



Published in final edited form as:

J Youth Adolesc. 2012 June ; 41(6): 764–775. doi:10.1007/s10964-011-9720-0.

The Stability of Perceived Pubertal Timing across Adolescence

Jessica Duncan Cance, M.P.H., Ph.D.¹, Susan T. Ennett, M.S.P.H., Ph.D.², Antonio A. Morgan-Lopez, Ph.D.³, and Vangie A. Foshee, Ph.D.²

¹Department of Kinesiology and Health Education, College of Education, The University of Texas at Austin

²Department of Health Behavior and Health Education, Gillings School of Global Public Health, University of North Carolina at Chapel Hill

³L.L. Thurstone Psychometric Laboratory, Department of Psychology, University of North Carolina at Chapel Hill

Abstract

It is unknown whether perceived pubertal timing changes as puberty progresses or whether it is an important component of adolescent identity formation that is fixed early in pubertal development. The purpose of this study is to examine the stability of perceived pubertal timing among a school-based sample of rural adolescents aged 11 to 17 (N=6,425; 50% female; 53% White). Two measures of pubertal timing were used, *stage-normative*, based on the Pubertal Development Scale, a self-report scale of secondary sexual characteristics, and *peer-normative*, a one-item measure of perceived pubertal timing. Two longitudinal methods were used: one-way random effects ANOVA models and latent class analysis. When calculating intraclass correlation coefficients using the one-way random effects ANOVA models, which is based on the average reliability from one time point to the next, both measures had similar, but poor, stability. In contrast, latent class analysis, which looks at the longitudinal response pattern of each individual and treats deviation from that pattern as measurement error, showed three stable and distinct response patterns for both measures: always early, always on-time, and always late. Study results suggest instability in perceived pubertal timing from one age to the next, but this instability is likely due to measurement error. Thus, it may be necessary to take into account the longitudinal pattern of perceived pubertal timing across adolescence rather than measuring perceived pubertal timing at one point in time.

Keywords

puberty; pubertal timing; measurement

Introduction

Puberty is not a one-time distinct event, but is a process of sequential events that has variable onset and progression (Hayward, 2003). This variability has prompted an interest in pubertal timing, defined as an adolescent's development relative to their peers of the same age and gender. Off-timing, compared with being in sync with peers, has been associated with anxiety and depression, and a number of deleterious health behaviors, including sexual risk taking, delinquency, and substance use (e.g., Archibald, Graber, & Brooks-Gunn, 2003;

Correspondence concerning this article should be addressed to Dr. Jessica Duncan Cance, Department of Kinesiology and Health Education, The University of Texas at Austin, 1 University Station D3700, Belmont Hall 222, Austin, Texas, 78712, jdance@mail.utexas.edu, phone: (512) 471-8184.

Negriff & Susman, 2011; Silbereise & Kracke, 1997). Questions regarding the stability of perceived pubertal timing are relevant not only to the measurement of pubertal timing but also to the period of risk associated with off-timing. If adolescents' perceptions of pubertal timing are unstable, as might be expected given the dynamic and individually variable process of pubertal development, then windows of risk associated with off-timing may be limited. In contrast, and as suggested by theories of identity development, if perceptions of pubertal timing are fixed early in adolescence and remain stable, the impact of these early perceptions of pubertal timing on behavior and well-being may persist throughout adolescence. Our purpose in this article is to examine the extent to which adolescents' perceptions of their pubertal timing are unstable or stable over the span of pubertal development, using a longitudinal sample of youth aged 11 to 17.

Overview of Pubertal Timing in Adolescence

Puberty is the process of developing from a child into a sexually mature adult (Hayward, 2003). Important physical changes occur during this process, from skeletal and nervous system growth to the development of the endocrine system, but the changes occurring to the reproductive organs and secondary sexual characteristics gain the most attention. Pubertal development during adolescence has been found to be highly salient, resulting in powerful emotions on the part of the adolescent as well as changing relationships with peers, parents, and teachers (Beausang & Razor, 2000; Forbes & Dahl, 2010; Lee, 1997; Silbereise & Kracke, 1997; Summers-Effler, 2004). There is wide variation in the onset and tempo of puberty by gender and race/ethnicity, as well as individual differences within groups (Archibald, Graber, & Brooks-Gunn, 2003; Chumlea et al., 2003; Marceau, Ram, Houts, Grimm, & Susman, 2011; Mendle, Harden, Brooks-Gunn, & Graber, 2010; Sisk & Foster, 2004;). This individual-level variation has drawn researchers to explore the impact of pubertal timing on adolescent risk behavior.

There are two common ways of establishing self-report pubertal timing. The first, referred to in this article as *stage-normative* pubertal timing, is based on an adolescent's pubertal status, which is a measure of how developed an adolescent is in relation to the pubertal development process. The adolescent's pubertal status is normed within study-defined demographic subgroups (typically, age, gender, and race/ethnicity). Adolescents can either be classified into categories as early, on-time, or late based on comparison with the average pubertal status of their demographic subgroup, or can be given a continuous score based on a regression including the demographic characteristics. Stage-normative pubertal timing is, therefore, based on the adolescent's assessment of his or her physical development, typically using several indicators. In contrast, the second measure, what is referred to in this article as *peer-normative* pubertal timing, is not based on pubertal status explicitly, but instead on the adolescent's perception of pubertal development timing relative to peers; adolescents are asked how they perceive their timing to be compared with their peers, typically using a Likert scaled measure. Thus, the peer-normative measure explicitly invokes a social comparison. In this article, both the stage-normative and peer-normative measures of pubertal timing are based on self-report and are therefore subject to bias compared to pubertal timing assessed by clinical means, such as physician examination or hormone concentrations (Dorn & Biro, 2011; Dorn, Dahl, Woodward, & Biro, 2006). But the peer-normative pubertal timing measure is considered to be a more subjective measure of self-report pubertal timing than stage-normative measures because it is not based on specific indicators of pubertal development but instead is based on the adolescent's overall assessment of her or his pubertal status and how that compares with peers. Thus, there is reason to believe that these two measures may be differentially assessing perceived pubertal timing.

Theoretical and Empirical Considerations on the Stability of Pubertal Timing

Most of the research on relationships between perceived pubertal timing and adolescent health has been cross-sectional and conducted with adolescents of varying ages, without consideration as to whether perceived pubertal timing is a stable construct throughout adolescence. If perceived pubertal timing is unstable, then we would not expect the relationships detected at one age to persist at other ages. As such, it would be important to examine the impact of pubertal timing on adolescent health and well-being at different times during adolescence. For example, we would expect that adolescents who perceive themselves as early developing would be the only adolescents at risk in early adolescence and would only be at risk until their peers “catch up” in pubertal development. Likewise, adolescents who perceive themselves as late developing would emerge as a risk group in late adolescence, when the majority of their peers have progressed through puberty. However, if perceived pubertal timing is stable, then the impact of perceived pubertal timing in early adolescence may be persistent throughout adolescence. In other words, how an adolescent feels about their pubertal timing in early adolescence could continue to impact their health and well-being throughout adolescence. It might then be possible to predict an adolescent’s risk profile based on perceived pubertal timing assessed in early adolescence.

