. G. Iacono, S. M. Malone, and M. McGue. 2006. Using the brain P300 response to identify novel phenotypes reflecting genetic vulnerability for adolescent substance misuse. *Addictive Behaviors* 31 (6): 1067–87.

94. Donny, E. C., S. T. Lanza, R. L. Balster, L. M. Collins, A. Caggiula, and P. P. Rowell. 2004. Using growth models to relate acquisition of nicotine self-administration to break point and nicotinic receptor binding. *Drug and Alcohol Dependence* 75 (1): 23–35.

95. Eissenberg, T., and R. L. Balster. 2000. Initial tobacco use episodes in children and adolescents: Current knowledge, future directions. *Drug and Alcohol Dependence* 59 Suppl. 1: S41–S60.

96. Chen, X., A. Stacy, H. Zheng, J. Shan, D. Spruijt-Metz, J. Unger, J. Gong, et al. 2003. Sensations from initial exposure to nicotine predicting adolescent smoking in China: A potential measure of vulnerability to nicotine. *Nicotine & Tobacco Research* 5 (4): 455–63.

97. DiFranza, J. R., J. A. Savageau, K. Fletcher, J. K. Ockene, N. A. Rigotti, A. D. McNeill, M. Coleman, and C. Wood. 2004. Recollections and repercussions of the first inhaled cigarette. *Addictive Behaviors* 29 (2): 261–72.

98. Hirschman, R. S., H. Leventhal, and K. Glynn. 1984. The development of smoking behavior: Conceptualization and supportive cross-sectional survey data. *Journal of Applied Social Psychology* 14 (3): 184–206.

99. Hu, M. C., B. Muthén, C. Schaffran, P. C. Griesler, and D. B. Kandel. 2008. Developmental trajectories of criteria of nicotine dependence in adolescence. *Drug and Alcohol Dependence* 98 (1–2): 94–104.

100. Abroms, L., B. Simons-Morton, D. L. Haynie, and R. Chen. 2005. Psychosocial predictors of smoking trajectories during middle and high school. *Addiction* 100 (6): 852–61.

101. Colder, C. R., P. D. Mehta, K. Balanda, R. T. Campbell, K. Mayhew, W. R. Stanton, M. Pentz, and B. R. Flay. 2001. Identifying trajectories of adolescent smoking: An application of latent growth mixture modeling. *Health Psychology* 20 (2): 127–35.

102. Juon, H. S., M. E. Ensminger, and K. D. Sydnor. 2002. A longitudinal study of developmental trajectories to young adult cigarette smoking. *Drug and Alcohol Dependence* 66 (3): 303–14.

103. Soldz, S., and X. Cui. 2002. Pathways through adolescent smoking: A 7-year longitudinal grouping analysis. *Health Psychology* 21 (5): 495–504.

104. White, H. R., R. J. Pandina, and P. H. Chen. 2002. Developmental trajectories of cigarette use from early adolescence into young adulthood. *Drug and Alcohol Dependence* 65 (2): 167–78.

105. Audrain-McGovern, J., D. Rodriguez, K. P. Tercyak, J. Cuevas, K. Rodgers, and F. Patterson. 2004. Identifying and characterizing adolescent smoking trajectories. *Cancer Epidemiology, Biomarkers & Prevention* 13 (12): 2023–34.

106. Orlando, M., J. S. Tucker, P. L. Ellickson, and D. J. Klein. 2004. Developmental trajectories of cigarette smoking and their correlates from early adolescence to young adulthood. *Journal of Consulting and Clinical Psychology* 72 (3): 400–410.

107. Stanton, W. R., B. R. Flay, C. R. Colder, and P. Mehta. 2004. Identifying and predicting adolescent smokers' developmental trajectories. *Nicotine & Tobacco Research* 6 (5): 843–52.

108. White, H. R., D. Nagin, E. Replogle, and M. Stouthamer-Loeber. 2004. Racial differences in trajectories of cigarette use. *Drug and Alcohol Dependence* 76 (3): 219–27.

109. Karp, I., J. O'Loughlin, G. Paradis, J. Hanley, and J. DiFranza. 2005. Smoking trajectories of adolescent novice smokers in a longitudinal study of tobacco use. *Annals of Epidemiology* 15 (6): 445–52.

