



Published in final edited form as:

Dev Psychol. 2014 December ; 50(12): 2715–2726. doi:10.1037/a0038096.

Modeling Pubertal Timing and Tempo and Examining Links to Behavior Problems

Adriene M. Beltz,

Department of Psychology, The Pennsylvania State University, University Park, PA 16802

Robin P. Corley,

Institute for Behavioral Genetics, University of Colorado, Boulder CO 80309

Josh B. Bricker,

Institute for Behavioral Genetics, University of Colorado, Boulder CO 80309

Sally J. Wadsworth, and

Institute for Behavioral Genetics, University of Colorado, Boulder CO 80309

Sheri A. Berenbaum

Departments of Psychology and Pediatrics, The Pennsylvania State University, University Park, PA 16802

Abstract

Research on the role of puberty in adolescent psychological development requires attention to the meaning and measurement of pubertal development. Particular questions concern the utility of self report, the need for complex models to describe pubertal development, the psychological significance of pubertal timing versus tempo, and sex differences in the nature and psychological significance of pubertal development. We used longitudinal self-report data to model linear and logistic trajectories of pubertal development, and used timing and tempo estimates from these models, and from traditional approaches (age at menarche and time from onset of breast development to menarche), to predict psychological outcomes of internalizing and externalizing behavior problems, and early sexual activity. Participants (738 girls, 781 boys) reported annually from ages 9 through 15 on their pubertal development, and they and their parents reported on their behavior in mid-to-late adolescence and early adulthood. Self reports of pubertal development provided meaningful data for both boys and girls, producing good trajectories, and estimates of individuals' pubertal timing and tempo. A logistic model best fit the group data. Pubertal timing was estimated to be earlier in the logistic compared to linear model, but linear, logistic, and traditional estimates of pubertal timing correlated highly with each other and similarly with psychological outcomes. Pubertal tempo was not consistently estimated, and associations of tempo with timing and with behavior were model dependent. Advances in modeling facilitate the study

Correspondence concerning this article should be addressed to Sheri Berenbaum, Department of Psychology, The Pennsylvania State University, University Park, PA 16802; sab31@psu.edu.

Portions of the data on age at sexual initiation were included in Josh Bricker's doctoral dissertation and were presented at the 2010 biennial meeting of the Society for Research on Adolescence; some of the work described in the paper was presented at the 2013 biennial meeting of the Society for Research in Child Development.

of some questions about pubertal development, but assumptions of the models affect their utility in psychological studies.

Keywords

growth curves; early puberty; puberty self report; adolescence; behavior problems

The pubertal transition is important for psychological development in adolescence (Dorn, Dahl, & Biro, 2006; Lenroot & Giedd, 2010; Susman & Dorn, 2009). There is a long-standing interest in psychological consequences of variations in pubertal development, particularly elevated risks of early maturation in girls for depression and externalizing behavior problems (e.g., delinquency, substance use, early sexual activity; for reviews, see Ge & Natsuaki, 2009; Graber, 2013; Negri & Susman, 2011). Recent work has highlighted the role of puberty in normative adolescent development (e.g., increased risk taking), and in triggering psychopathology in vulnerable individuals; these changes are hypothesized to reflect effects of sex hormones and social experiences acting on the developing brain and stress systems (e.g., Gunnar, Wewerka, Frenn, Long, & Griggs, 2009; Trotman et al., 2013).

Progress on understanding mechanisms linking puberty to behavior depends on defining the aspects of puberty that matter and measuring them well, and this is the issue we address in this paper. Puberty is not unitary (Styne & Grumbach, 2011) and it is likely that aspects affect psychological development in different ways. For example, psychological risk associated with early pubertal timing mediated by peers would be associated with an aspect of pubertal development apparent to others, whereas risk associated with development of limbic-mediated motivational systems would be associated with an early-developing aspect. The argument that pubertal measures should be chosen on the basis of conceptual links to outcomes of interest is not new (Dorn et al., 2006), but is often overlooked.

