Running head: Academic Achievement and Adolescent Smoking

Academic Achievement and Smoking Initiation in Adolescence: A General Growth Mixture Analysis.

Alexandre J.S. Morin *, Ph.D.
Department of Psychology, University of Sherbrooke, Canada
and Educational Excellence and Equity (E3) Research Program, Center for Educational Research
(University of Western Sydney)

Daniel Rodriguez *, Ph.D.
University of Pennsylvania

Jean-Sébastien Fallu, Ph.D.
School Environment Research Group, School of Psychoeducation, University of Montreal

Christophe Maïano, Ph.D.
Cyberpsychology Laboratory, Department of Psychoeducation and Psychology, University of Quebec in Outaouais, Canada

Michel Janosz, Ph.D.
School Environment Research Group, School of Psychoeducation, University of Montreal

* The first two authors (A.J.S.M. and D.R.) contributed equally to this article and their order was determined at random: both should thus be considered first authors.

This is the prepublication version of:

Morin, A.J.S., Rodriguez, D., Fallu, J.-S., **Maïano, C.,** & Janosz, M. (2012). Academic Achievement and Adolescent Smoking: A General Growth Mixture Model. *Addiction*, *107*, 819-828.

The definitive version is available at: DOI: 10.1111/j.1360-0443.2011.03725.x

Corresponding author:
Alexandre J.S. Morin
University of Sherbrooke
Department of Psychology
2500 boulevard de l'Université
Sherbrooke, QC J1K 2R1 Canada
E-mail: Alexandre.Morin@USherbrooke.ca

Abstract

Aims. This study aims to: (a) explore the relations between smoking initiation and different profiles of academic achievement trajectories in early to mid-adolescence; (b) to investigate whether background characteristics (gender, ethnicity, grade repetition, parental education) and proximal processes (parental practices, extracurricular involvement) predicted class membership and smoking initiation.

Design: Four year longitudinal cohort study (7th to 10th grade).

Setting: Adolescents completed the questionnaires during school hours.

Participants. 741 adolescents with no history of smoking in grade 7 participating in the Montreal Adolescent Depression Development Project.

Measurements. Self-report questionnaires were used to assess predictors and previous smoking in year 1, and smoking initiation by the end of the study. Grade point average (GPA) was obtained twice yearly from school records.

Results. Three academic achievement trajectories were identified and found to significantly differ in rates of smoking initiation: persistently high achievers (7.1% smoking), average achievers (15.1% smokers), and unstable low achievers (49.1% smoking). Further, results showed that general parenting practices and parental education indirectly reduced the likelihood of smoking by reducing the risk of membership in classes with lower GPA.

Conclusions. Adolescents who do well in school are less likely to smoke and it may be cost-effective for smoking prevention to focus on the few (12%) easy to identify unstable low achievers who form 35% of smoking onsets. In addition as parental support and democratic control reduced the likelihood of poor academic performance, promoting essential generic parenting skills from a young age may also prevent future onsets of smoking in adolescence.

Keywords. Smoking, onset, adolescence, grade point average, GPA, growth mixture, parenting, trajectories, latent classes.

Tobacco use, typically initiated in adolescence [1-2], is the leading cause of preventable death worldwide [3]. Thus, there is a critical need to identify factors that reduce the likelihood of smoking initiation in adolescence. One such factor is academic achievement [4-12]. Having a higher grade point average (GPA) was found to predict a 50% decrease in the odds of smoking experimentation, and a 75% decrease in the odds of being an early smoker [4] even after controlling for many smoking risk factors [13, 14]. This suggests that one efficient way to reduce adolescent smoking is to increase academic performance.

Targeting prevention efforts solely at increasing GPA without also intervening on social factors affecting GPA would likely, however, bear little fruit. Social factors such as parenting style and extracurricular involvement have been found to relate to both GPA and smoking in adolescence [13, 15-21]. Thus, GPA may mediate the relation between social factors that both increase the propensity to do well in school and not smoking [14]. Indeed, research supports GPA as a mediator in the relation between factors such as parental education, and smoking [13]. Accordingly, targeting GPA in smoking prevention programs may prove less fruitful than targeting factors increasing the risk of underachievement. For instance, generic parenting skills such as parental support and democratic control are relatively stable characteristics of the parent-youth relationship [20-23] known to positively impact many important developmental outcomes [24-27], along with GPA and smoking. Consequently, early prevention programs targeting parenting skills may have lasting effects on youth development. Conversely, it has been argued that it may be even more important to teach domain-specific academic parenting skills if the objective is to improve GPA, such as providing academic support without unduly pressuring youth to perform [22, 28-32]. Finally, parents can motivate adolescents to get involved in extracurricular activities where interactions with prosocial peers and adult models may help adolescents to value education and avoid smoking or other unhealthy behaviors [33-37].

Although GPA may mediate the relation between social factors and smoking, GPA trajectories present a substantial level of developmental heterogeneity and inter-individual variability. Part of this heterogeneity may be explained by the presence of subpopulations in which the profiles of associations between smoking, GPA and predictors differ. Investigating this possibility is important for many reasons. First, the predictors of membership into a specific subgroup (e.g., low versus increasing achievers) may differ from the predictors of fluctuations in GPA levels over time across all individuals. Thus, membership into GPA subgroups may itself mediate the relationship between predictors and smoking if the subgroups are also found to differ in smoking initiation rates. Likewise, the relation between these predictors, GPA, and smoking may also differ by subgroup. Indeed, different factors may influence fluctuations in GPA over time and smoking for a class of persistently high achievers, than for low achievers. For example, parental academic support may explain fluctuations in GPA levels for low achievers struggling to perform but not for high achievers who may not need such support. Similarly, high achievers may be more likely to come from families were they can discuss openly personal issues and thus may be less impacted by exposure to prosocial peers and adults in extracurricular activities. By contrast, low achievers may come from families characterized by harsher control and may thus benefit from opportunities to openly discuss issues and be exposed to prosocial peers in extracurricular activities. This would represent a form of moderated mediation [38], where the mediated relationship between predictors, GPA, and smoking will change according to the student's latent trajectory class (subgroup).

The purpose of the present study is fourfold. First, we test for the presence of multiple latent trajectory classes of adolescents with different patterns of GPA across grades 7 through 10 (ages 12 to 15). Using general growth mixture modeling (GGMM), we explore the number of latent classes of GPA trajectories, what factors characterize latent class membership, and whether latent classes differ on smoking initiation at the end of grade 10. Second, we assess whether several background variables (gender, nationality, grade repetition, low parental education) and proximal processes in the interpersonal context of the developing youth (parenting skills, extracurricular involvement) have direct associations with class membership. Third, we verify whether these factors also have direct associations with within-class fluctuations in GPA and smoking initiation and whether these relations differ by class membership. To our knowledge, this is the first study to assess the prospective relation between trajectories of GPA and smoking, and factors that may influence this relation. Although research suggests two to four latent classes of math and English achievement [39-41], no study to date has assessed the relation between GPA trajectory classes and smoking. Furthermore, this is one of few studies to use an objective measure of GPA assessed from school records [8, 42]. Finally, although research suggests that poor academic performance predicts smoking [4, 43], it is also possible that prior smoking increases the risk of poor performance [42]. To control for that possibility in a preventive perspective, this study targets non-smokers in order to predict smoking initiation.

Methods

Participants and Procedures

Participants were 741 non-smoking secondary school adolescents (54% male), aged 12 (SD= .56) at 7th grade baseline, taking part in the Montreal Adolescent Depression Development Project (MADDP) [44, 45]. This project was initially designed as a one-year follow-up study, with three measurement points in fall, winter and summer 2000-2001. All seventh-grade students from five Montreal-area secondary schools were recruited, right after the school transition. A total of 1370 participants (88.2% of the eligible students) consented to participate, completed Time 1 measures and at least one follow-up. These participants were contacted in 2001-2002 to participate in a longer-term study comprising three additional yearly measurements (Time 4-6, with Time 4 one year after Time 2). From those, 1034 (75.7%) were included in the longer follow-up study (58 failed to consent; 142 were impossible to locate; 136 were excluded by parental refusal), and 741 (71.7%) had never smoked and form the sample used in this study.

Measures

Controls. Sex (0=Male, 1=Female) and nationality [0= French/English speaking youths of North American Decent (NAD; 89% of the sample) and 1=non-NAD] were obtained from school records. Parental education is an average of maternal and paternal reports of their education level. Missing data were imputed from adolescents' self-reports. Grade repetition in elementary school was assessed through self-reports (0=never, 1=yes).

Extracurricular involvement. Extracurricular involvement was assessed at Time 1 using a self-reported two-item scale (e.g. "I spend many hours a week in extracurricular activities"). This scale was validated on a population sample of nearly 35,000 Canadian adolescents from the Quebec Province and was found to present adequate validity and scale-score reliability (α = .76) [46].