How pubertal timing is measured may have implications for the likely stability or instability of the construct. Because pubertal development for most adolescents is ongoing, not only is the adolescent changing but their referent peer group is also changing. And there is emerging evidence that the onset of puberty is inversely related to puberty tempo, the time it takes to complete the pubertal development process (Biro et al., 2006; Marceau et al., 2011; Martí-Hennenberg & Vizmanos, 1997; Mendle et al, 2010). That is, early developing adolescents take longer to complete the pubertal development process compared with their later developing peers. This interaction between pubertal onset and tempo could have an impact on the stability of pubertal timing. For example, an adolescent who is classified as early developing at age 11 could be re-classified as on-time at age 15 if her peers have a faster tempo of puberty and thus “catch up” to the early developing adolescent. Therefore, it is plausible that stage-normative pubertal timing is unstable across adolescence because an adolescent’s pubertal status is changing and how his/her status compares to peers is changing as well.

However, empirical evidence only partially supports the hypothesis that stage-normative measures of pubertal timing lack stability. This could be because the studies, with one exception, only have used two waves of data to assess stability, and the analyses have not been stratified by age, despite a wide age range in the sample. Combining ages within a sample could mask the differences expected at the younger and older ages of the pubertal development process. For example, one study of adolescent males between the ages of 12 and 16 found that the correlation of stage-normative pubertal timing measured one year apart was .63 (Drapela, Gebelt, & McRee, 2006). Another study of adolescents aged 12 to 16 found that the correlation, this time measured two years apart, was strong and slightly different across gender (.82 for males and .87 for females; Wichstrom, 2001). In contrast, and more in line with the expectation that stage-normative pubertal timing is unstable, one study of early adolescents (aged 10 to 13) using a stage-normative measure found that about half of the off-time adolescents at baseline were reclassified as on-time one year later (a stability coefficient was not reported) (Wiesner & Ittel, 2002). Another study of stage-normative pubertal timing among twins 12 years of age at baseline found a substantial proportion switched from one category of timing (early, on-time, or late) to another at a two-year follow-up; although a stability coefficient was not reported, the proportion was deemed great enough that the authors chose to use seven categories of timing (consistently on-time, consistently early, consistently late, and four groups reflecting change from off-time to on-time and vice versa) as predictors of behavior (Dick, Rose, Pulkkinen, & Kapriro, 2001).

Thus, it is plausible that stage-normative pubertal timing is unstable but variation in sample ages has prevented a thorough test of the hypothesis.

Contrary to stage-normative pubertal timing, theory suggests that peer-normative pubertal timing should remain stable throughout adolescence. As discussed previously, peer-normative pubertal timing is not explicitly based on pubertal status but instead is reliant on a social comparison process. The importance of peer comparison is supported by person-in-context theory, which postulates that an adolescent's identity is formed based on an understanding of the contexts in which she or he is embedded (Adams & Marshall, 1996). In regards to peer-normative pubertal timing, an adolescent must engage in social comparison to determine how her or his pubertal status compares with peers. It is likely that such social comparison introduces a psychosocial component to the peer-normative pubertal timing measure that is missing from the stage-normative pubertal timing measure.

According to Erikson's theory of psychosocial development, adolescence is a developmental stage focused on the formation of personal identity, of which puberty plays an important role (Erikson, 1950). In order to establish ego identity – knowledge of who you are and how you fit into the broader society – the adolescent interacts and compares himself or herself to significant others, a process known as *psychosocial reciprocity* (Finkenauer, Engels, Meeus, & Oosterwegel, 2002; Muuss, 1996). According to Erickson, the social interactions occurring during pubertal development influence the adolescent's identity formation such that the perception of pubertal timing during this formative time is internalized and considered constant, regardless of actual pubertal development. It is possible that these early experiences in pubertal development become a part of adolescent identity, such that peer-normative pubertal timing would remain stable throughout adolescence.

However, two studies that used a peer-normative measure of perceived pubertal timing found lower stability than reported for stage-normative measures (Dubas, Graber, & Petersen, 1991; Graber, Lewinsohn, Seeley, & Brooks-Gunn, 1997). One study of high school students found the Kappa statistic for peer-normative pubertal timing assessed one year apart to be .61 for females and .48 for males (Graber et al., 1997). Only one study to date has examined the stability of peer-normative pubertal timing across more than two waves, and the analyses were limited to comparisons of two waves at a time (Dubas et al., 1991). This study found that females appeared to be more consistent in reporting their peer-normative pubertal timing over time compared with males, and that the correlation appears to strengthen as the age of the adolescent increases. The empirical evidence regarding the stability of peer-normative pubertal timing does not support the hypothesis that peer-normative pubertal timing is stable, and there appear to be stability differences depending on adolescent age and gender.

Methodological Considerations for the Measurement of Pubertal Timing Stability

When dealing with longitudinal data there are two conceptual approaches: the variable-centered approach and the person-centered approach (Laursen & Hoff, 2006). The underlying assumption in variable-centered approaches, such as regression and correlation analysis, is that the population is homogenous with respect to the variables of interest. In other words, the timing and tempo of pubertal development is the same for all adolescents. But, as discussed previously, pubertal development is not homogenous among adolescents but rather is a process that is individually variable. Person-centered approaches to longitudinal data allow for sample heterogeneity and are thus a better fit for understanding the stability of pubertal timing throughout adolescence.

Another important consideration in determining the stability of perceived pubertal timing throughout adolescence is the underlying assumptions of different person-centered

approaches. At issue is whether pubertal timing is a construct that has a reliable and distinguishable pattern across adolescence. Calculating stability using the intraclass correlation coefficient (ICC) in a random-effects ANOVA model is useful for longitudinal data (Bland & Altman, 1996). The ICC makes no assumption of an underlying pattern of responses over time, but instead calculates the average reliability of the measure of interest from one time point to the next. In contrast, the assumption of another person-centered analytic technique, latent class analysis (LCA), is that there are underlying response patterns in a sample; variation from the underlying patterns is treated as measurement error (B. Muthén & L. Muthén, 2000).