In particular, conceptual and practical considerations have kept researchers focused on two main measures of puberty: age at menarche and self-reported overall pubertal maturation (derived from an average rating of development on several features). On the one hand, these measures may not tap important processes; for example, menarche is not an external social signal and it occurs late in puberty after the major physical changes and hormone surge, and a summary typically includes features differently linked to adrenarche (e.g., body hair) and gonadarche (e.g., breast development in girls, voice change in boys). On the other hand, these measures do tap processes with psychological significance; for example, menarche has meaning for the girl experiencing it, and a summary score accurately reflects a youth's overall development.

Issues of conceptualization and measurement of puberty have received renewed attention recently, likely due both to methodological advances (e.g., hormone measurement, neuroimaging, growth curve modeling) and to increasing interest in the psychological significance of puberty from developmentalists and neuroscientists (Graber, 2013; Lenroot & Giedd, 2010; Negri & Susman, 2011; Trotman et al., 2013). Some work has focused on comparing different pubertal measures, particularly self report, pubertal hormones, and physical exams that describe development in terms of Tanner stages (e.g., Tanner, 1962).

Results show that different measures have different strengths, that self report may suffice in certain cases (Shirtcliff, Dahl, & Pollak, 2009), and that self report is more accurate at some stages than at others (Huang et al., 2012). But, most studies comparing methods involve selected samples, particularly children willing to undergo physical exam and hormone sampling. Further, there remains a need to examine further the value of self report, given the logistics and costs involved in physical exams and hormone assays, and the number of children who refuse invasive measures (Shirtcliff et al., 2009).

Other work has taken advantage of longitudinal data to provide estimates of timing and tempo by modeling trajectories of development. Approaches include both a simple linear model (Castellanos-Ryan, Parent, Vitaro, Tremblay, & Seguin, 2013; Mendle, Harden, Brooks-Gunn, & Graber, 2010) and a logistic (S-shaped) model. A logistic model is thought to reflect physiological changes more accurately than a linear model (Eaves et al., 2004; Huang, Biro, & Dorn, 2009; Marceau, Ram, Houts, Grimm, & Susman, 2011) and does fit better (Marceau et al., 2011), but models have not been compared with respect to links to behavior.

A benefit of modeling is the direct estimation of tempo or “how quickly or slowly [children] develop” (Mendle, 2014, p. 215). But tempo estimates from these methods only operationalize rate of pubertal development and are inconsistent with each other, and with other methods of operationalizing tempo. Studies using traditional pediatric measures of tempo (interval between two stages of puberty, e.g., time from Tanner 2 breast development to menarche) show that tempo is adjusted when onset is off time, so that tempo is slow for children who enter puberty early, and fast for children who enter late; this is best-documented in girls (Biro et al., 2006; Martí-Henneberg & Vizmanos, 1997; Pantiotou et al., 2008). Studies estimating tempo from linear or logistic models find varying links of tempo with timing and with behavior (Castellanos-Ryan et al., 2013; Marceau et al., 2011; Mendle et al., 2010). Inconsistencies likely reflect different ways of estimating tempo: Traditional measures reflect interval between pubertal events; the linear model estimates rate of change per year between two consecutive puberty stages; and the logistic model estimates peak change rate at the midpoint of puberty.

Another focus of recent work regarding measurement and meaning of puberty concerns sex differences. Pubertal timing is seen to affect development differently in boys than in girls, with early-maturing girls but both early- and late-maturing boys showing increased depression or adjustment problems compared to on-time peers (Graber, 2013; Mendle & Ferrero, 2012). The sexes may also be differently affected by pubertal hormones; for example, hormones that are higher in boys than in girls (e.g., testosterone) may be important in risk taking (Paus et al., 2010), whereas hormones that are higher than in boys (e.g., estradiol) may be important in disordered eating (Klump, Keel, Sisk, & Burt, 2010) and depression (Angold, Costello, & Worthman, 1998). The lack of a simple measure of pubertal development in boys analogous to menarche in girls means that boys are studied less often than are girls.

Thus, despite advances in understanding the nature and psychological significance of pubertal development, uncertainty remains about how self report reflects pubertal

development, and how findings depend on the model of pubertal development and youth's sex. To address these issues, we used longitudinal self-report data to model different trajectories of pubertal development and then used the timing and tempo estimates to predict outcomes of internalizing and externalizing behavior problems, and early sexual activity. We compared model estimates to traditional measures of development.