Parenting. Parenting was assessed from adolescents' self-reports at Time 1. Parental support and democratic control were assessed with 26 items forming two scales validated on a representative sample of Quebec adolescents and found to present adequate validity and reliability (α = .82-.95) [47]. Parental support assesses the support provided by the parents to the adolescent, parental awareness of their adolescent's activities and the quality of the parent-adolescents relations (15 items; e.g. "How often do your parents support you or praise you for things you have done?"). Parental democratic control assesses the presence of rules at home, fairness of these rules, the possibility to discuss rules and whether rules can be enforced without relying on harsh punishment (11 items; e.g. "In your home, is there a rule concerning how often you can go out with your friends?"). Parental academic support was evaluated with 7 items (e.g. "I can count on my parents when I have school difficulties") [46] and parental academic pressure with 4 items (e.g. "When I get low grades, my parents make me feel guilty") [45]. Validation studies showed that these scales have adequate validity and reliability (α =.70-.84) [45, 46].

Grade point average (GPA). Semester-specific GPA was obtained from school records at the beginning and end of each school year across four years, on a 1-100 scale as per norms in vigor in the Canadian Province of Quebec, with 60 as the achievement threshold.

Smoking. Smoking was assessed at the end of the study using a self-reported item measuring the frequency of smoking (from never smoked to over 20 cigarettes each day), recoded into a binary indicator of smoking. A similar item was used at Time 1 to eliminate youths with a smoking history. Youths with smoking experiences, regardless of frequency, were considered smokers.

Analyses

Analyses were performed using Mplus 5.1 [48]. Models with 1 to 7 latent trajectory classes were estimated with intercepts, linear/quadratic slopes, time-specific residuals and distal outcome (smoking) freely estimated in all classes. Predictors were directly integrated to the final model. A baseline model predicting only class membership was first estimated. Then, models in which predictors were allowed to influence GPA fluctuations and the outcome, and in which these effects were allowed to vary across classes were estimated. These models were compared to select the model providing optimal representation of the data. Additional technical details, as well as variable correlations and descriptive statistics are reported in the online supplements.

Results

Developmental Trajectories

The results converged on a three class solution (see Figure 1a): Elevated (n = 267), Moderate (n = 388), and Low-unstable (n = 86). The elevated class (Figure 1b) was characterized by an elevated level of GPA (intercept = 83.73) showing a slight but significant average decreasing trend over time (linear slope = -0.28, p = .045) and a non-significant average quadratic trend (quadratic slope = -0.004, $p \geq .05$). This average trend appeared common to most members of this class as the within-class variability of these estimates (expressed in SD units) was low, albeit significantly different from zero (intercept S.D. = 3.74, $p \leq .01$; linear slope SD = 0.89, $p \leq .01$; quadratic slope SD = 0.13, $p \leq .01$). These estimates indicate that there is significant variability within each class, i.e. that members of a class differ from one another, while remaining within class boundaries. In addition, most class members follow continuous linear or slightly curvilinear trends as seen from low standardized time-specific residuals (reflecting deviations in GPA SD units of the observed measure from the modeled linear/curvilinear trajectory), varying from 1.88 to 3.75 GPA points (yet still significant, $p \leq .01$, suggesting that model estimates include some predictive error).

The moderate class (Figure 1c) was characterized by a moderate GPA level (intercept = 74.74) showing a significant average decreasing trend (linear slope = -0.62, $p \leq .01$) and a non-significant quadratic trend (quadratic slope = -0.03, $p \geq .05$). In addition, greater variability was observed than in the elevated class, especially for the estimated intercepts (intercept SD = 5.44, $p \leq .01$; linear slope SD = 1.85, $p \leq .01$; quadratic slope SD = 0.28, $p \leq .01$), meaning that students from this class present moderately low to moderately high stable levels of GPA. Although time-specific residuals remained significant, indicating the presence of some predictive error, they remained small (3.27 to 4.89, $p \leq .01$), suggesting that most class members did follow reasonably linear/curvilinear trajectories.

Finally, the low-unstable class (Figure 1d) was the most peculiar. Indeed, although students from this class presented generally low GPA levels (intercept = 65.19; with significant variability SD = 9.55, $p \leq .01$) that appeared stable over time (non-significant linear slope = -1.31, $p \geq .05$; SD = 4.22, $p \geq .05$; non-significant quadratic slope = 0.14, $p \geq .05$; SD = 0.45, $p \geq .05$), these results are misleading given the significant and elevated levels of time-specific residuals observed from the second to eighth waves (7.93 to 17.08, $p \leq .01$). This indicates that this class mostly included students presenting highly unstable GPAs that cannot be adequately represented by continuous, linear, or quadratic trends. The fact that the residuals observed in the first wave were non-significantly different from zero does not contradict this conclusion as inter-individual differences on the initial time-point in unstable trajectories are generally embedded into the intercept.

Perhaps the strongest evidence in favor of the added-value of this GGMM model comes from rates of smoking initiation. Indeed, the proportion of smokers was significantly different between the three latent trajectory classes ($p \leq .01$) with 7.1% of smoking onsets in the elevated class, 15.1% in the moderate class, and 49.1% in the low-unstable class. This means that by targeting the small proportion (11.61%) of low-unstable achievers who have a 50% chance of initiating smoking in adolescence, more than a third of future smokers could be identified. Interestingly, this model resulted in a 30 to 40% increase in the percentage of explained variance in smoking onset relative to classical variable-centered analyses (see online supplements).

Conditional models

Predictors were then included to the model and significantly predicted class membership, intercept, linear slope and distal outcome (see Tables 1 and 2). These predictions did not change as a function of class membership (see online supplements), thus not supporting the moderated mediation hypothesis. However, the proposal that class membership did fully mediate the association between some predictors and smoking is supported. Indeed (1) class membership was significantly related to smoking initiation as shown in the previous section and (2) predictors (grade repetition, parental education, parental support and parental control) with significant zero-order correlations with the outcome and GPA levels (see online supplements) had direct effects on class membership but did not predict the outcome once their effect on class membership was considered.

More specifically, the results show that females were more likely to be members of the elevated versus moderate class, whereas nationality had no effect on class membership. Grade repetition and low parental education increased the risk of membership in classes with lower GPA. Parental support and democratic control were both associated with a greater likelihood of membership in the elevated class. Interestingly, once the effects of generic parental practices were taken into account, parental academic practices and extracurricular involvement had no remaining effect on class membership. The results also show that once class membership is considered, few predictors have direct residual effects on within-class GPA variability and smoking. Furthermore, being a female is associated with greater linear increases in GPA, and with an increased risk of smoking initiation. Previous results showed that females had a higher probability of membership in the high-GPA class that included the fewest smokers. The current results show that females are also at higher risk of smoking than males with similar GPA levels. Finally, higher parental education also predicted higher initial GPA levels above its effect on membership in the moderate and high classes.

Discussion

This study examined the presence of multiple GPA trajectories in adolescence, and the association between these trajectory and later tobacco use. Consistent with previous studies, our results revealed three latent trajectories [39-41]. The first one included persistently high achievers (36%) and a very low proportion of smokers (7%). The second included average achievers whose GPA followed relatively stable trends (52%) and a moderately low proportion of smokers (15%). The third included low achievers presenting highly unstable GPA trajectories (12%). This small class included a very high proportion of smokers (49%). Thus, these person-centered analyses confirmed the GPA-smoking association, and improved the proportion of explained variance in smoking by 30-40%.

These results have important practical implications. They suggest that targeting adolescents at elevated risks for multiple adaptation problems due to low/unstable GPA could reach 35% of future smokers. However, identification of this subpopulation requires longitudinal assessments since their main characteristic is unstable GPA levels that are low on the *average*, but not necessarily all the time. Clearly, this study cannot support causal interpretations of these associations despite the longitudinal design. Indeed, multiple factors related to adolescents' psychosocial backgrounds are most likely at play in determining adolescents' involvement in academic activities and smoking. However, these results have heuristic value as GPA is probably the most easily accessible indicator of psychosocial adaption for school systems. Further, the observation that these students present unstable trajectories, rather than systematically low trajectories, suggests that their achievement may be more reactive to the environment and modifiable through preventive efforts. Even in the moderate class significant inter-individual variability was observed, suggesting that even if class membership is established from Year 1, it remains possible to help students reach higher GPAs overlapping those from the elevated class.

Several predisposing factors were also found to have direct relations with class membership and within-class variations in GPA and smoking once class membership was considered. The results supported a complete mediation hypothesis and replicated previous findings [13, 17-21] showing that sex (male), grade repetition, low parental education, and low parental support and democratic control increased the risk of membership in lower GPA classes. Conversely, extracurricular involvement and parental academic practices had no effects on class membership once these other characteristics were taken into account, supporting the importance of sustaining generic parenting practices during early adolescence rather than domain-specific academic parenting practices. This does not mean that parental academic practices are unimportant but that their impact on GPA is likely to be better accounted for by generic parenting practices. Given the stability of generic parenting practices, early preventive intervention may yield long-term benefits. Once predictors of class membership were taken into account, very few direct residual effects on GPA and smoking remained, suggesting that the effects of the predictors on GPA and smoking were completely mediated by their effects on class membership.