Due to the different assumptions, the two analytic techniques could result in different conclusions concerning measurement stability. For example, using the peer-normative pubertal timing measure, an adolescent may respond early at age 12, on-time at age 12.5, and early at age 13. There would thus be variation from ages 12 to 12.5 and from ages 12.5 to 13, but no variation from age 12 to 13. Many adolescents could have this slight variation in their perceived pubertal timing across adolescence. Using random effects ANOVA models, the within-subject variance would be high compared with the total variance, resulting in a lower ICC, leading to the conclusion that peer-normative pubertal timing is unstable. In contrast, using latent class analysis, the observed variation in pubertal timing is thought of as measurement error. That is, an adolescent has an underlying perception of pubertal timing and deviation from this perception is not a result of a change in perception but rather a random departure. This hypothesis can be tested in LCA by treating the adolescent as the unit of analysis and examining whether there are distinct classes of perceived pubertal timing that remain stable across adolescence. In the case above, the adolescent would have a high probability of being in an early developing class because two of the three responses were early. If there was a consistent pattern of change in perceived pubertal timing across adolescence (for instance, if a large proportion of adolescents believed they were early developers until they started high school when they switched to believing they were late developers) then the LCA would identify this response pattern as a class. Thus, there is reason to believe that stability of perceived pubertal timing may vary between the two person-centered analytic approaches.

Study Purpose and Hypotheses

The purpose of this study is to examine the stability of two measures of perceived pubertal timing in a diverse sample of rural adolescents aged 11 to 17, using two person-centered approaches. Person-centered approaches using longitudinal data are important because they allow for the assessment of dynamic relationships and provide the ability to understand the heterogeneity among subjects (Frees, 2004). However, assessment of the stability of pubertal timing has thus far been based on variable-centered approaches using limited longitudinal samples. First, the stability of the two pubertal timing measures, stage-normative pubertal timing and peer-normative pubertal timing, is examined using random effects ANOVA modeling. Latent class analysis is then conducted to explore the stability of both measures of pubertal timing by determining whether distinct patterns of perceived pubertal timing exist. We hypothesize that stage-normative pubertal timing will be less stable than peer-normative pubertal timing, due primarily to variations in the tempo of pubertal development, and there will be five distinct classes – always early, always on-time, always late, early in early adolescence moving to on-time in mid-adolescence, and on-time in early adolescence moving to late in mid-adolescence. We hypothesize that peer-normative pubertal timing will be stable and there will be three distinct classes of pubertal timing – early, on-time, and late.

This study also expands on previous research by examining whether demographic differences play a role in the stability of perceived pubertal timing. Based on the previously mentioned gender differences in pubertal development (e.g., Archibald et al., 2003; Marceau et al., 2011) and pubertal timing stability (Graber et al., 1997; Wichstrom, 2001), we hypothesize that both measures of perceived pubertal timing will be more stable among females compared with males. Past research has found racial and ethnic differences in pubertal status and pubertal timing (e.g., African-American adolescents develop earlier than White adolescents and perceive their development to be earlier compared with White youth; Archibald et al., 2003; Biro et al., 2010; Chumlea et al., 2003; Obeidallah et al., 2000; Sun et al., 2002), but no studies have looked at racial or ethnic differences in the stability of pubertal timing. The results from this study will thus contribute a better understanding of whether, and how, demographic differences in pubertal development may impact the stability of pubertal timing across adolescence.

Method

The Context Study

This study was conducted through the secondary analysis of five waves of data from the Context of Adolescent Substance Use study (Context Study), a school-based longitudinal study of three cohorts of adolescents from three North Carolina counties. Wave 1 began in the Spring of 2002 when adolescents were enrolled in the 6th to 8th grades and data collection occurred every semester until the Spring of 2004 (Wave 5). All adolescents in the grades of interest in the sampled schools (eight middle schools, two K-8 schools, six high schools, and three alternative schools) were considered eligible for participation. Adolescents in self-contained special education classes and adolescents who had English as a second language and had insufficient reading skills to complete the questionnaire in English were excluded from the study. Response rates ranged from 88% at Wave 1 to 76% at Wave 5.

The Context Study was approved by UNC's School of Public Health IRB in the Office of Human Research Ethics. The study received a waiver of written parental consent; written adolescent assent was obtained. Data were collected in a group setting in the schools using self-administered questionnaires. Each classroom had at least one data collector from the research team and larger classrooms were assigned two data collectors. Data collectors returned to the school on as many as four additional days after primary data collection to attempt to reach absent adolescents. Teachers remained in the classroom to maintain order among the students but, to protect confidentiality, teachers were not allowed to walk around the classroom during the data collection or answer student questions about the study. The completion time for the questionnaire was approximately one hour and there was no monetary compensation for participation in the study.

Study Sample

The current study is based on data from adolescents who participated in at least one wave of data collection (N=6,892). Approximately 13% of adolescents participated in one wave, 13% in two waves, 15% in three waves, 17% in four waves, and the majority, 42%, participated in all five waves of data collection. Participants missing information on age, gender, or race/ethnicity were excluded from analyses (N=295 excluded) and the sample was limited to adolescents aged 11 to 17 to only include students who were within the typical age range for their grade (N=172 excluded). Excluded adolescents (N=467, 7 percent of the total sample) were less likely to be White ($p<.001$), less likely to be in the other racial/ethnic category ($p<.05$), and more likely to be male ($p<.001$). Excluded adolescents were also less likely to have participated in all five waves of data collection ($p<.001$). The final

sample included 6,425 respondents (50 percent male, 53 percent White, 36 percent African-American, 4 percent Latino, and 7 percent indicating another racial/ethnic category). The mean age at Wave 1 was 13.1 (SD=0.97) and at Wave 5 was 15.0 (SD=0.92).

Measures

Stage-normative pubertal timing—Stage-normative pubertal timing was calculated based on a revised version of the Pubertal Development Scale (PDS) (Petersen, Crockett, Richards, & Boxer, 1988). The PDS consists of five questions each for boys and girls assessing development of body hair growth, skin changes and height for males and females, facial hair and voice changes for males, and breast development for females. The range of the items is 1=not yet started to 4=seems complete. Females are also asked if they started menstruating (1=no, 4=yes). The items were averaged to obtain a mean PDS score (alphas by wave ranged from 0.68 to 0.73 for females and 0.76 to 0.81 for males). To measure stage-normative pubertal timing, we first calculated the mean pubertal stage among adolescents in the sample by age, gender, and race/ethnicity. We then compared each adolescent's pubertal stage to the mean for their demographic subgroup. Adolescents were classified as "early" (1=more than one standard deviation above the mean pubertal stage), "on-time" (0), or "late" (-1=more than one standard deviation below the mean pubertal stage) based on the norm for their demographic subgroup.

Peer-normative pubertal timing—Peer-normative pubertal timing is based on adolescent perceptions of their pubertal development relative to their peers. Adolescents were asked one item about how they believe their physical development compared with others their own age and sex (1=much earlier to 5=much later). Adolescents indicating their development was much or somewhat earlier than their peers were classified as "early" (1), about the same as their peers as "on-time" (0), and somewhat or much later than their peers as "late" (-1) developers.

Demographic variables—Age was calculated using adolescent date of birth and the date of the interview. Age was recoded into twelve half-year categories, ranging from 11 to 16.5. Race/ethnicity was recoded into four categories: White, Black or African-American, Hispanic or Latino, and Other (including American Indian or Native American, Asian or Pacific Islander, multiracial, other, and adolescents who answered don't know).