110. Brook, J. S., K. Pahl, and Y. Ning. 2006. Peer and parental influences on longitudinal trajectories of smoking among African Americans and Puerto Ricans. *Nicotine & Tobacco Research* 8 (5): 639–51.

111. Maggi, S., C. Hertzman, and T. Vaillancourt. 2007. Changes in smoking behaviors from late childhood to adolescence: Insights from the Canadian National Longitudinal Survey of Children and Youth. *Health Psychology* 26 (2): 232–40.

112. Maggi, S. 2008. Changes in smoking behaviours from late childhood to adolescence: 4 years later. *Drug and Alcohol Dependence* 94 (1–3): 251–53.

113. Lessov-Schlaggar, C. N., H. Hops, J. Brigham, K. S. Hudmon, J. A. Andrews, E. Tildesley, D. McBride, L. M. Jack, H. S. Javitz, and G. E. Swan. 2008. Adolescent smoking trajectories and nicotine dependence. *Nicotine & Tobacco Research* 10 (2): 341–51.

114. Wills, T. A., J. A. Resko, M. G. Ainette, and D. Mendoza. 2004. Smoking onset in adolescence: A person-centered analysis with time-varying predictors. *Health Psychology* 23 (2): 158–67.

115. Zucker, R. A. 1986. The four alcoholisms: A developmental account of the etiologic process. *Nebraska Symposium on Motivation* 34:27–83.

116. Blitstein, J. L., L. A. Robinson, D. M. Murray, R. C. Klesges, and S. M. Zbikowski. 2003. Rapid progression to regular cigarette smoking among nonsmoking adolescents: Interactions with gender and ethnicity. *Preventive Medicine* 36 (4): 455–63.

117. Curran, P. J., and M. T. Willoughby. 2003. Implications of latent trajectory models for the study of developmental psychopathology. *Development and Psychopathology* 15 (3): 581–612.

118. Wohlwill, J. F. 1991. Relations between method and theory in developmental research: A partial–isomorphism view. In *Annals of Theoretical Psychology*, vol. 7, ed. P. van Geert and L. P. Mos, 91–138.

119. Rogosa, D. R., and J. B. Willett. 1985. Understanding correlates of change by modeling individual differences in growth. *Psychometrika* 50 (2): 203–28.

120. Bauer, D. J. 2003. Estimating multilevel linear models as structural equation models. *Journal of Educational and Behavioral Statistics* 28 (2): 135–67.

121. Curran, P. J. 2003. Have multilevel models been structural equation models all along? *Multivariate Behavioral Research* 38 (4): 529–69.

122. Mehta, P. D., and M. C. Neale. 2005. People are variables too: Multilevel structural equations modeling. *Psychological Methods* 10 (3): 259–84.

123. Raudenbush, S. W. 2001. Toward a coherent framework for comparing trajectories of individual change. In *New methods for the analysis of change*, ed. L. M. Collins and A. G. Sayer, 35–64. Washington, DC: American Psychological Association.

124. Willett, J. B., and A. G. Sayer. 1994. Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin* 116 (2): 363–81.

125. Bryk, A. S., and S. W. Raudenbush. 1987. Application of hierarchical linear models to assessing change. *Psychological Bulletin* 101 (1): 147–58.

126. Meredith, W., and J. Tisak. 1990. Latent curve analysis. *Psychometrika* 55 (1): 107–22.

127. McArdle, J. J. 1988. Dynamic but structural equation modeling of repeated measures data. In *The handbook of multivariate experimental psychology*, vol. 2, ed. J. R. Nesselroade and R. B. Cattell, 561–614. New York: Plenum Press.

128. McArdle, J. J. 1989. Structural modeling experiments using multiple growth functions. In *Learning and individual differences: Abilities, motivation and methodology*, ed. P. Ackerman, R. Cudeck, and R. Kanfer, 71–117. Hillsdale, NJ: Lawrence Erlbaum.

129. McArdle, J. J. 1991. Structural models of developmental theory in psychology. In *Annals of Theoretical Psychology*, vol. VII, ed. P. Van Geert and L. P. Mos, 139–60. New York: Plenum Press.