The results also show that being a female is associated with greater increases in GPA over time, but also with increased risks of smoking, and that high parental education is associated with higher initial levels of GPA. However, since these effects did not differ across classes, the results did not support the mediated-moderation proposal. These additional results might be explained by recent observations suggesting that the known gender-related gap in achievement deepens across adolescence [49]. Similarly, a survey conducted in the Canadian province of Quebec reveals a higher rate of smoking in adolescent females at the time the present data were collected [50]. These results suggest that there are likely additional factors contributing to the increased rates of smoking observed in females. For instance, anxiety (and self-medication) tends to be more prevalent in females [51-52]. Similarly, female adolescents have a higher drive for thinness than males and suffer more from the weight gains associated with puberty [53-56]. Thus, females may be more prone to use cigarettes (or drugs) as a weight management strategy [57-61].

Despite the strengths of this study, limitations remain. First, this study relied on a short-term follow-up following secondary school transition making the findings hard to generalize to other developmental periods. Further, the attrition rate, albeit consistent with similar studies, remains high, and its impact on the generalizability of results is unknown. These findings thus need to be replicated using representative samples. Some of the measures used in this study were also suboptimal. For instance, most measures assessed generic constructs, which may have hidden domain-specific relations. Thus, although we relied on an official GPA measure, it did not allow us to differentiate specific subject matters. Likewise, we considered a binary indicator of self-reported smoking without relying on more specific or objective indicators of smoking frequency, addiction, or cessation attempts. Similarly, our measures of generic parenting addressed global rather than specific skills (e.g. monitoring, communication, support, praise). Additionally, our extracurricular involvement measure was based on only two indicators. Thus, our conclusion regarding the absence of effects of this variable remains tentative and should be thoroughly assessed in future studies considering specific activities (e.g. sports versus social) or dimensions (e.g. time spent, commitment, relations with the adults in charge) of involvement. A final potential limitation is the use of baseline measures of predictors instead of allowing these measures to vary across time with GPA. Indeed, it is possible that the relations between predictors, GPA and smoking change over time, altering the results. However, as our interest was in establishing temporal precedence in order to better guide preventive efforts, it was critical to measure these variables before trajectory class membership was established.

In conclusion, the results suggest that when adolescents do well in school, they are less likely to smoke. Practically, this important result suggests that it may be cost-effective for smoking prevention to focus on the few easy-to-identity unstable low achievers who represent more than a third of smoking onsets. Further, as low parental education is a risk factor for poor educational outcomes and adolescent smoking, when adolescents do poorly in school, they may predispose their own children to poor developmental outcomes, including smoking. Thus, the relationship between low GPA and smoking is serious and may even affect future generations. However, as parental support and democratic control reduce the likelihood of poor academic performance, teaching parents, especially those with low levels of education, essential parenting skills may be one way to reduce smoking and its intergenerational transmission.

Acknowledgements

This study was made possible by a grant from the Social Sciences and Humanities Research Council of Canada awarded to the last author. The authors thank Herbert W. Marsh for advice on the first draft of the results, as well as every student, staff and research assistant involved in the MADDP.

References

1. Biener L., Albers A.B. Young adults: Vulnerable new targets of tobacco marketing. *Am J Public Health* 2004; **94**: 326-330.
2. Hu M.C., Davies M., Kandel D.B. Epidemiology and correlates of daily smoking and nicotine dependence among young adults in the United States. *Am J Public Health* 2006; **96**: 299-308.
3. World Health Organization. *WHO report on the global tobacco epidemic, 2009: Implementing smoke-free environments*. Geneva: World Health Organization; 2009.
4. Audrain-McGovern J., Rodriguez D., Tercyak K.P., Cueva J., Rodgers K., Patterson F. Identifying and characterizing adolescent smoking trajectories. *Cancer Epidemiol Biomarkers Prev* 2004; **13**: 2023-2034.
5. Carvajal S.C., Granillo T.M. A prospective test of distal and proximal determinants of smoking initiation in early adolescents. *Addict Behav* 2006; **31**: 649-660.
6. Cox R.G., Zhang L., Johnson W.D., Bender D.R. Academic performance and substance use: Findings from a state survey of high school students. *J School Health* 2007; **77**: 109-115.
7. Ellickson P.L., Tucker J.S., Klein D.J. High-risk behaviors associated with early smoking: Results from a 5-year follow-up. *J Adolescent Health* 2001; **28**: 465-473.
8. Hishinuma E.S., Else R.N., Chang J.Y., Goebert D.A., Nishimura S.T., Choi-Misailidis S. et al. Substance use as a robust correlate of school outcome measures for ethnically diverse adolescents of Asian/Pacific ancestry. *School Psychol Quart* 2006; **21**: 256-322.
9. Khuder S.A., Price J.H., Jordan T., Khuder S.S., Silvestri K. Cigarette smoking among adolescents in northwest Ohio: correlates of prevalence and age at onset. *Int J Environ Res Public Health* 2008; **5**: 278-289.
10. Pederson L.L., Koval J.J., McGrady G.A., Tyas S.L. The degree and type of relationship between psychosocial variables and smoking status for students in grade 8: is there a dose-response relationship? *Prev Med* 1998; **27**: 337-347.
11. Schnohr C.W., Kreiner S., Rasmussen M., Due P., Diderichsen F. School-related mediators in social inequalities in smoking: a comparative cross-sectional study of 20 399 adolescents. *Int J Equity Health* 2009; **8**: 17.
12. Tucker J.S., Martinez J.F., Ellickson P.L., Edelen M.O. Temporal associations of cigarette smoking with social influences, academic performance, and delinquency: a four-wave longitudinal study from ages 13-23. *Psychol Addict Behav* 2008; **22**: 1-11.
13. Fergusson D.M., Horwood L.J., Boden J.M., Jenkin G. Childhood social disadvantage and smoking in adulthood: results of a 25-year longitudinal study. *Addiction* 2007; **102**: 475-482.
14. Heaven P.C., Ciarrochi J. Parental styles, conscientiousness, and academic performance in high school: a 3-wave longitudinal study. *Pers Soc Psychol B* 2008; **34**: 451-461.
15. Barnes G.M., Hoffman J.H., Welte J.W., Farrell M.P., Dintcheff B. Adolescents' time use: Effects on substance use, delinquency, and sexual activity. *J Youth Adolescence* 2007; **36**: 697-710.
16. Castrucci B.C., Gerlach K.K. Understanding the association between authoritative parenting and adolescent smoking. *Matern Child Hlth J* 2006; **10**: 217-224.
17. Chassin L., Presson C., Seo D.C., Sherman S.J., Macy J., Wirth R.J. et al. Multiple Trajectories of cigarette smoking and the intergenerational transmission of smoking: A multigenerational, longitudinal study of a community sample. *Health Psychol* 2008; **27**: 819-828.
18. Edelen M.O., Tucker J.S., Ellickson P.L. A discrete time hazards model of smoking initiation among West Coast youth from age 5 to 23. *Prev Med* 2007; **44**: 52-54.
19. Foster S.E., Jones D.J., Olson A.L., Forehand R., Gaffney C.A., Zens M.S. et al. Family socialization of adolescent's self-reported cigarette use: The role of parents' history of regular smoking and parenting style. *J Pediatr Psychol* 2007; **32**: 481-493.
20. Avenevoli S., Merikangas K.R. Family influences on adolescent smoking. *Addiction* 2003; **98**: S1-S20.
21. Darling N., Cumsille P. Theory, measurement, and methods in the study of family influences on adolescent smoking. *Addiction* 2003; **98**: 21-36.
22. Darling N., Steinberg L. Parenting style as context: An integrative model. *Psychol Bull* 1993; **113**: 487-496.
23. O'Connor T.G. The 'effects' of parenting reconsidered: Findings, challenges, and applications. *J Child Psychol Psychiatry* 2002; **43**: 555-572.