Analyses

For all analyses, the sample was configured to use age as the unit of time instead of wave of data collection. Because The Context Study was a longitudinal study of cohorts in three different grades at baseline, there is wide variation in age at each wave, which would be ignored in analyses based on data collection wave.

To test whether an individual's perceived pubertal timing is stable across the ages of 11 to 16.5, a series of one-way random effects ANOVA models was conducted using SAS 9.1 with each measure of pubertal timing separately. The sample was then split by gender and race/ethnicity to determine if there were demographic differences in the stability of pubertal timing.

A random-effects ANOVA model is different from a standard one-way ANOVA model such that the grouping variable is treated as a level of nesting, not a fixed effect. In longitudinal studies such as the current research, the grouping variable is the individual. There are a number of benefits to using the random-effects ANOVA model for calculating stability over traditional methods, such as the Pearson correlation coefficient (Baumgartner, 2000). The random effects model allows for more than two scores per individual. Individuals are not

assumed to have equally spaced measurement and do not have to have values on the same number of time points. It is therefore possible to include adolescents who have missed one or more of the measurements, which allows for the retention of the full analytic sample, decreasing the chances of selection bias.

Between-group and within-group differences were determined using the intraclass correlation coefficient (ICC). The range of the ICC is from 0 to 1. An ICC closer to 1 indicates that the adolescent's perception of their pubertal timing (early, on-time, or late) does not change over time. An ICC below .40 indicates poor stability, between .40 and .59 is fair, between .60 and .74 is good, and between .75 and 1.00 indicates the measure shows excellent stability (Cicchetti, 1994).

Latent class analysis (LCA) was used to determine stability via the underlying patterns of stage-normative and peer-normative perceived pubertal timing. The goal of LCA is to determine if subgroups or classes of individuals exist based on their patterns of item response (B. Muthén & L. Muthén, 2000). The result is a set of latent classes where the membership within a class is more homogenous than between classes. However, individual membership in a specific class is not definite but is stated in terms of a probability estimate. In other words, LCA indicates how likely it is that each individual belongs to each class.

While LCA is typically used with cross-sectional data, and is often used to assess the reliability of multiple measures within a cross-sectional dataset, LCA can be expanded to longitudinal data (see Biemer & Wiesen, 2002; Flaherty, 2002 for examples). With LCA, it is possible to determine if there are classes, using the longitudinal response pattern as the unit of analysis, while accounting for potential measurement error. LCA can also be used to analyze categorical data, such as the perceived pubertal timing measures in this study. As with the random-effects ANOVA model, it is possible to retain the full analytic sample, even if adolescents have missed one or more measurements, by using full information maximum likelihood estimation.

The first step in the latent class analysis was to test a single-class latent growth curve model to assess the underlying structure of the overall means. The next step was to determine the number of classes for each measure of perceived pubertal timing. One methodological debate regarding LCA concerns whether the determined number of classes is accurate or is biased by the properties of the measure under analysis (Bauer & Curran, 2003). To lessen the likelihood of misspecification, the number of classes was determined using theoretical justifications in combination with fit indices (Bauer & Curran, 2003; Jung & Wickrama, 2007). The fit indices used in this study included the Bayesian information criteria value (BIC) and the Lo, Mendell, and Rubin likelihood ratio test (LMR-LRT). The model with the lowest BIC and a significant LMR-LRT *p*-value compared with a model with one fewer classes was considered the best fitting model. In addition, the best fitting model should successfully converge, have an entropy value close to 1, have greater than 1% of the population in each class, and have posterior probabilities close to 1 (Jung & Wickrama, 2007). For all models the variances of the slope and intercept were set to zero for all classes. After determining the number of classes, the sample and estimated means of peer-normative and stage-normative perceived pubertal timing at each age for the three classes were examined. If the estimated classes are a perfect fit the sample means and estimated means should not differ. The posterior probabilities, which can be interpreted as the reliability of class assignment, were also examined. The latent class analyses were conducted using MPlus Version 5.1 (B. Muthén & L. Muthén, 2001).

Results

Table 1 presents the means and standard deviations for the two perceived pubertal timing measures by age and gender. The means of stage-normative pubertal timing were close to zero at all ages, which is to be expected because the stage-normative measure was normed by age, gender, and race/ethnicity. The means of the peer-normative pubertal timing measure were in general more positive than the stage-normative measure, suggesting that adolescents were on average likely to perceive themselves as early developing compared with their peers. In general, with both measures, the means increased with increasing age, indicating that adolescents were more likely to perceive themselves as early developing compared with their peers as they aged. The standard deviations were higher with the peer-normative measure compared with the stage-normative measure. Overall, based on the ICCs from the random effects ANOVA models, the two measures of pubertal timing showed poor to fair stability (ICC=.400 for stage-normative and ICC=.388 for peer-normative, Table 2).

Latent Class Analysis

Because the sample means tended to increase with increasing age, a linear model was tested to assess the underlying structure of peer-normative and stage-normative perceived pubertal timing for the overall sample. The linear model was a good fit for both the peer-normative pubertal timing data (CFI=.98, TLI=.99, RMSEA=.015 (.011-.019)) and the stage-normative pubertal timing data (CFI=.95, TLI=.96, RMSEA=.028 (.025-.031)).

As shown in Table 3, the three-class solution was the best fit for both measures of pubertal timing. In order to interpret the latent classes, the sample and estimated means of peer-normative and stage-normative perceived pubertal timing at each age for the three classes are presented in Figures 1 and 2. For the stage-normative measure, the difference between the sample and estimated means was greatest in the late developing class, compared with the on-time and early developing classes. For the peer-normative measure, the estimated means were most different from the sample means at the youngest ages in the sample.

Based on an examination of the estimated means, the three classes were interpreted as “always early” (Class 1), “always on-time” (Class 2), and “always late” (Class 3). Table 4 presents the percentage of adolescents in each class and the average probability of membership for each class for both the peer-normative and the stage-normative measures of perceived pubertal timing. More adolescents had a probability of being in the early class using the peer-normative measure (28%) compared with the stage-normative measure (13%). However, there was little difference between the two measures in the probability of being in the late class (12% using the peer-normative measure vs. 13% using the stage-normative measure). The posterior probabilities for class membership were relatively high for both measures (above .80), but were higher for the stage-normative measure, indicating that the stage-normative measure may be more stable than the peer-normative measure of perceived pubertal timing.

A final exploratory step was to determine if gender or race/ethnicity predicted class membership. There were no gender or racial/ethnic differences in class membership for the stage-normative pubertal timing measure. Using the peer-normative measure, however, Black adolescents were more likely than White adolescents ($p<.001$) and Latino adolescents ($p=.022$) to be classified as early developing. Female adolescents were more likely than male adolescents to be classified as late developing ($p<.001$).