We investigated several specific questions about the measurement and modeling of pubertal development, and how different pubertal indicators relate to behavioral outcomes. First, are complex curves statistically and conceptually necessary to model pubertal change, particularly in relation to behavior? Second, what is the relative importance for behavior of pubertal timing versus tempo, and does the answer depend on how timing and tempo are conceptualized and estimated? Third, how do the answers to these questions vary for boys and girls?

Method

Participants

Participants were enrolled in two longitudinal genetically informative studies, the Colorado Longitudinal Twin Study (LTS) and the Colorado Adoption Project (CAP), which focused on genetic and environmental contributors to variations in cognition, personality, and behavior problems (Plomin & DeFries, 1985; Rhea, Gross, Haberstick, & Corley, 2006). The original LTS sample consisted of 966 individuals from 483 twin pairs and the original CAP sample consisted of 732 subjects from 490 families.

Most participants provided at least one pubertal self assessment: 89% of LTS, 91% of CAP. Those providing pubertal data were similar to those who did not on the outcome measures of interest. The current sample (total $N = 1519$) included 854 participants from LTS and 665 participants from CAP: from LTS, 234 monozygotic (MZ) and 192 dizygotic (DZ) girls, 213 MZ and 213 DZ boys, and 2 boys of unknown zygosity; from CAP, 137 adopted and 175 non-adopted girls, and 153 adopted and 200 non-adopted boys. Most participants were White (92%) and not Hispanic (95%). The samples of boys and girls were each divided into two replicates to address within-family dependencies (so that each replicate contained one member of each family) and to permit cross validation, as has been done elsewhere (e.g., Bleidorn, Kandler, Riemann, Angleitner, & Spinath, 2009). As shown below, results were generally consistent across replicates.

Participants were assessed on multiple occasions from infancy through young adulthood; the focus here is on assessments of puberty throughout adolescence and behavioral outcomes in mid-to-late adolescence/early adulthood. Puberty was assessed annually from the end of grade 3 (average age (SD in parentheses): 9.46 (.37) years, range 8.25–10.92) to the end of grade 9 (average age: 15.37 (.32) years, range 13.75–17.25), with an in-person visit after grade 6, and telephone interviews at other ages. The interviews also included assessments of various psychological and health-related characteristics. Most behavioral outcomes of interest here were assessed between ages 16 and 18, although age at sexual initiation was assessed in some participants at age 21; the interviews also included other assessments not reported here.

Measures

Pubertal development—Puberty was assessed by annual self report on the Pubertal Development Scale (PDS; Petersen, Crockett, Richards, & Boxer, 1988). Youth answered five questions about the development of secondary sexual characteristics: body hair, skin changes, and growth spurt in both sexes, facial hair and deepening voice in boys, and breast development and menarche in girls. All items except menarche were rated on a 4-point scale: 1 = “no development,” 2 = “yes, barely,” 3 = “yes, definitely,” 4 = “development completed.” Menarche was rated as absent (1) or completed (4); age at menarche was recorded for those who had reached it. Items were averaged to produce a summary PDS score at each age.

The psychometric properties of the PDS are well studied (Petersen et al., 1988). Correlations with pubertal stage rated by health professionals are generally .70 (Schmitz et al., 2004; Shirtcliff et al., 2009). PDS scores correlated with salivary hormone levels to the same extent as physical exam did (Shirtcliff et al., 2009). The PDS has been considered to be “most appropriate for broad estimates of development, or for use in longitudinal studies” (Coleman & Coleman, 2002, p. 547), although this is not without controversy (e.g., Dorn et al., 2006; Shirtcliff et al., 2009); we return to issues surrounding self report in the Discussion.

Behavior problems—Parent reports of child behavior problems at age 16 were obtained with the Child Behavior Checklist (CBCL; Achenbach, 1991). We used unstandardized scores for higher-order scales of Internalizing and Externalizing Problems, which have been shown to relate to a variety of clinical conditions.