24. Berg-Nielsen T.S., Vikan A., Dahl A.A. Parenting related to child and parental psychopathology: A descriptive review of the literature. *Clin Child Psychol Psychiatry* 2002; **7**: 529-552
25. McKee L., Colletti C., Rakow A., Jones D.J., Forehand R. Parenting and child externalizing behaviors: Are the associations specific or diffuse? *Aggress Violent Behav* 2008; **13**: 201-215.
26. McLeod B.D., Wood J.J., Weisz J.R. Examining the associations between parenting and childhood anxiety: A Meta Analysis. *Clin Psychol Rev* 2007; **27**: 155-172.
27. Newman K., Harrison L., Dashiff C., Davies S. Relationships between parenting styles and risk behaviors in adolescent health: An integrative review. *Rev Lat Am Enfermagem* 2008; **16**: 142-150.
28. Ginsburg G.S., Bronstein P. Family factors related to children’s intrinsic/extrinsic motivational orientation and academic performance. *Child Dev* 1993; **64**: 1461-1474.
29. Grolnick W.S., Slowiaczek M.L. Parents’ involvement in children’s schooling: A multidimensional conceptualization and motivational model. *Child Dev* 1994; **65**: 237-252.
30. Grolnick W.S., Friendly R.W., Bellas, V.M. Parenting and Children’s Motivation at School. In: Wenzel K. & Wigfield A., editors. Handbook of motivation at school. New York: Routledge/Taylor & Francis Group; 2009. p. 171-195.
31. Hill N.E., Taylor L.C. Parental school involvement and children’s academic achievement: pragmatic and issues. *Curr Dir Psychol Sci* 2004; **13**: 161-164
32. Pomerantz E.M., Eaton M.M. (2001). Maternal intrusive support in the academic context: Transactional socialization processes. *Dev Psychol* 2001; **37**: 174-186.
33. Eccles J.S., Roeser R.W. Schools as developmental contexts during early adolescence. *J Res Adolesc* 2011; **21**: 225-241.
34. Fredricks J.A., Eccles J.S. Developmental benefits of extracurricular involvement: Do peer characteristics mediate the link between activities and youth outcomes? *J Youth Adolesc* 2005*;* **34**: 507-520.
35. Fredricks, J.A., Eccles J.S. Extracurricular involvement and adolescent adjustment: Impact of duration, number of activities, and breadth of participation. *Appl Dev Sci* 2006; **10**: 132-146.
36. Fredricks J.A., Eccles J.S. Participation in extracurricular activities in the middle school years: Are there developmental benefits for African American and European American youth? *J Youth Adolesc* 2006; **37**: 1029-1043.
37. Simpkins S.D., Eccles J.S., Becnel J.N. The mediational role of adolescents' friends in relations between activity breadth and adjustment. *Dev Psychol* 2008; **44,** 1081-1094.
38. Muller D., Judd C. M., Yzerbyt V.Y. When moderation is mediated and mediation is moderated. *J Pers Soc Psychol* 2005; **89**: 852-863.
39. Ding C.S. Variations in academic performance trajectories during high school transition: Exploring change profiles via multidimensional scaling growth profile analysis. *Educ Res Eval* 2008; **14**: 305-319.
40. Espy K.A., Fang H., Charak D., Minich N., Taylor H.G. Growth mixture modeling of academic achievement in children of varying birth weight risk. *Neuropsychology* 2009; **23**: 460-474.
41. Kreisman M.B. Evaluating academic outcomes of Head Start: An application of general growth mixture modeling. *Early Child Res Q* 2003; **18**: 238-254.
42. Jeynes W.H. The relationship between the consumption of various drugs by adolescents and their academic achievement. *Am J Drug Alcohol Ab* 2002; **28**: 15-35.43. Bryant A.L., Schulenberg J., Bachman J.G., O'Malley P.M., Johnston L.D. Understanding the links among school misbehavior, academic achievement, and cigarette use: A national panel study of adolescents. *Prev Sci* 2000; **1**: 71-87.
44. Morin A.J.S., Janosz M., Larivée S. The Montreal Adolescent Depression Development Project (MADDP): School life and depression following high school transition. *Psychiat Res J* 2010; **1**: 1-50.
45. Morin A.J.S., Maïano C., Nagengast B., Marsh H.W., Morizot J., Janosz M. General Growth Mixture Analysis of Adolescents’ Developmental Trajectories of Anxiety: The Impact of Untested Invariance Assumptions on Substantive Interpretations. *Struct Equ Modeling* in press.
46. Janosz M., Archambault I., Lacroix M., Lévesque J. *Screening package for school dropout risk - User's manual. School Environment Research Group*. Montreal: University of Montreal; 2007.
47. Leblanc M. *Manual for the questionnaire package for the assessment of the social and personal*

adaption of Quebec adolescents. Montreal: University of Montreal; 1998.
48. Muthén L.K., Muthén B.O. *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén; 2008.
49. Chouinard R., Roy N. Changes in high-school students' competence beliefs, utility value and achievement goals in mathematics. *Brit J Educ Psychol* 2008; **78**: 31-50.
50. Dubé G., Bordeleau M., Cazale L., Fournier C., Traoré I, Plante N. et al. *Enquête québécoise sur le tabac, l'alcool, la drogue et le jeu chez les élèves du secondaire, 2008* [*Quebec survey on tobacco, alcohol, drugs, and gambling in secondary school youths, 2008*]. Sainte Foy: Institut de la statistique du Québec; 2008.
51. Fallu J.-S., Brière F.N., Descheneaux A., Keegan V., Maguire J., Chabot A., et al. *Consommation d'amphétamines chez les adolescents et les adolescentes: étude des facteurs associés avec centration sur les différences entre les sexes* [*Amphetamine use among adolescents: Analysis of associated factors an gender differences*]. GRIP-Montréal report to the Quebec Ministry of Health and Social Services; 2008.
52. Semple S.J., Patterson T.L., Grant, I. The context of sexual risk behavior among heterosexual methamphetamine users. *Addict Behav* 2004; **29**: 807-810.
53. Alsaker F.D. Timing of puberty and reactions to pubertal changes. In: Rutter M., editor, Psychosocial disturbances in young people: Challenges for prevention. New York: Cambridge University; 1995. p. 37-.82.
54. Carson N. J., Rodriguez D., Audrain-McGovern J. Investigation of mechanisms linking media exposure to smoking in high school students. *Prev Med* 2005; **41**: 511-520.
55. Morin A.J.S., Maïano C., Marsh H.W., Janosz M., Nagengast B. The longitudinal interplay of adolescents' self-esteem and body image: A conditional autoregressive latent trajectory Analysis. *Multivar Behav Res* 2011; **46**: 157-201.
56. Steinberg L., Morris A. (2001). Adolescent development. *Annu Rev Psychol* 2001; **52**: 83-110.
57. Boys A., Marsden J., Strang J. Understanding reasons for drug use amongst young people: A functional perspective. *Health Educ Res* 2001, **16**; 457-469.
58. Brecht M.-L., O'Brien A., Von Mayrhauser C., Anglin M.D. Methamphetamine use behaviors and gender differences. *Addict Behav* 2004, **29**; 89-106.
59. Camp D.E., Klesges R.C., Relyea G. The relationship between body weight concerns and adolescent smoking. *Health Psychol* 1993; **12**: 24-32.
60. Degenhardt L., Coffey C., Moran P., Carlin J.B., Patton G.C. The predictors and consequences of adolescent amphetamine use: Findings from the Victoria Adolescent Health Cohort Study. *Addiction* 2007, **102**; 1076-1084.
61. Seo D.C., Jiang N. Associations between smoking and extreme dieting among adolescents. *J Youth Adolesc* 2009; **38**: 1364-1373.

Table 1 Results from the multinomial logistic regression analyses predicting class membership.

	Moderate Versus High		Low-unstable versus High		Low-unstable versus Moderate	
	Coefficient (s.e.)	OR	Coefficient (s.e.)	OR	Coefficient (s.e.)	OR
Gender	**-0.69 (0.25)****	0.50	-0.60 (0.35)	0.55	0.09 (0.30)	1.09
Nationality	-0.18 (0.39)	0.83	0.23 (0.48)	1.26	0.41 (0.40)	1.51
Repeating a grade	**1.23 (0.45)****	3.42	**2.22 (0.47)****	9.21	**0.99 (0.30)****	2.69
Parental education	**-0.34 (0.07)****	0.72	**-0.63 (0.10)****	0.54	**-0.29 (0.09)****	0.75
Parental support	**-0.32 (0.12)****	0.73	**-0.44 (0.17)****	0.65	-0.12 (0.13)	0.89
Parental control	**-0.20 (0.10)***	0.82	**-0.31 (0.14)***	0.74	-0.11 (0.11)	0.90
Parental academic support	-0.02 (0.15)	0.98	-0.31 (0.29)	0.74	-0.29 (0.26)	0.75
Parental academic pressure	0.34 (0.28)	1.41	-0.43 (0.44)	0.65	-0.77 (0.41)	0.46
Extracurricular involvement	0.11 (0.11)	1.11	0.22 (0.17)	1.24	0.11 (0.15)	1.12

Note: * $p \leq .05$; ** $p \leq .01$; s.e. = standard error of the coefficient (the coefficient divided by its standard error is equivalent to a t score and indicate the significance of the effect); OR = odds ratio.

Table 2 Results from the multiple regressions predicting the trajectory factors and from the logistic regression predicting the outcome.

	Predicting the Intercept	Predicting the linear slope	Predicting the outcome	
	Coefficient (s.e.)	Coefficient (s.e.)	Coefficient (s.e.)	OR
Gender	0.73 (0.52)	**0.29 (0.070)****	**0.90 (0.23)****	**2.469**
Nationality	-0.09 (0.79)	0.02 (0.11)	-0.37 (0.34)	0.694
Repeating a grade	-1.67 (0.91)	0.21 (0.15)	0.41 (0.22)	1.508
Parental education	**0.58 (0.14)****	0.02 (0.02)	0.12 (0.06)	1.126
Parental support	0.02 (0.26)	0.03 (0.03)	-0.17 (0.09)	0.845
Parental control	0.12 (0.19)	-0.05 (0.03)	0.09 (0.08)	1.089
Parental academic support	-0.66 (0.36)	0.05 (0.05)	0.21 (0.15)	1.233
Parental academic pressure	1.05 (0.63)	-0.05 (0.09)	-0.02 (0.25)	0.976
Extracurricular involvement	0.28 (0.23)	-0.03 (0.03)	0.02 (0.12)	1.016

Note: * $p \leq .05$; ** $p \leq .01$; s.e. = standard error of the coefficient (the coefficient divided by its standard error is equivalent to a t score and indicate the significance of the effect); OR = odds ratio.