Discussion

It is well established that pubertal development is an individually variable process (e.g., Archibald et al., 2003; Hayward, 2003). What is less understood is whether perceptions of pubertal timing, the relative development of an adolescent compared with same age and gender peers, also vary throughout adolescence. And there is reason to expect that conclusions reached about the stability of perceived pubertal timing may vary based on how perceived pubertal timing is measured (Dubas et al., 1991; Graber et al., 1997). The purpose of this study was to determine the stability of two measures of perceived pubertal timing throughout adolescence, and results confirmed the complexity of assessing pubertal development timing in adolescence. We found important differences between stage-normative and peer-normative pubertal timing, which suggest that these measures assess different aspects of pubertal development. We also found that the question of whether pubertal timing was stable across adolescence depended on the analytic method used to determine stability. Perceived pubertal timing does vary from one time point to the next, and it appears that the highest variability in assessment occurs in early adolescence. As such, researchers should be cautious in assuming that cross-sectional findings using pubertal timing are comparable across adolescence. However, when taking into account the full pattern of perceived pubertal timing across adolescence, three stable classes emerge, suggesting that at least some of the variability in perceived pubertal timing is due to measurement error. The findings from this study emphasize the need to use analyses that can take into account the longitudinal patterns of perceived pubertal timing rather than looking at pubertal timing at one age. Furthermore, the results from this study suggest that how an adolescent perceives their pubertal timing in early adolescence holds throughout adolescence, supporting the hypothesis that pubertal timing is part of identity development.

Both the stage-normative and peer-normative measures of perceived pubertal timing showed poor to fair stability over time in the random effects ANOVA models. This means that perceived pubertal timing, either stage-normative or peer-normative, is likely different depending on the age of assessment. The extreme instability in both measures of perceived pubertal timing was contrary to our hypotheses based on previous literature (Dubas et al., 1991; Graber et al., 1997). This could be due in part to the larger age range of our study sample, but is more likely the result of using person-centered analyses versus variable-centered analyses. Unlike the previous studies of pubertal timing stability that looked at the sample mean differences from one time point to the next, this study was able to look at individual differences in perceived pubertal timing across adolescence. While there appeared to be some differences in the ICCs across pubertal timing measure and demographic subgroups, both measures showed poor to fair stability overall so the differences likely do not have practical implications.

Despite the poor stability of the two perceived pubertal timing measures in the random effects ANOVA models, the latent class analyses (LCA) showed three distinct response patterns – always early, always on-time, and always late. This finding of stability in the latent class analysis was contrary to hypotheses for the stage-normative measure but confirmed the hypotheses for the peer-normative measure. The key reason for these differing results from the random effects ANOVA models is that LCA takes into account measurement error. The measurement error in this study can be thought of as an adolescent's deviation from their "true", or most commonly answered, pubertal timing response. The difference in the two analyses demonstrates the importance of utilizing more sophisticated analyses to understand measurement stability. The results from this study show that, in general, adolescents may have variation in their perceived pubertal timing (both stage-normative and peer-normative) from one time point to the next, which results in low ICCs.

But when looking at the full pattern of responses across adolescence using LCA, perceived pubertal timing remains relatively stable.

In some ways, the LCA models support the multilevel stability analyses, in that the stage-normative measure appeared to be more stable than the peer-normative measure. The sample and estimated means were more closely aligned with the stage-normative measure compared with the peer-normative measure. And the posterior probabilities, which can be thought of as a test of the reliability of classification, were higher with the stage-normative measure compared with the peer-normative measure. It is possible that the truncated age range of the sample resulted in the appearance of more stability in the stage-normative measure. However, the posterior probability differences in the LCA were not that profound between the stage-normative and peer-normative measures. While there may be slightly more variation in adolescent responses in the peer-normative measure compared with the stage-normative measure, both can be good assessments of adolescent perceived pubertal timing when measurement error is taken into account with LCA.

An important finding of the latent class analysis was the difference in proportions in latent class membership and difference in predictors of class membership between the two measures of pubertal timing. The proportions of adolescents who were classified as late developing were similar across the two measures but adolescents were twice as likely to classify themselves as early developing using the peer-normative measure compared with the more objective classification of the stage-normative measure. This could be due to a social desirability for earlier development, especially for male and Black adolescents (Cohane & Pope, 2001; Graber et al., 1997; Lynne et al., 2007; Obeidallah et al., 2000; Siegel, Yancey, Aneshensel, & Schuler, 1999). These differences also demonstrate that while the stability of the two measures may be similar, the two measures are assessing different aspects of pubertal development.

There are limitations to the current research. The youngest adolescents in the study sample were 11 years of age, but the first stages of pubertal development typically begin by age 9 or 10 (Tanner, 1962). In addition, the oldest adolescents in the study sample were up to 17 years old, which is, on average, prior to the completion of pubertal development. Therefore, this study is an examination of the stability of perceived pubertal timing during the midst of pubertal development. Future studies should be conducted to include mid-childhood and early-adulthood ages to determine if the stability of pubertal timing differs when including the full pubertal development process. It is possible that we did not find the two additional latent classes proposed for the stage-normative measure (transitioning from early to on-time and from on-time to late) because this sample is lacking information from late childhood and late adolescence. It is possible that there were demographic differences in the predictors of latent class probability between the stage-normative and peer-normative measures of pubertal timing because the peer-normative measure only asked adolescents to compare their development to peers of the same age and gender and did not mention race or ethnicity while the stage-normative measure was developed within age, gender, and racial/ethnic group. It would be worthwhile to compare the peer-normative measure used in this study with a measure that asks specifically about age, gender, and race/ethnicity.

Another consideration to note is the decision to trichotomize both measures of pubertal timing into early, on-time, and late. Because one of the goals of this study was to better understand if pubertal timing assessed at one age could be compared with pubertal timing assessed at another age, we chose to use the stage-normative and peer-normative measures as they are most commonly used in the literature (e.g., Bratberg et al., 2007; Dubas et al., 1991; Ge et al. 2003; Siegel et al., 1999; Tremblay & Frigon, 2005). However, we tested the analyses using a continuous measure of stage-normative pubertal timing (based on the

regression residuals including age, gender, and race/ethnicity) and using the five-category peer-normative measure and found similar results. More research is needed to further explore how other calculations of pubertal timing (e.g., using cutoffs other than one standard deviation) may impact the stability of pubertal timing.

The most controversial aspect of LCA is the determination of the number of classes. Misspecification of the number of classes could dramatically alter the study conclusions. However, in this study, we followed recommendations of using a combination of theoretical justifications and statistical tests in order to determine the number of classes (Bauer & Curran, 2003; Jung & Wickrama, 2007; Nylund, Asparouhov, & Muthén, 2007). Due to computational burden the bootstrap likelihood ratio test (BLRT) could not be calculated for the 4-class or higher models (for the 2-class and 3-class models the p-value was less than 0.001, suggesting there were at least 3 classes). But the BIC has been shown to be accurate 100 percent of the time when the sample size is 1,000, much lower than this study's sample size of over 6,000 (Nylund et al., 2007). The 3-class model also makes theoretical sense when looking at the estimated means (e.g., the majority of adolescents had the greatest likelihood of being in the on-time class). Based on this information, we are confident in the conclusion that, in this sample of adolescents, the 3-class model was the best fitting model for both measures of pubertal timing.