Depression—At age 17, participants described their mood over the past week using The Center for Epidemiologic Studies at the National Institute of Mental Health Depression scale (CES-D; Radloff, 1977). This 20-item survey has very good reliability (internal consistency reliability of .85–.90) and validity (discriminating between a clinical and general population sample, and correlating with other scales of depression).

Drug symptom counts—Involvement with substances, including alcohol, cannabis, amphetamines, opiates, cocaine, sedatives, inhalants, PCP, and hallucinogens, was assessed with the Composite International Diagnostic Interview-Substance Abuse Module (CIDI-SAM; Cottler & Keating, 1990) at ages 16–18. For each substance, participants retrospectively recalled whether they had any of seven dependence symptoms during assessments between ages 16 and 18. This measure has discriminative and convergent validity, as determined in adolescents with substance use and conduct problems (Crowley, Mikulich, Ehlers, Whitmore, & MacDonald, 2001). An overall measure of substance involvement was obtained from the average lifetime number of symptoms experienced during adolescence across all substances (Button, Hewitt, Rhee, Corley, & Stallings, 2010; Stallings et al., 2003).

Conduct disorder—Lifetime symptoms of conduct disorder (CD) were assessed at age 17 with the Diagnostic Interview Schedule for Children (DISC-IV; Shaffer, Fisher, Lucas,

Dulcan, & Schwab-Stone, 2000). Symptoms range from truancy and disobeying parental curfews to initiating fights and committing crimes (e.g., breaking into a car). CD symptom severity was measured as the number of endorsed symptoms, with scores standardized by participants' sex and age. Substance use and conduct disorder measures were normed with regard to age and sex on adolescents from several samples (Rhea et al., 2006; Stallings et al., 2003).

Age at first sex—Sexual history was assessed at age 17 in LTS participants, and at age 17 or 21 in CAP participants. Although methods varied across age, all participants provided information on the age of their first sexual experience if it had occurred (Bricker et al., 2006).

Analysis Plan

The overall analytic strategy was to estimate timing and tempo of pubertal development from longitudinal growth curve models and traditional measures, and to relate estimates to each other and to psychological outcomes. Analyses were conducted separately for boys and girls.

Linear and logistic modeling of pubertal development—The longitudinal data (average PDS score at each of seven waves of assessment) were used to calculate group trajectories of development separately for girls and boys, allowing individual deviations, and providing estimates of pubertal timing (PDS score of 2.5, corresponding to Tanner stage 3) and tempo. Development was represented and estimated separately by linear and logistic models because the models incorporate different assumptions. A linear model represents constant development, whereas a logistic model represents development that is symmetric but not uniform across time, starting slowly, increasing to maximum growth at the midpoint, and then decreasing in rate.

Models were compared statistically with AIC and BIC; lower values indicate better fit. They were evaluated conceptually by consistency of timing and tempo parameters across models, and links between parameters and behavioral outcomes. A linear model is represented as

$$PDS = g_{0i} + g_{1i} * Age_{it} + r_{it},$$

where g_{0i} is the intercept, g_{1i} is the slope, and r_{it} is the residual for an individual i at assessment t , having a normal distribution with 0 mean and constant variance at each t . Timing at a specific stage of development was calculated algebraically; for example, age at PDS 2.5 = $(2.5 - g_{0i}) / g_{1i}$. The dashed lines and text in Figure 1A depict how the linear timing and tempo parameters were calculated (overlaid on the mean results for girls).

A logistic model is represented as

$$PDS = ((\beta_0 + (\beta_1 - \beta_0)) / (1 + e^{-\alpha_i(Age_{it} - \lambda_i)})) + r_{it},$$

where β_0 is 1, the lower bound for PDS scores; β_1 is 4, the upper bound for PDS scores; e is the exponential function; λ_i is the age at PDS 2.5; α_i is the slope of the function at the mid-pubertal age; as above, r_{it} is the normally-distributed residual for an individual i at assessment t . Logistic estimates of timing at mid puberty and maximum tempo coincide, so the slope of the curve is maximized at the mid-pubertal age and is the first derivative of this function with respect to Age. The dashed lines and text in Figure 1B depict how the logistic timing and tempo parameters were calculated (overlaid on the mean results for girls).