130. Bollen, K. A., and P. J. Curran. 2005. *Latent curve models: A structural equation perspective.* Hoboken, NJ: Wiley.

131. Muthén, B. O., and P. J. Curran. 1997. General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods* 2 (4): 371–702.

132. Muthén, B. O., and K. Shedden. 1999. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* 55 (2): 463–69.

133. Nagin, D. S. 1999. Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods* 4 (2): 139–57.

134. Bauer, D. J., and P. J. Curran. 2003. Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods* 8 (3): 338–63.

135. Bauer, D. J., and P. J. Curran. 2004. The integration of continuous and discrete latent variable models: potential problems and promising opportunities. *Psychological Methods* 9 (1): 3–29.

136. Muthén, B. O. 2004. Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In *Handbook of quantitative methodology for the social sciences*, ed. D. Kaplan, 345–68. Newbury Park, CA: Sage Publications.

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

138. Muthén, B. O. 1989. Latent variable modeling in heterogeneous populations. *Psychometrika* 54:557–85.

139. MacQueen, J. B. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 281–97. Berkeley, CA: Univ. of California Press.

140. Clogg, C. C. 1995. Latent class models. In *Handbook of statistical modeling for the social and behavioral sciences*, ed. G. Arminger, C. C. Clogg, and M. E. Sobel, 311–59. New York: Plenum.

141. Lazarsfeld, P. F., and N. W. Henry. 1968. *Latent structure analysis.* Boston, MA: Houghton Mifflin.

142. McLachlan, G., and D. Peel. 2000. *Finite mixture models.* New York: John Wiley and Sons.

143. Arminger, G., and P. Stein. 1997. Finite mixtures of covariance structure models with regressors. *Sociological Methods & Research* 26 (2): 148–82.

144. Arminger, G., P. Stein, and J. Wittenberg. 1999. Mixtures of conditional mean- and covariance-structure models. *Psychometrika* 64 (4): 475–94.

145. Dolan, C. V., and H. L. J. van der Maas. 1998. Fitting multivariate normal finite mixtures

subject to structural equation modeling. *Psychometrika* 63 (3): 227–53.

146. Gibson, W. A. 1959. Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. *Psychometrika* 24 (3): 229–52.
147. Jedidi, K., H. S. Jagpal, and W. S. DeSarbo. 1997. Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science* 16 (1): 39–59.
148. Pearson, K. 1894. Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London A* 185:71–110.
149. Yung, Y.-F. 1997. Finite mixtures in confirmatory factor-analysis models. *Psychometrika* 62 (3): 297–330.
150. Nagin, D. S., and K. C. Land. 1993. Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson model. *Criminology* 31 (3): 327–62.
151. Nagin, D. S., and R. E. Tremblay. 2001. Analyzing developmental trajectories of distinct but related behaviors: A group-based method. *Psychological Methods* 6 (1): 18–34.
152. Lazarsfeld, P. F. 1950. Some latent structures. In *The American soldier: Studies in social psychology in World War II*, vol. 4, ed. S. A. Stouffer. Princeton, NJ: Princeton Univ. Press.
153. Muthén, B. O. 2001. Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In *New methods for the analysis of change*, ed. L. M. Collins and A. Sayer, 291–322. Washington, DC: American Psychological Association.
154. Muthén, B. O. 2002. Beyond SEM: General latent variable modeling. *Behaviormetrika* 29 (1): 81–117.
155. Bauer, D. J., and P. J. Curran. 2003. Overextraction of latent trajectory classes: Much ado about nothing? *Psychological Methods* 8 (3): 384–93.
156. Muthén, B. O. 2003. Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran (2003). *Psychological Methods* 8 (3): 369–77; discussion 384–93.
157. Nagin, D. S., and R. E. Tremblay. 2005. Developmental trajectory groups: Fact or a useful statistical fiction. *Criminology* 43 (4): 873–904.
158. Beauchaine, T. P., and E. Waters. 2003. Pseudotaxonicity in MAMBAC and MAXCOV analyses of rating-scale data: Turning continua into classes by manipulating observer's expectations. *Psychological Methods* 8 (1): 3–15.
159. Markon, K. E., and R. F. Krueger. 2006. Information-theoretic latent distribution modeling: Distinguishing discrete and continuous latent variable models. *Psychological Methods* 11 (3): 228–43.
160. Meehl, P. E. 1967. Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science* 34 (2): 103–15.
161. Meehl, P. E. 1968. *Detecting latent clinical taxa, II: A simplified procedure, some additional hitmax cut locators, a single-indicator method, and miscellaneous theorems* (Report No. PR-68-2). Minneapolis, MN: Univ. of Minnesota, Research Laboratories of the Department of Psychiatry.
162. Meehl, P. E. 1992. Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality* 60 (1): 117–74.
163. Waller, N. G., and P. E. Meehl. 1998. *Multivariate taxometric procedures: Distinguishing types from continua.* Thousand Oaks, CA: Sage.
164. Cudeck, R., and S. J. Henly. 2003. A realistic perspective on pattern representation in growth data: Comment on Bauer and Curran (2003). *Psychological Methods* 8 (3): 378–83; discussion 384–93.
165. Shaw, P., D. Greenstein, J. Lerch, L. Clasen, R. Lenroot, N. Gogtay, A. Evans, J. Rapoport, and J. Giedd. 2006. Intellectual ability and cortical development in children and adolescents. *Nature* 440 (7084): 676–79.
166. MacCallum, R. C. 2003. 2001 presidential address: Working with imperfect models. *Multivariate Behavioral Research* 38 (1): 113–39.
167. Eggleston, E. P., J. H. Laub, and R. J. Sampson. 2004. Methodological sensitivities to latent class analysis of long-term criminal trajectories. *Journal of Quantitative Criminology* 20 (1): 1–26.
168. Jackson, K. M., and K. J. Sher. 2006. Comparison of longitudinal phenotypes based on number and timing of assessments: A systematic comparison of trajectory