This study has important implications for future research involving pubertal development. The results confirm that while both of the measures of perceived pubertal timing, peer-normative and stage-normative, are based on self-report, they are assessing different aspects of pubertal development. More research is needed to understand whether and how these two measures may be differentially related to adolescent risk behavior. Furthermore, the differing results of the stability analyses demonstrate that assessment of pubertal timing at one age is not necessarily the same as at a different age. However, when incorporating the longitudinal patterns of responses, and treating the variability as measurement error, stable pubertal timing classes emerge. The results from this study imply that the longitudinal patterns of pubertal timing should be used as the predictor of adolescent health, not pubertal timing assessed at a single point in time. And these longitudinal patterns suggest that the experiences of pubertal development early in adolescence remain relatively constant throughout adolescence. We would then expect that the impact of pubertal timing on adolescent well-being and health risk behaviors may persist throughout adolescence; future research should test this hypothesis. Another implication of this research is that interventions aimed at alleviating the deleterious effects of off-timing could be relevant throughout adolescence. In conclusion, the findings from this study demonstrate that while pubertal development is a dynamic process, perceptions of pubertal timing based on early adolescent experiences are stable throughout adolescence and contribute to adolescent identity development.

Acknowledgments

This research was supported by the National Institutes of Health under Ruth L. Kirschstein National Research Service Award F31-DA023326 from the National Institute on Drug Abuse (Cance). The Context of Adolescent Substance Use Study was funded by the National Institute on Drug Abuse R01-DA13459 (Ennett).

References

- Adams GR, Marshall SK. A developmental social psychology of identity: understanding the person-in-context. *Journal of Adolescence*. 1996; 19:429–442. [PubMed: 9245296]
- Archibald, AB.; Graber, JA.; Brooks-Gunn, J. Pubertal processes and physiological growth in adolescence. In: Adams, GR.; Berzonsky, MD., editors. *Handbook of Adolescence*. Oxford, UK: Blackwell Publishers; 2003. p. 24-48.

- Bauer DJ, Curran PJ. Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods*. 2003; 8:338–363. [PubMed: 14596495]
- Baumgartner TA. Estimating the stability reliability of a score. *Measurement in physical education and exercise science*. 2000; 4:175–178.
- Beausang CC, Razor AG. Young western women's experiences of menarche and menstruation. *Health Care for Women International*. 2000; 21:517–528. [PubMed: 11235283]
- Biemer PP, Wiesen C. Measurement error evaluation of self-reported drug use: a latent class analysis of the US National Household Survey on Drug Abuse. *Journal of the Royal Statistical Society: Series A*. 2002; 165:97–119.
- Biro FM, Galvez MP, Greenspan LC, Succop PA, Vangeepuram N, Pinney SM, Teitelbaum S, Windham GC, Kushi LH, Wolff MS. Pubertal assessment method and baseline characteristics in a mixed longitudinal study of girls. *Pediatrics*. 2010; 126(3):e583–e590. [PubMed: 20696727]
- Biro FM, Huang B, Crawford P, Lucky A, Stiegel-Moore R, Barton BA, et al. Pubertal correlates in black and white girls. *The Journal of Pediatrics*. 2006; 148:234–240. [PubMed: 16492435]
- Bland JM, Altman DG. *Statistics Notes: Measurement error and correlation coefficients*. *British Medical Journal*. 1996; 313:41–42. [PubMed: 8664775]
- Bratberg GH, Nilsen TIL, Holmen TL, Vatten LJ. Perceived pubertal timing, pubertal status, and the prevalence of alcohol drinking and cigarette smoking in early and late adolescence: A population based study of 8950 Norwegian boys and girls. *Acta Paediatrica*. 2007; 96:292–295. [PubMed: 17429923]
- Chumlea WC, Schubert CM, Roche AF, Kulin HE, Lee PAMD, Himes JH, Sun SS. Age at menarche and racial comparisons in US girls. *Pediatrics*. 2003; 111:110–113. [PubMed: 12509562]
- Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normal and standardized assessment instruments in psychology. *Psychological Assessment*. 1994; 6:284–290.
- Cohane GH, Pope HG Jr. Body image in boys: A review of the literature. *International Journal of Eating Disorders*. 2001; 29:373–379. [PubMed: 11285574]
- Dick DM, Rose RJ, Pulkkinen L, Kaprio J. Measuring puberty and understanding its impact: A longitudinal study of adolescent twins. *Journal of Youth and Adolescence*. 2001; 30:385–399.
- Dorn LD, Biro FM. Puberty and its measurement: A decade in review. *Journal of Research on Adolescence*. 2011; 21:180–195.
- Dorn LD, Dahl RE, Woodward HR, Biro F. Defining the boundaries of early adolescence: A user's guide to assessing pubertal status and pubertal timing in research with adolescents. *Applied Developmental Science*. 2006; 10:30–56.
- Drapela LA, Gebelt JL, McRee N. Pubertal development, choice of friends, and smoking initiation among adolescent males. *Journal of Youth and Adolescence*. 2006; 35:717–727.
- Dubas JS, Graber JA, Petersen AC. The effects of pubertal development on achievement during adolescence. *American Journal of Education*. 1991; 99:444–460.
- Erikson, EH. *Childhood and society*. New York: W.W. Norton & Company; 1950.
- Finkenauer, C.; Engels, RCME.; Meeus, WHJ.; Oosterwegel, A. Self and identity in early adolescence. In: Brinthaupt, TM.; Lipka, RP., editors. *Understanding Early Adolescent Self and Identity*. Albany, NY: State Univ of New York Press; 2002. p. 25-35.
- Flaherty BP. Assessing the reliability of categorical substance use measures with latent class analysis. *Drug and Alcohol Dependence*. 2002; 68:S7–S20. [PubMed: 12324172]
- Forbes EE, Dahl RE. Pubertal development and behavior: Hormonal activation of social and motivational tendencies. *Brain and Cognition*. 2010; 72:66–72. [PubMed: 19942334]
- Frees, EW. *Longitudinal and Panel Data: Analysis and Applications for the Social Sciences*. Cambridge, UK: Cambridge University Press; 2004.
- Ge X, Kim IJ, Brody GH, Conger RD, Simons RL, Gibbons FX, Cutrona CE. It's about timing and change: Pubertal transition effects on symptoms of major depression among African American youths. *Developmental Psychology*. 2003; 3:430–439. [PubMed: 12760513]