Parameter estimation—Linear and logistic mixed-effects curves were estimated separately for boys and girls using SAS 9.3 PROC NLMIXED with the full information maximum likelihood method to handle missing data (SAS Institute, Cary, NC; see Grimm & Ram, 2009). The equations above represent Level 1 models for fixed group-level effects. Individual-level random effects were defined and estimated in the following Level 2 models. For the linear model, $g_{0i} = \gamma_{00} + v_{0i}$ and $g_{1i} = \gamma_{10} + v_{1i}$, where γ_{00} is the mean intercept; γ_{10} is the mean slope; v_{0i} and v_{1i} are individual deviations from the respective means. For the logistic model, $\lambda_i = \delta_{00} + \varepsilon_{0i}$ and $\alpha_i = \delta_{10} + \varepsilon_{1i}$, where δ_{00} is the mean midpoint; δ_{10} is the mean slope at the midpoint; ε_{0i} and ε_{1i} are individual deviations from the respective means. For both models, the deviation parameters have multivariate normal distributions with 0 means, variances, and covariances. The random effects parameters were used in subsequent analyses.

It is important to emphasize the assumptions of the models that affect parameter estimates or limit their use. The linear model has a constant slope, so tempo is estimated to be uniform across development. The logistic model is centered and symmetrical at mid puberty, so timing is estimated at mid puberty (PDS 2.5), and tempo is estimated as instantaneous rate of change at that point, representing the maximum speed of development. Neither model permits *independent* estimation of pubertal onset (PDS 1.5) and mid puberty (PDS 2.5).

Traditional measures of pubertal development in girls—Timing was assessed by age at menarche (assessed close in time to the event). Tempo was assessed by the difference between age at pubertal onset (PDS 1.5 estimated from linear models) and menarche. Thus, tempo is conceptualized differently by the growth curve models and the traditional approach: model parameters reflect rate of change, whereas the traditional measure reflects time between pubertal events; they are thus inversely related.

Comparing estimates of pubertal timing and tempo—Estimates of pubertal timing and tempo were compared by testing mean differences with matched-pairs t -tests, and by correlating estimates across method. Links between timing and tempo within method were examined with correlations.

Links between puberty and psychological function—Psychological outcome measures were correlated with estimates of pubertal development, including linear and logistic estimates of pubertal timing and tempo for both sexes, and traditional measures for girls. Analyses in girls included only those who had data on both age at menarche and estimated trajectories (N=613).

Missing data—Complete data were not available for all participants, either because of missing puberty or behavioral data. PDS data were incomplete because of missed assessments or skipped items at some assessments. Most youth had six or seven assessments (76.7% of girls, 74.4% of boys), with small numbers having fewer assessments (three or four assessments, 9.5% of girls and 9.2% of boys; two assessments, 3.0% of girls and 3.7% of boys). Most missing assessments were at the last wave for youth close to age 16, and therefore administered the age 16 assessment, which did not include the PDS. Participants were included in the linear and logistic trajectory analyses if they had at least two different PDS scores; this excluded 3.9% of girls and 3.8% of boys. The percentage of individuals with incomplete PDS data (missing one or more items) at a given assessment was low: for boys, it was less than 1% at all ages; for girls, it was less than 4% at all except the final assessment when it was 6%. Participants were included if they reported on at least four features. Some girls did not have information on age at menarche ($N = 116$, 15.7%). In most cases, this was because they had not yet reached menarche. Behavioral data were missing primarily because of the assessment schedule. Amount of missing behavioral data was not significantly associated with pubertal development (with a representative measure of pubertal timing, for girls, $r = .04$, $p > .05$; for boys, $r = -.01$, $p > .05$). Further, participants were excluded from correlations (i.e., timing, tempo, and behavior) in which their puberty parameters were outliers, i.e., three or more standard deviations from the mean or $\text{tempo} < 10$ (which are implausible). Across all models, 1.1% of girls were outliers and 6.5% of boys were outliers, most due to very slow development estimated from the linear model.