approaches II. *Psychology of Addictive Behaviors* 20 (4): 373–84.

169. Olsen, M. K., and J. L. Schafer. 2001. A two-part random-effects model for semicontinuous longitudinal data. *Journal of the American Statistical Association* 96 (454): 730–45.
170. Chassin, L., C. C. Presson, M. Bensenberg, E. Corty, R. W. Olshavsky, and S. J. Sherman. 1981. Predicting adolescents' intentions to smoke cigarettes. *Journal of Health and Social Behavior* 22 (4): 445–55.
171. Chassin, L., C. C. Presson, S. J. Sherman, and D. A. Edwards. 1990. The natural history of cigarette smoking: Predicting young-adult smoking outcomes from adolescent smoking patterns. *Health Psychology* 9 (6): 701–16.
172. Lugaila, T. A. 1998. Marital status and living arrangements: March 1998 (update). Current Population Reports, vol. P20-514. Washington, DC: U.S. Government Printing Office. http://www.census.gov/prod/99pubs/p20-514.pdf.
173. Day, J. C., and A. E. Curry. 1998. Educational attainment in the United States: March 1998 (update). Current Population Reports, P20–513. Washington, DC: U.S. Government Printing Office. http://www.census.gov/prod/3/98pubs/p20-513.pdf.
174. Holtzman, D., E. Powell-Griner, J. C. Bolen, and L. Rhodes. 2000. State- and sex-specific prevalence of selected characteristics—Behavioral Risk Factor Surveillance System, 1996 and 1997. *Morbidity and Mortality Weekly Report* 49 (6): 1–39.
175. Rose, J. S., L. Chassin, C. C. Presson, and S. J. Sherman. 1996. Demographic factors in adult smoking status: Mediating and moderating influences. *Psychology of Addictive Behaviors* 10 (1): 28–37.
176. Jones, B. L., D. S. Nagin, and K. Roeder. 2001. A SAS procedure based on mixture models for estimating developmental trajectories. *Sociological Methods & Research* 29 (3): 374–93.
177. Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6 (2): 461–64.
178. Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. In *Proceedings of the Second International Symposium on Information Theory*, ed. B. N. Petrov and F. Csaki, 267–81. Budapest: Akademiai Kiado.
179. Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19 (6): 716–23.
180. Dolan, C. V., V. D. Schmittmann, G. H. Lubke, and M. C. Neale. 2005. Regime switching in the latent growth curve mixture model. *Structural Equation Modeling* 12 (1): 94–119.