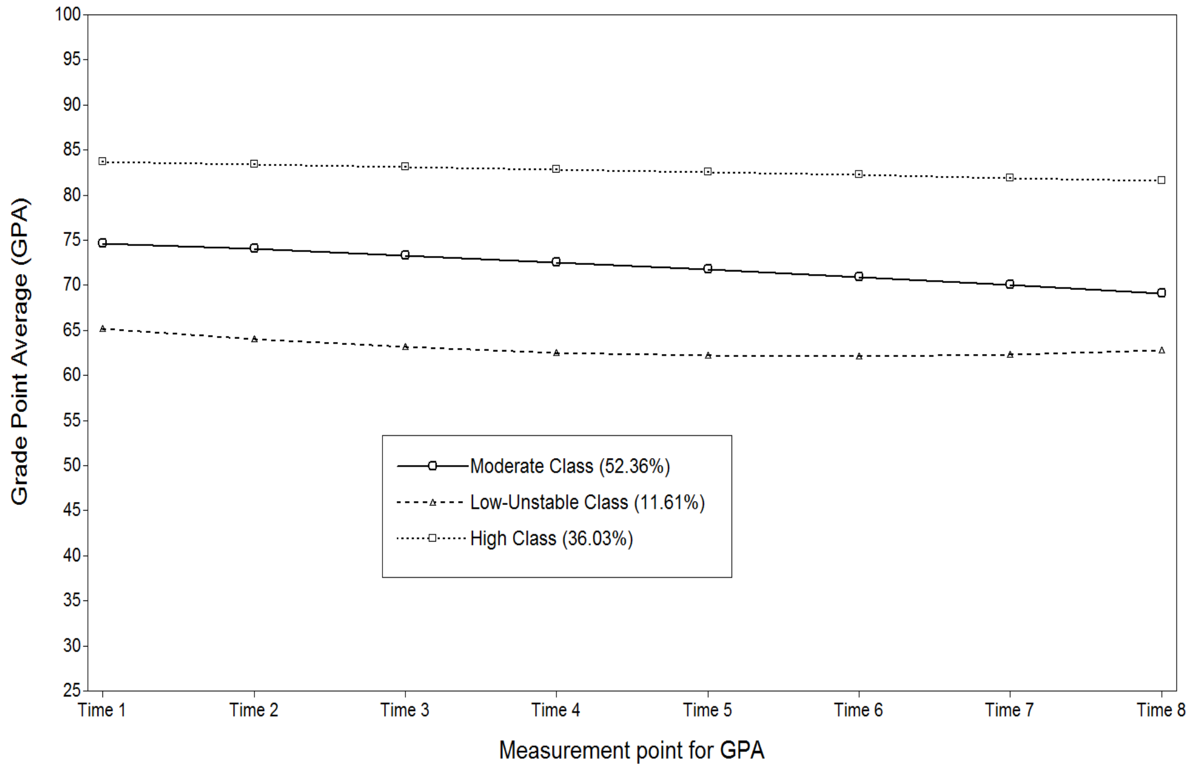


Figure 1 a Estimated latent trajectory classes.

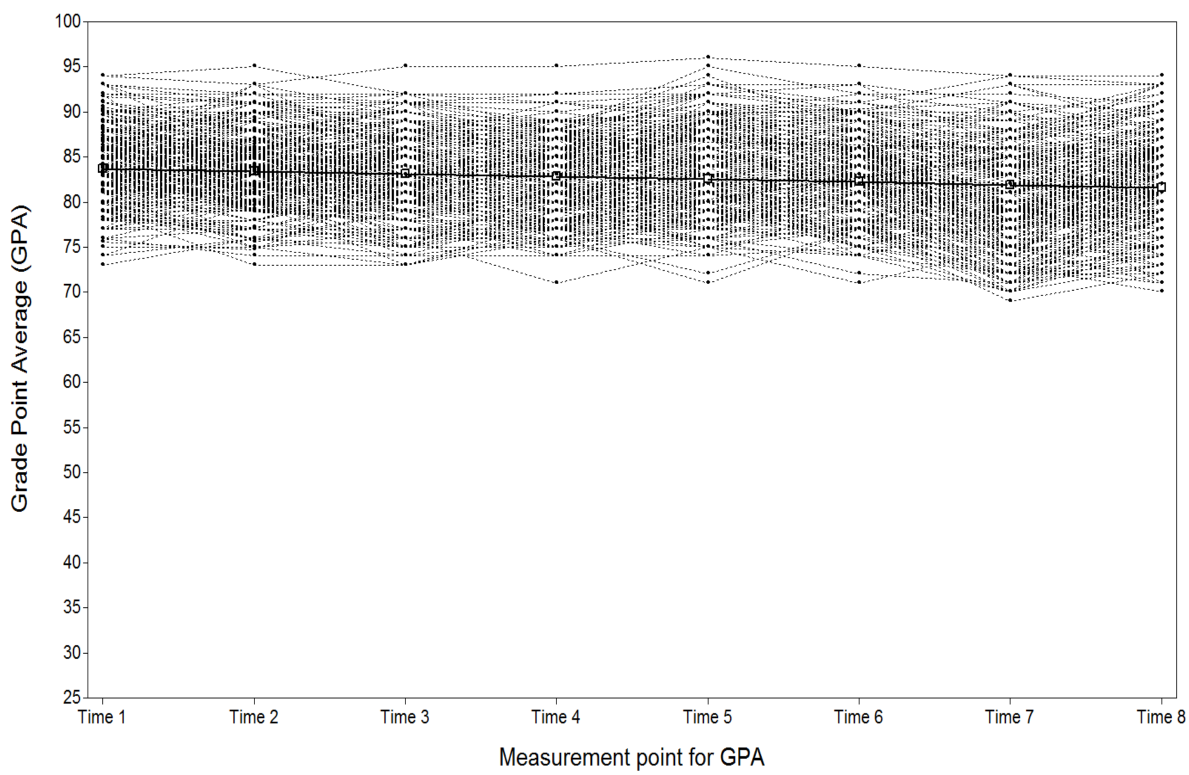


Figure 1 b Estimated trajectory and observed individual values in the elevated class.

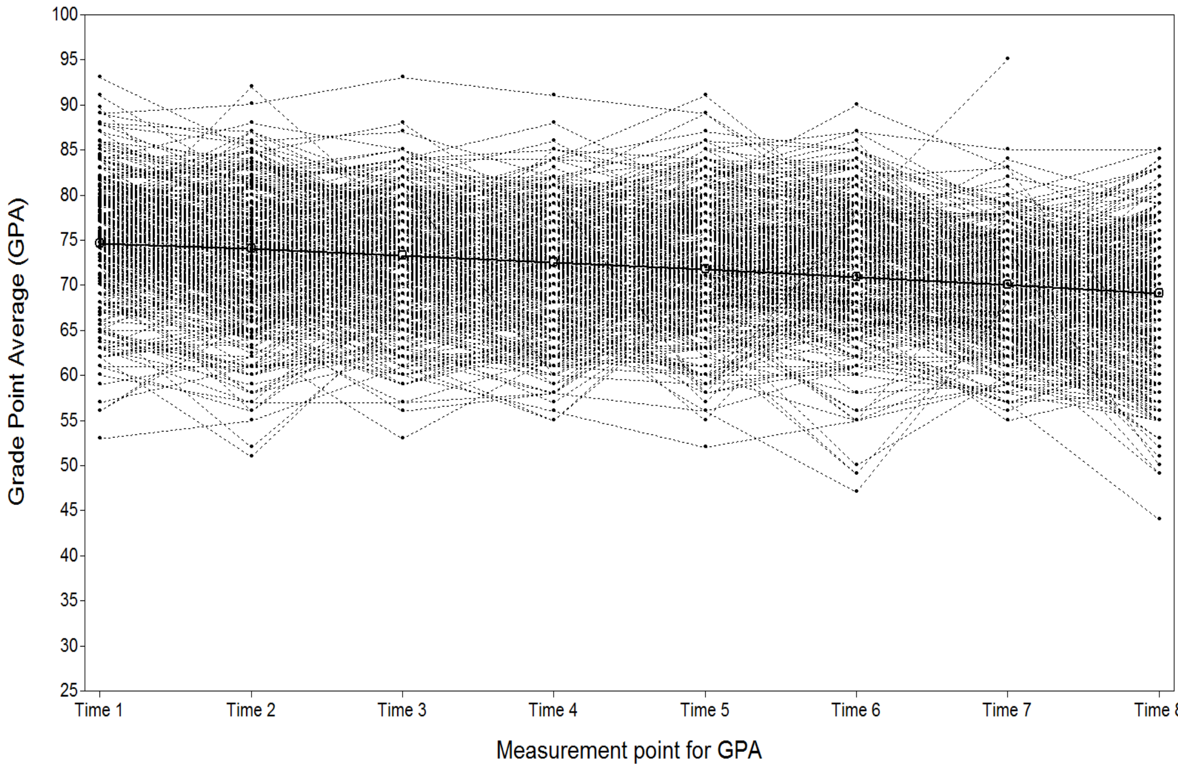


Figure 1 c Estimated trajectory and observed individual values in the moderate class.