Results

As expected, there was considerable variability, including sex differences, in both pubertal development and in psychological outcomes. Expected sex differences were seen on most outcome measures, as shown in Table 1: Boys scored higher than girls on CBCL externalizing problems and on conduct disorder symptoms, and girls scored higher than boys on CBCL internalizing problems and depressed mood; there were no significant sex differences on drug symptom counts or age at sexual initiation.

Trajectory Results: Model Fit

Both linear and logistic models were successfully fit to the data. Group-level mean curves for each model type by sex are shown in Figure 1. All estimated curves for replicate 1 (black) and replicate 2 (gray) overlapped, reflecting nearly identical mean group estimates. The logistic model fit better than the linear model, for both replicates for both sexes, as shown in Table 2.

Consistency of Estimates of Timing and Tempo

In general, the methods produced similar estimates of pubertal timing, but divergent estimates of pubertal tempo. Results were consistent across replicates.

Individual differences in pubertal timing were estimated similarly across methods, as shown in Table 3: linear and logistic parameters were highly correlated for both sexes, and age at menarche correlated well and to a similar degree with those estimates. Girls were estimated

to reach mid puberty earlier than boys, as expected. Timing was estimated to be later in the linear versus logistic model for both sexes: Pairwise differences were significant in matched pairs *t*-tests; *d*'s (adjusted for correlated measures) were small to moderate for girls, and large for boys.

Individual differences in tempo were not estimated consistently across method, as shown in Table 4: parameters were weakly correlated for girls, and moderately correlated for boys. (Tempo is opposite in direction for traditional versus linear and logistic models, reflecting time for the former, and rate for the latter.) Estimated mean tempo varied across method, although all showed girls progressing through puberty more quickly than boys.

Timing-Tempo Links

Links between timing and tempo differed for the logistic model compared to other methods, as shown in Table 5. (Note again opposing direction of tempo for traditional versus linear and logistic models.) Early timing was associated with fast linear tempo for both sexes and traditional tempo for girls. But, logistic timing and tempo correlated in opposite directions for boys and girls, with the pattern for boys similar to that of the other methods (early timing associated with fast tempo) and the opposite pattern for girls (early timing associated with slow tempo). Results were consistent across replicates. The varying tempo estimates reflect model assumptions (i.e., uniform tempo in the linear model, instantaneous tempo at PDS 2.5 in the logistic model, and the need to estimate PDS 1.5 to obtain the traditional measure of tempo).

Links between Pubertal Development and Psychological Development

Pubertal timing correlated better with some psychological outcomes (i.e., drug symptom counts, age at first sex) than others, with links more likely to be significant in girls (Figure 2) than in boys (Figure 3). Figures show correlations with timing estimated from the different methods, with areas conveying differential predictions. Points on the abscissa do not lie on a continuum, but peaks and valleys reflect distinct correlations. Behavior was linked in similar ways to all estimates of pubertal timing: early puberty was associated with problems in both sexes.

Pubertal tempo was not clearly correlated with psychological outcomes, as shown in Table 6. There was some suggestion that fast tempo was associated with externalizing problems in both sexes. This was particularly the case for the association of linear and traditional tempo measures with age at first sex.

Correlations between outcome measures and pubertal timing and tempo were consistent across replicates. No correlations significantly differed between replicates. Nevertheless, given the small size of the correlations, some were significantly different from zero in one replicate but not the other.

Discussion

This study was focused on examining different models of pubertal development and their links with psychological problems in mid-to-late adolescence. Results confirm the utility of

self report in both sexes (Shirtcliff et al., 2009), and highlight the ways in which different methods facilitate and constrain our ability to study the psychological significance of puberty. Confidence in findings is increased by consistency of results across replicates here and with other studies (Castellanos-Ryan et al., 2013; Marceau et al., 2011; Mendle et al., 2010).

Modeling Pubertal Development and Significance for Psychological Outcomes

We focused on three questions regarding the measurement and meaning of pubertal development. The answers should facilitate study of mechanisms underlying adolescent psychological change at puberty. We consider how the results bear on those questions.