- Graber JA, Lewinsohn PM, Seeley JR, Brooks-Gunn J. Is psychopathology associated with the timing of pubertal development? *Journal of the American Academy of Child and Adolescent Psychiatry*. 1997; 36:1768–1776. [PubMed: 9401339]
- Hayward, C., editor. *Cambridge studies on child and adolescent health*. New York: Cambridge University Press; 2003. Gender differences at puberty.
- Jung T, Wickrama KAS. Recent advances in longitudinal data analysis in social and psychological research: An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*. 2007; 2:302–331.
- Laursen B, Hoff E. Person-centered and variable-centered approaches to longitudinal data. *Journal of Developmental Psychology*. 2006; 52:377–389.
- Lee J. Never innocent: Breast experiences in women's bodily narratives of puberty. *Feminism & Psychology*. 1997; 7:453–474.
- Lynne S, Graber J, Nichols T, Brooks-Gunn J, Botvin G. Links between pubertal timing, peer influences, and externalizing behaviors among urban students followed through middle school. *Journal of Adolescent Health*. 2007; 40:181.e7–181.e13. [PubMed: 17259062]
- Marceau K, Ram N, Houts RM, Grimm KJ, Susman EJ. Individual differences in boys' and girls' timing and tempo of puberty: Modeling development with nonlinear growth models. *Developmental Psychology*. 2011
- Martí-Hennenberg C, Vizmanos B. The duration of puberty in girls is related to the timing of its onset. *The Journal of Pediatrics*. 1997; 131:618–621. [PubMed: 9386670]
- Mendle J, Harden KP, Brooks-Gunn J, Graber JA. Development's tortoise and hare: Pubertal timing, pubertal tempo, and depressive symptoms in boys and girls. *Developmental Psychology*. 2010; 46:1341–1353. [PubMed: 20822243]
- Muthén BO, Muthén LK. Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*. 2000; 24:882–891.
- Muthén, BO.; Muthén, LK. *MPlus User's Guide*. Los Angeles: Authors; 2001.
- Muuss, RE. Erik Erikson's theory of identity development. In: Muuss, RE., editor. *Theories of Adolescence*. 6. New York: McGraw-Hill; 1996. p. 42-57.
- Negriff S, Susman EJ. Pubertal timing, depression, and externalizing problems: A framework, review, and examination of gender differences. *Journal of Research on Adolescence*. 2011; 21:717–746.
- Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*. 2007; 14:535–569.
- Obeidallah DA, Brennan RT, Brooks-Gunn J, Kindlon D, Earls F. Socioeconomic status, race, and girls' pubertal maturation: Results from the Project on Human Development in Chicago neighborhoods. *Journal of Adolescent Research*. 2000; 10:443–464.
- Petersen AC, Crockett L, Richards M, Boxer A. A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*. 1988; 17:117–133.
- Siegel J, Yancey A, Aneshensel C, Schuler R. Body image, perceived pubertal timing, and adolescent mental health. *Journal of Adolescent Health*. 1999; 25:155–165. [PubMed: 10447043]
- Silbereise, RK.; Kracke, B. Self-reported maturational timing and adaptation in adolescence. Health risks and developmental transitions during adolescence. In: Schulenberg, J.; Maggs, JL.; Hurrelmann, K., editors. *Health Risks and Developmental Transitions During Adolescence*. New York, NY, US: Cambridge University Press; 1997. p. 85-109.
- Sisk CL, Foster DL. The neural basis of puberty and adolescence. *Nature Neuroscience*. 2004; 7:1040–1047.
- Summers-Effler E. Little girls in women's bodies: Social Interaction and the strategizing of early breast development. *Sex Roles: A Journal of Research*. 2004; 51:29–44.
- Sun SS, Schubert CM, Chumlea WC, Roche AF, Kulin HE, Lee PAMD, Himes JH, Ryan AS. National estimates of the timing of sexual maturation and racial differences among US children. *Pediatrics*. 2002; 110:911–919. [PubMed: 12415029]

- Tanner, JM. Growth at Adolescence: With a General Consideration of Effects of Hereditary and Environmental Factors upon Growth and Maturation from Birth to Maturity. Oxford, England: Blackwell Scientific Publications; 1962.
- Tremblay L, Frigon JY. The interaction role of obesity and pubertal timing on the psychosocial adjustment of adolescent girls: Longitudinal data. *International Journal of Obesity*. 2005; 29:1204–1211. [PubMed: 16155588]
- Wichstrom L. The impact of pubertal timing on adolescents' alcohol use. *Journal of Research on Adolescence*. 2001; 11:131–150.
- Wiesner M, Ittel A. Relations of pubertal timing and depressive symptoms to substance use in early adolescence. *Journal of Early Adolescence*. 2002; 22:5–23.

Biographies

Jessica Duncan Cance is an Assistant Professor in the Department of Kinesiology and Health Education at the University of Texas at Austin. Dr. Cance received her Ph.D. in Health Behavior and Health Education from the Gillings School of Global Public Health at the University of North Carolina at Chapel Hill and her Master's in Public Health in Behavioral Sciences and Health Education from the Rollins School of Public Health at Emory University. She also received a predoctoral Ruth L. Kirschstein National Research Service Award (NRSA) fellowship from the National Institute on Drug Abuse. Her research is focused on how the longitudinal interaction of biological, psychological, and social factors impacts adolescent health risk behaviors, including substance use and aggression.

Susan Ennett is a Professor in the Department of Health Behavior and Health Education at the Gillings School of Global Public Health at the University of North Carolina at Chapel Hill. Her research focuses on understanding and preventing health risk behaviors among adolescents in the general population as well as among high-risk youth. Her current research includes a longitudinal study of the social context of adolescent alcohol, tobacco, and other drug use and a national study of substance use prevention practices in middle schools. In other recent research, she has studied substance use and risky sexual behavior among runaway and homeless youth and collaborated on the development and evaluation of a family intervention to prevent youth tobacco and alcohol use.

Antonio Morgan-Lopez is an Associate Professor of Psychology in the L.L. Thurstone Psychometric Laboratory in the Department of Psychology, UNC-Chapel Hill. His overall program of research has focused on advances in measurement and analysis in behavioral drug and alcohol etiology, prevention, and treatment. He has published in the areas of experimental and quasi-experimental design in substance use/abuse prevention and treatment, advances in statistical mediation analysis, substance use and mental health among vulnerable populations and advances in modeling non-ignorable missing data.

Vangie Foshee is a Professor in the Department of Health Behavior and Health Education at the Gillings School of Global Public Health at the University of North Carolina at Chapel Hill. Her research is on adolescent problem behaviors. Her etiological research has included identifying determinants of violence between adolescent dating couples, adolescent cigarette smoking and alcohol use, and adolescent sexual behavior. Intervention studies have included development and evaluation of *Safe Dates*, a school and community-based adolescent dating violence prevention program, and development and evaluation of a family-based adolescent substance abuse prevention program *Family Matters*.