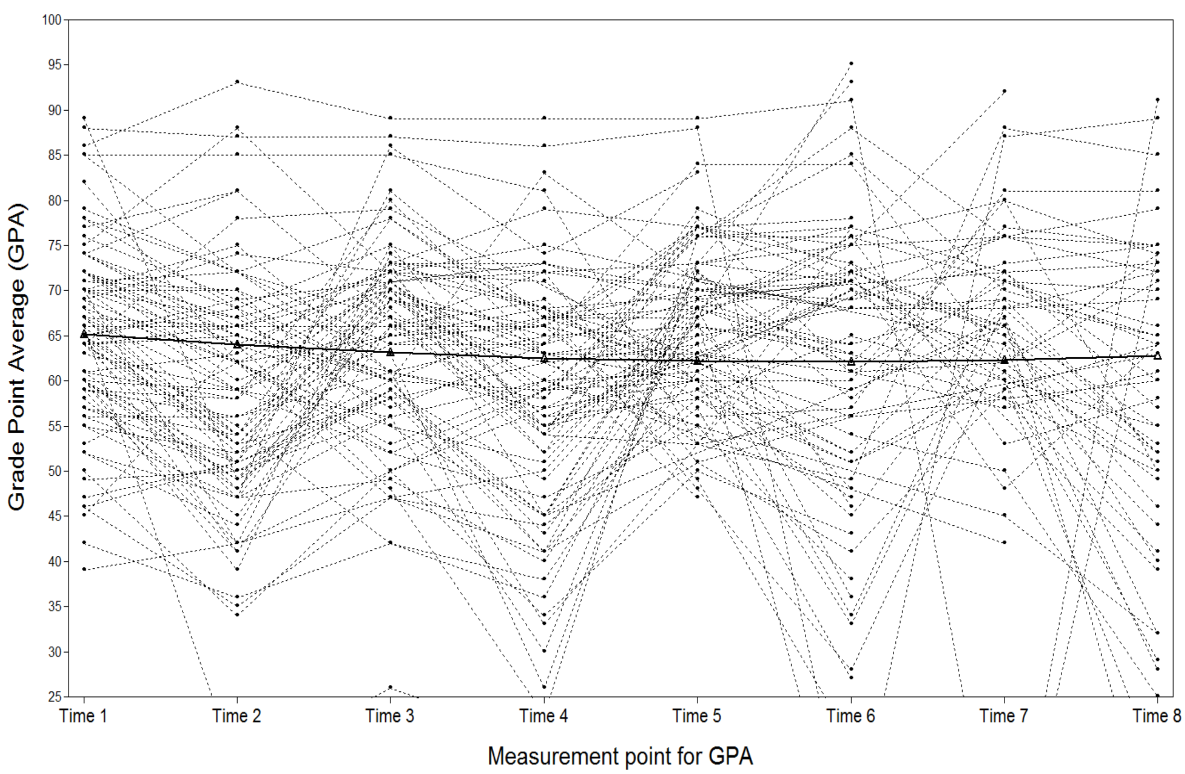


Figure 1 d Estimated trajectory and observed individual values in the low-unstable class.

Supplemental Technical Materials for:
Academic Achievement and Smoking Initiation in Adolescence: A General Growth Mixture Analysis.

The variables correlations and descriptive statistics are presented in the supplemental Table 1.

In order to determine if meaningful latent classes of students could be identified based on their grade point average (GPA) trajectories, general growth mixture modeling (GGMM) [1-3] were used. GGMM combine latent curve models, used to analyse developmental trends [4], and latent class analyses [5-6] used to extract qualitatively distinct latent classes. Predictors and outcomes of class membership and trajectory factors (intercept, linear slope and quadratic slope) may be directly integrated in the model and their effects may vary across classes.

The analyses were performed using Mplus 5.1 [7], which relies on the expectation-maximization algorithm of the robust maximum likelihood estimator (MLR) to estimate GGMM model parameters [3]. All models were estimated with 1000 random sets of start values and converged on a well replicated solution [8]. Models with 1 to 7 latent trajectory classes of GPA were estimated with intercepts, linear slopes, and quadratic slopes specified as fully variant across latent classes (i.e. latent means and variance-covariances). The time-specific residuals were also freely estimated in all classes. These specifications are in line with the results from recent studies showing that assuming linearity or invariance may result in biased estimates when these constraints do not hold in the population [9-14]. By default, Mplus relies on Full-Information Maximum Likelihood estimation to handle missing data on the repeated (GPA measures), which is generally considered to be as effective as multiple imputations except in scenarios including far more missing values than in the present study [15-17]. Indeed, given that GPA ratings were obtained from school records, there were less than 10% of missing data on all of these repeated measures

More specifically, GGMM were estimated as a quadratic growth model for the outcome y_{it} where *i* is the index for individual and *t* is the index for time [9]. This model also includes *c*, a categorical latent variable with *k* levels (k = 1, 2, …, K) that is estimated from the data, with each individual *i* having a probability of membership in each of the *k* levels representing the latent trajectory classes.

$$y_{it} = \sum_{k=1}^{K} p(c=k)[\alpha_{iyk} + \beta_{1iyk}\lambda_t + \beta_{2iyk}\lambda^2{}_t + \varepsilon_{yitk}] \quad (1)$$

$$\alpha_{iyk} = \mu_{\alpha yk} + \zeta_{\alpha yik} \quad (2)$$

$$\beta_{1iyk} = \mu_{\beta 1yk} + \zeta_{\beta 1yik} \quad (3)$$

$$\beta_{2iyk} = \mu_{\beta 2yk} + \zeta_{\beta 2yik} \quad (4)$$

The *k* subscript indicates that most parameters were freely estimated in all of the estimated latent trajectory classes and that each latent trajectory class could thus be defined by its own latent curve model based on independent covariance matrices and mean vectors. More precisely, α_{iyk}, β_{1iyk} and β_{2iyk} respectively represent the random intercept, random linear slope and random quadratic slope of the trajectory for individual *i* in latent trajectory class *k*; ε_{yitk} represents the time- individual- and class-specific errors; $\mu_{\alpha yk}$, $\mu_{\beta 1yk}$ and $\mu_{\beta 2yk}$ represent the average intercept, linear slope and quadratic slope in latent trajectory class *k*; and $\zeta_{\alpha yik}$, $\zeta_{\beta 1yik}$ and $\zeta_{\beta 2yik}$ are disturbances reflecting the variability of the estimated intercepts and slopes across cases within latent trajectory classes. These disturbances have a mean of zero and a variance-covariance matrix represented by Φ_{yk}:

$$\Phi_{yk} = \begin{bmatrix} \psi_{\alpha\alpha yk} & & \\ \psi_{\alpha\beta 1yk} & \psi_{\beta 1\beta 1yk} & \\ \psi_{\alpha\beta 2yk} & \psi_{\beta 1\beta 2yk} & \psi_{\beta 2\beta 2yk} \end{bmatrix} \quad (5)$$

Errors (ε_{yitk}) are assumed to have a mean of 0 and to be uncorrelated over time, across cases or with the other model parameters, but were allowed to vary across periods. Time is indicated by λ_t, which represents the loadings of the time-specific measurement points on the slope factor and is coded to reflect the equal intervals between the measurement points. More precisely, in the current study, time was coded so that the intercepts of the trajectories were estimated at Time 1 [$E(\alpha_{iyk}) = \mu_{y1k}$]: $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 2$, $\lambda_4 = 3$, $\lambda_5 = 4$, $\lambda_6 = 5$, $\lambda_7 = 6$, and , $\lambda_8 = 7$. As the latent classes are unknown but estimated from the data, GGMA estimates a probability of membership in each latent trajectory class for all individuals, which is reflected in the first part of the equation $\sum_{k=1}^{K} p(c = k)$. These probabilities add up to one for each individual across all classes and unconditionally over all classes.

The distal outcome (i.e. smoking) was also directly incorporated into the models, to avoid relying on a sub-optimal two-step process [18]. Doing so involves including the outcome as an additional mixture indicator and may influence the nature of the latent classes, as well as inflating their relation with the outcome. To ensure that no such biases were introduced, the models were estimated with and without the outcome and were found to converge on identical results [18].