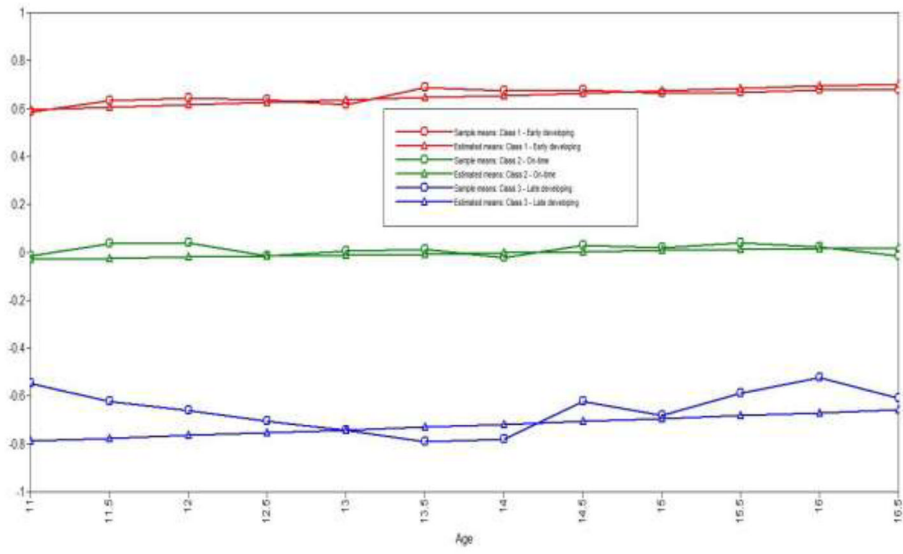


Figure 1. Sample and estimated means of stage-normative pubertal timing by class (n=6,392)

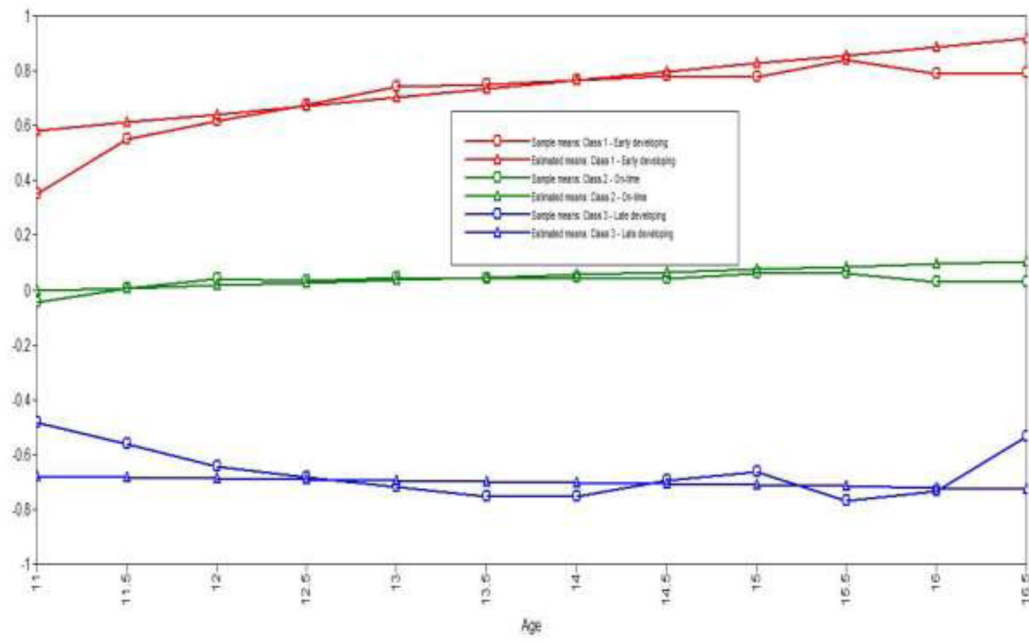


Figure 2.
Sample and estimated means of peer-normative pubertal timing by class (n=6,292)

Table 1

Means and standard deviations of stage-normative and peer-normative pubertal timing, by age and gender

Age	Stage-normative pubertal timing		Peer-normative pubertal timing		Males		Females	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
11	-.029	.484	.010	.618	.063	.710	-.125	.692
11.5	.008	.617	.044	.617	.104	.651	.002	.702
12	.061	.558	-.014	.601	.118	.667	.074	.690
12.5	.020	.583	-.065	.572	.135	.663	.057	.711
13	.037	.593	-.063	.513	.163	.659	.087	.692
13.5	-.025	.514	.020	.534	.135	.686	.122	.686
14	-.032	.521	-.028	.512	.159	.674	.131	.672
14.5	.028	.518	.040	.530	.184	.688	.148	.672
15	.003	.520	.037	.523	.168	.684	.203	.671
15.5	.044	.539	.039	.549	.172	.670	.215	.647
16	.002	.517	.058	.621	.198	.655	.195	.658
16.5	.034	.540	-.102	.443	.243	.664	.143	.645

Range is -1 to 1 (-1=late, 0=on-time, 1=early)

Table 2

Intraclass correlations (ICC) for stage-normative and peer-normative pubertal timing, by gender and race/ethnicity

	N	Stage-normative pubertal timing	Peer-normative pubertal timing
Full sample	6425	.400	.388
<i>White</i>	3393	.424	.443
<i>Black</i>	2335	.374	.323
<i>Latino</i>	254	.302	.213
<i>Other</i>	443	.386	.328
Female	3212	.425	.441
<i>White</i>	1672	.437	.493
<i>Black</i>	1185	.423	.383
<i>Latino</i>	117	.339	.194
<i>Other</i>	238	.388	.238
Male	3213	.374	.327
<i>White</i>	1721	.411	.391
<i>Black</i>	1150	.316	.243
<i>Latino</i>	137	.275	.228
<i>Other</i>	205	.387	.182

Note: ICC<.40 indicates poor stability, .40–.59 is fair, .60–.74 is good, and .75–1.00 is excellent

Table 3

Fit indices for Latent Curve Analysis (LCA) models with 1–5 classes, by measure of perceived pubertal timing

	LL	BIC	LMR-LRT	Entropy
<i>Stage-normative pubertal timing</i>				
1	-17892.2	35907.1	--	--
2	-16523.6	33196.2	<.0001	.779
3	-15488.7	31152.7	<.0001	.808
4	-15417.8	31037.1	.4646	.779
5	Did not converge			
<i>Peer-normative pubertal timing</i>				
1	-21444.8	43012.1	--	--
2	-20055.8	40260.3	<.0001	.588
3	-19437.5	39049.9	<.0001	.687
4	-19426.2	39053.7	.3487	.669
5	-19419.4	39066.3	1.0000	.693

LL=log likelihood, BIC=Bayesian information criteria value, LMR-LRT=Lo, Mendell, and Rubin likelihood ratio test

Bolded row indicates best fitting model

Table 4

Membership in latent classes and posterior probabilities

	Stage-normative pubertal timing		Peer-normative pubertal timing	
	%	Posterior probability	%	Posterior probability
Early	13%	.86	28%	.85
On-time	74%	.94	60%	.86
Late	13%	.86	12%	.82