An important challenge in GGMM is determining the number of latent classes in the data. To this end we inspected (i) the substantive meaning of the extracted classes [19, 20]; (ii) statistical adequacy of the solution (e.g. absence of negative variance) [11]; (iii) information criteria based on the model loglikelihood (a lower value suggests a better fitting model) such as the Akaïke information criterion (AIC), the consistent AIC (CAIC), the Bayesian information criterion (BIC), and the sample-size adjusted BIC (ABIC); (iv) the Lo, Mendel and Rubin's [21] likelihood ratio test (LMR) and the bootstrap likelihood ratio test (BLRT) based on 100 bootstrap samples (significance suggest that the *k*-class model fits significantly better than the *k-1*-class model). Simulation studies indicate that the ABIC, CAIC, BIC and BLRT are effective in choosing the model which best recovers the sample's true parameters in GGMM [22-24]. Previous authors [9, 18] suggested graphically representing information criteria through elbow plots illustrating the gains associated with each added class: the first angle indicates the optimal number of latent classes in the data. Additional indices include the entropy, for which higher values (closer to 1) indicates the precision with which the cases are classified into the various extracted latent profiles. The entropy should not in itself be used to determine the model with the optimal number of classes, but is important because it summarizes the extent to which a model generates classification errors [25].

The fit indices from the models are reported in the supplemental Table 2. Most of the efficient indicators (CAIC, BIC, ABIC, BLRT) converged on the GGMM with four trajectory classes. The only divergent results came from the AIC and LMR, indicators with a known tendency for overestimation. However, when elbow plots (see the supplemental Figure) were drawn for the recommended indices (CAIC, BIC and ABIC), these plots converged on a three-class solution. This is consistent with the fact that the four-class GGMM converged on an improper solution (with negative variances estimates and non positive definite first order derivative product matrix), apparently related to the extraction of a very small class including only 2.29% of participants. This suggests the inadequacy of the 4-class solution [11, 22] and the superiority of the three-class solution.

To further illustrate the added value of this model, we conducted a multiple logistic regression analysis in which all eight GPA indicators and the posterior probabilities of GGMM class membership were used to predict smoking to verify whether the addition of the class probabilities significantly improved the prediction. A similar analysis was also conducted in which student-specific trajectory factors scores estimated from classical latent curve models were used instead of the eight GPA indicators. In both cases, the percentage of explained variance [26], following the inclusion of the class probabilities was close to 130-140% what it was before (i.e. it increased from 12% to 16% for the analyses with GPA indicators and from 13% to 18% for the analyses with latent curve factors).

Predictors were then added to the model selected as presenting the optimal number of classes [9, 18, 19, 27]. As Mplus does not allow for missing data on the exogenous predictors, they were imputed with ML estimates using the EM algorithm [17] of SPSS 15.0 "missing values" module. Imputed estimates were conditional on all predictors used in the study. Given the low levels of missing data (0% to 4.05%, mean = 0.85%, s.d. = 1.27%), multiple imputation was not warranted.

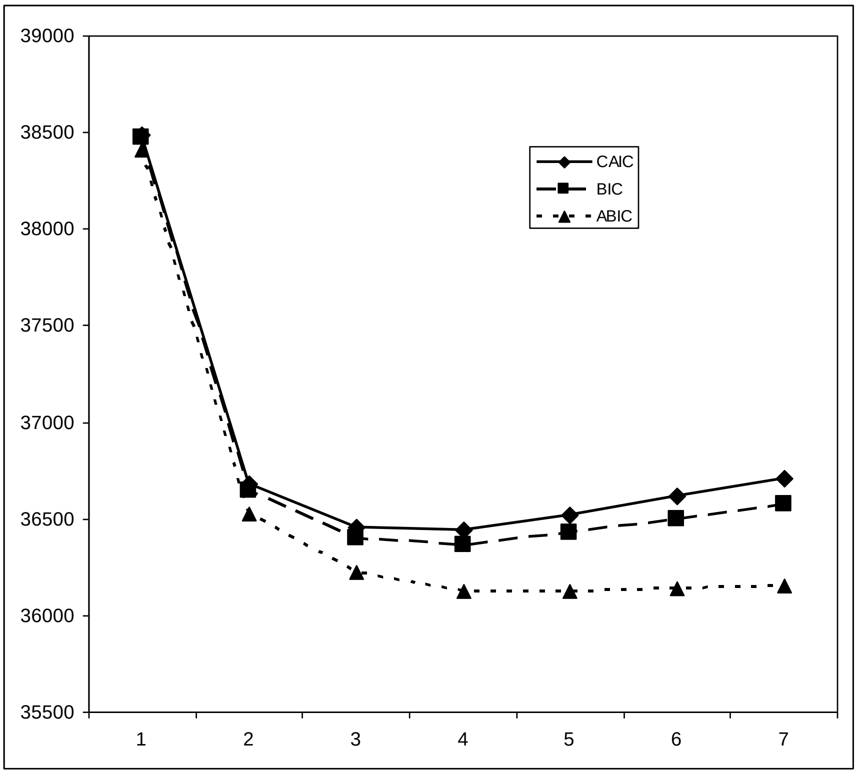
A baseline model was first estimated in which the predictors predicted only class membership through a multinomial logistic regression. Following this, models in which predictors were allowed to influence the trajectory factors and the outcomes, and in which these effects were also allowed to vary across classes were estimated. These models were compared on the basis of information criteria and robust likelihood ratio tests (LRT) [7], that may be used in the comparison of nested models [9, 12, 18]. LRTs are calculated as minus two times the difference in the log likelihood of the two nested models and can be interpreted as a chi square with degrees of freedom equal to the difference between the two models in terms of free parameters. Since Mplus relies on the MLR estimator, scaling factors may also be taken into account in this comparison [7] by dividing the LRT by its scaling correction composite, cd, were: (i) cd = (p0 *co – p1 *c1) / (p0-p1); (ii) p0 and p1 are, respectively the number of free parameters in nested and comparison models; (iii) c0 and c1 are, respectively, the scaling correction factors for nested and comparison models.

Fit indices from the conditional models including the predictors are also presented in the supplemental Table. As a first step, predictors only predicted class membership (Model 1). Then, their effects on the latent trajectory factors were also estimated. Although the information criteria generally appeared to favor the baseline model, the LRT tests showed that allowing the predictors to influence the intercepts (Model 2) and linear slope (Model 3), but not the quadratic slope (Model 4), significantly improved model fit. Similarly, the LRT tests indicated that these effects did not need to vary across latent classes (Model 5). Finally, the results showed that allowing predictors to have a direct effect on the outcome (Model 6), without allowing this effect to vary across classes (Model 7), also improved model fit. According to AIC and ABIC, Model 6 was the best fitting of all. The specific results from this model are reported in the main manuscript. It could be argued that the inclusion of additional third variables known to predict both GPA and smoking, such as peer deviancy, could have resulted in a fuller picture of the mechanisms involved in the observed relations. We note however, that adding covariates would not have changed the bulk of the results regarding trajectory class identification as this process was done without any covariates in the model.

Supplemental References

1. Muthén B.O. Beyond SEM: General latent variable modeling. *Behaviormetrika* 2002; **29**: 81-117.
2. Muthén B.O. Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In: Kaplan D., editor. *Handbook of quantitative methodology for the social sciences*. Newbury Park, CA: Sage; 2004, p. 345-368.
3. Muthén B.O., Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics* 1999; **55**: 463-469.
4. Bollen K.A., Curran, P.J. *Latent curve models*. Hoboken, NJ: Wiley; 2006.
5. Bartholomew D.J. *Latent variable models and factor analysis*. London, UK: Griffin; 1987.
6. Lazarfeld P.F., Henry N.W. *Latent structure analysis*. Boston, MA: Houghton-Mifflin; 1968.
7. Muthén L.K., Muthén B.O. *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén; 2008.
8. Hipp J.R., Bauer D.J. Local solutions in the estimation of growth mixture models. *Psychol Methods* 2006; **11**: 36-53.
9. Morin A.J.S., Maïano C., Nagengast B., Marsh H.W., Morizot J., Janosz M. General Growth Mixture Analysis of Adolescents' Developmental Trajectories of Anxiety: The Impact of Untested Invariance Assumptions on Substantive Interpretations. *Struct Equ Modeling* in press.
10. Bauer D.J. Observations on the use of growth mixture models in psychological research. *Multivar Behav Res* 2007; **42**: 757-786.
11. Bauer D.J., Curran P.J. Distributional assumptions of growth mixture models over-extraction of latent trajectory classes. *Psychol Methods* 2003; **8**: 338-363.
12. Bauer D.J., Curran P.J. The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychol Methods* 2004; **9**: 3-29.
13. Enders C.K., Tofighi D. The impact of misspecifying class-specific residual variances in growth mixture models. *Struct Equ Modeling* 2008; **15**: 75-95.
14. Voelkle M.C. Reconsidering the use of autoregressive latent trajectory (ALT) models. *Multivar Behav Res* 2008; **43**: 564-591.
15. Enders, C.K. *Applied missing data analysis*. New York, NY: Guilford; 2010.
16. Graham J.W. Missing data analysis: Making it work in the real world. *Ann Rev Psych 2009;* **60**: 549-576.
17. Little R.J., Rubin D.B. *Statistical analysis with missing data*. Hoboken, NJ: Wiley; 2002.
18. Petras H., Masyn K. General growth mixture analysis with antecedents and consequences of change. In: Piquero A., Weisburd D., editors. *Handbook of Quantitative Criminology*. New

York, NY: Springer; 2010, p. 69-100
19. Marsh H.W., Lüdtke O., Trautwein U., Morin A.J.S. Latent profile analysis of academic self-concept dimensions: Synergy of person- and variable-centered approaches to internal/external frame of reference model. *Struct Equ Modeling* 2009; **16**: 1-35.
20. Muthén B.O. Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran (2003). *Psychol Methods* 2003; **8**: 369-377.
21. Lo Y., Mendell N.R., Rubin D.B. Testing the number of components in a normal mixture. *Biometrika* 2001; **88**: 767-778.
22. Nylund K.L., Asparouhov T., Muthén B.O. Deciding on the number of classes in latent class analysis and growth mixture modeling. A Monte Carlo simulation study. *Struct Equ Modeling* 2007; **14**: 535-569.
23. Tofighi D., Enders C.K. Identifying the correct number of classes in growth mixture models. In: Hancock G.R., Samuelsen K.M., editors. *Advances in latent variable mixture models*. Charlotte, NC: Information Age; 2008, p. 317-341.
24. Tolvanen A. *Latent growth mixture modeling: A simulation study*. Unpublished doctoral thesis. Jyvaskyla: University of Jyvaskyla; 2007.
25. McLachlan G., Peel, D. *Finite mixture models*. New York, NY: Wiley, 2000.
26. Nagelkerke N.J.D. A note on a general definition of the coefficient of determination. *Biometrika* 1991; **78**: 691-692.
27. Morin A.J.S., Morizot J., Boudrias J.-S., Madore I. A multifoci person-centered perspective on workplace affective commitment: A latent profile/factor mixture analysis. *Organ Res Methods* 2011; **14**: 58-90.



Supplemental Figure Elbow plots of the CAIC, BIC and ABIC for the unconditional GGMM.

Supplemental Table 1 Descriptive statistics and correlations of the study variables.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.
1. Gender	1																	
2. Nationality	-0.034	1																
3. Repeat. grade	-0.030	0.130*	1															
4. Par. Educ.	-0.005	0.032	-0.171*	1														
5. GPA1-Year1	0.120*	-0.015	-0.284*	0.376*	1													
6. GPA2-Year1	0.202*	-0.077*	-0.324*	0.377*	0.832*	1												
7. GPA1-Year2	0.176*	-0.047	-0.189*	0.380*	0.753*	0.777*	1											
8. GPA2-Year2	0.137*	-0.064	-0.211*	0.376*	0.737*	0.781*	0.850*	1										
9. GPA1-Year3	0.193*	0.011	-0.234*	0.379*	0.666*	0.706*	0.747*	0.734*	1									
10. GPA2-Year3	0.222*	0.004	-0.211*	0.353*	0.601*	0.647*	0.631*	0.635*	0.737*	1								
11. GPA1-Year4	0.181*	-0.033	-0.118*	0.285*	0.593*	0.570*	0.633*	0.598*	0.630*	0.599*	1							
12. GPA2-Year4	0.233*	-0.041	-0.156*	0.347*	0.555*	0.603*	0.653*	0.630*	0.674*	0.615*	0.655*	1						
13. Par. Supp.	0.125*	0.012	-0.132*	0.111*	0.218*	0.234*	0.260*	0.217*	0.236*	0.224*	0.179*	0.205*	1					
14. Par. control	0.069	-0.042	-0.081*	0.041	0.144*	0.166*	0.146*	0.137*	0.100*	0.064	0.097*	0.118*	0.153*	1				
15. Par. acad. Supp.	0.087*	-0.067	-0.119*	0.141*	0.184*	0.230*	0.236*	0.193*	0.214*	0.186*	0.124*	0.100*	0.405*	0.142*	1			
16. Par. acad. Press.	-0.121*	0.137*	0.015	0.069	-0.078*	-0.098*	-0.081*	-0.064	-0.044	-0.002	-0.046	-0.061	-0.162*	-0.335*	-0.183*	1		
17. Extra. Involv.	-0.024	0.029	0.139*	-0.045	-0.041	-0.056	-0.017	-0.018	-0.034	-0.017	-0.006	-0.076	0.051	-0.042	0.070	-0.024	1	
18. Smoking	0.093*	-0.006	0.192*	-0.064	-0.241*	-0.246*	-0.259*	-0.277*	-0.200*	-0.221*	-0.180*	-0.228*	-0.145*	-0.033	-0.092*	0.041	0.038	1
Mean	0.458	0.112	0.130	6.017	77.111	75.278	75.490	74.110	75.89	75.280	73.130	72.910	7.072	7.278	3.566	2.002	2.198	0.163
SD	0.499	0.316	0.415	1.976	8.595	10.655	9.017	10.465	9.123	10.557	9.262	11.232	1.290	1.396	0.439	0.769	0.984	0.370
Variance	0.249	0.100	0.172	3.903	73.876	113.518	81.314	109.521	83.223	111.449	85.787	126.149	1.663	1.949	0.192	0.591	0.969	0.137
Range	0-1	0-1	0-3	1-9	39-94	17-95	26-95	20-95	14-96	4-95	0-95	8-94	1.7-10	2.5-10	1.8-4	1-4	1-4	0-1
Items	1	1	1	1	1	1	1	1	1	1	1	1	15	11	7	4	2	1
Compilation	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na	Mean	Mean	Mean	Mean	Mean	Na

Note. * = $p \leq .05$; GPA= Grade Point Average; Na= Not applicable; SD = Standard Deviation.

Supplemental Table 2 Fit indices from the various unconditional and conditional Models.

Model	LL	#fp	SF	CM	LRT (df)	AIC	CAIC	BIC	ABIC	Entropy	LMR	BLRT
Unconditional models												
1 Class	-19175	18	4.67	NA	NA	38387	38488	38470	38412	NA	NA	NA
2 Class	-18199	37	1.66	NA	NA	36473	36680	36643	36526	.89	$\leq .01$	$\leq .01$
3 Class	-18016	56	1.35	NA	NA	36143	36457	36401	36224	.81	$\leq .01$	$\leq .01$
4 Class	-17935	75	0.36	NA	NA	36020	**36441**	**36366**	**36128**	.83	$\leq .01$	$\leq .01$
5 Class	-17903	94	0.99	NA	NA	35993	36520	36426	36128	.83	$\leq .01$	**$\geq .05$**
6 Class	-17877	113	0.01	NA	NA	35980	36614	36501	36142	.84	$\leq .01$	$\leq .01$
7 Class	-17852	132	0.89	NA	NA	**35968**	36708	36576	36157	.72	$\leq .01$	$\leq .01$
Conditional models (3-class)												
Model 1. P -> class membership	-17887	74	1.283	NA	NA	35923	36338	36264	36029	0.831	NA	NA
Model 2. P -> class membership and trajectory factors I (INV.)	-17858	83	1.252	1	-58 (9)*	35882	36347	36264	36001	0.816	NA	NA
Model 3. P -> class membership and trajectory factors I S (INV.)	-17844	92	1.233	2	-26 (9)*	35871	36387	36295	36003	0.817	NA	NA
Model 4. P -> class membership and trajectory factors I S Q (INV.)	-17834	101	1.219	3	-18 (9)	35870	36436	36335	36014	0.817	NA	NA
Model 5. P -> class membership and trajectory factors I S (VAR.)	-17817	128	1.197	3	-49 (36)	35890	36608	36480	36074	0.826	NA	NA
Model 6. P -> class membership, trajectory factors I S (INV.), and outcome (INV.)	-17830	101	1.213	3	-28 (9)*	35861	36428	36327	36006	0.816	NA	NA
Model 7. P -> class membership, trajectory factors I S (INV.), and outcome (VAR.)	-17821	119	1.198	6	-18 (16)	35881	36549	36430	36052	0.817	NA	NA

Note. $p \leq .01$ (p was a priori fixed at .01 due to the known sensitivity of LRT tests to sample size and to account for the multiple comparisons that were conducted); LL = model loglikelihood; #fp = number of free parameters; SF: scaling factor of the robust Maximum Likelihood estimator; CM = comparison model for the robust likelihood ratio tests; LRT = robust likelihood ratio tests; AIC = Akaïke Information Criterion; CAIC = consistent AIC; BIC = bayesian information criterion; ABIC = sample-size adjusted BIC; LMR = Lo-Mendel and Rubin's likelihood ratio test; BLRT = bootstrap likelihood ratio test; P -> = the predictors were allowed to influence…; I = intercept factor of the latent trajectories; S = linear slope factor of the latent trajectories; Q = quadratic slope factor of the latent trajectories; INV. = the prediction was constrained to invariance across the latent trajectory classes; VAR. = the prediction was allowed to vary across the latent trajectory classes.