Are complex curves necessary to model pubertal change?—A logistic model is better than a linear model to describe average pubertal development, as seen in data from self report (here) and physical exam by a trained health professional (Marceau et al., 2011). This likely reflects the constraint of nonlinear symmetry in the logistic model, which is biologically reasonable and facilitates description of development for youth who have not completed puberty. The superiority of the logistic model was not due to our intensive (seven-wave) assessment, because not all participants completed development, especially boys.

Estimates of timing of mid puberty from the two models were highly correlated with each other for both sexes, and with menarche in girls. Not surprisingly, then, psychological outcomes were not differentially predicted by the methods. Although all methods preserved individuals' relative ranks, the linear model predicted later pubertal timing than did the logistic model. This reflects the application of the linear model to data from youth who have not completed development; because timing calculations include tempo, they are estimated to have slow tempo and exaggerated timing (algebraic derivation of the linear equation). A presumed advantage of the logistic model over the linear model – more between-individual variation in tempo – did not translate into stronger links with behavior.

What is the relative importance for behavior of pubertal timing versus tempo?

—Results confirm others in showing that there is a small risk associated with early and fast puberty. Early puberty is weakly linked to a host of behaviors, including increased risk for internalizing problems in girls only and externalizing problems, such as conduct disorder and early sex in both sexes. These findings are generally consistent with the early timing (or stage termination) hypothesis: early maturing individuals – particularly girls – are at the greatest risk for behavior problems because they experience the greatest asynchrony of all youth between physical and cognitive development (Ge, Brody, Conger, Simons, & Murry, 2002; Negri & Susman, 2011). Fast tempo appears to be linked to externalizing problems here and elsewhere (Castellanos-Ryan et al., 2013; Marceau et al., 2011). This is consistent with the maturation compression hypothesis: quick development exacerbates the challenges that accompany the pubertal transition, and as a consequence, is associated with more problems than is slow development (Mendle, 2014; Mendle et al., 2010).

Findings from this study combine with others to emphasize the varied ways in which tempo is conceptualized and measured, and how estimates depend on model assumptions and

available data. Logistic tempo reflects instantaneous development at the midpoint of puberty, not early in the process. Linear tempo is averaged across development, so may be inaccurate if the full range of development is not measured. Traditional tempo is calculated as time between the onset of two key features known to bracket development in girls; there is no parallel in boys. Moreover, logistic and linear measures operationalize tempo as the rate of development, with estimates determined not only by a participant's data, but also by information from group means (in mixed effects models) and model parameters (especially in logistic curves). But the traditional measure operationalizes tempo as the length of time between pubertal events, so it is susceptible to methodological issues affecting two-time-point measures (e.g., regression to the mean). Nonetheless, traditional tempo has an intuitive interpretation, and it may be sufficient if pubertal change is constant across time (an assumption also made in the linear model). Results from this study provide some evidence for this: Traditional tempo was significantly related to linear (but not logistic) tempo, and traditional and linear tempo were related to externalizing behavior in similar ways.

Correlations between timing and tempo vary with the model used, explaining previous inconsistencies. Linear timing and tempo were negatively correlated (here and Mendle et al., 2010); this may reflect the constraints of a random effects growth curve model applied to large samples. Logistic timing and tempo were correlated in different ways by sex: correlations were weakly positive (early timing, slow tempo) for girls, and moderately negative (early timing, fast tempo) for boys (here and Marceau et al., 2011). The logistic association in girls is also consistent with studies using the traditional measure (Biro et al., 2006; Martí-Henneberg & Vizmanos, 1997; Pantisiotou et al., 2008). Across methods and studies, fast tempo is associated with externalizing, but not internalizing, problems in both sexes (here and Castellanos-Ryan et al., 2013; Marceau et al., 2011).

How do the answers to the questions about pubertal development vary for boys and girls?—For both sexes, trajectories could be obtained, and the average logistic model fit better than the average linear model. There were sex differences in timing and tempo and in their association with each other and with psychological outcomes. Compared to girls, boys reached mid puberty later and progressed through puberty more slowly. Timing-tempo correlations for the logistic model were positive for girls (early maturers develop slowly, consistent with the endocrine literature; Biro et al., 2006; Martí-Henneberg & Vizmanos, 1997; Pantisiotou et al., 2008), but negative for boys (early maturers develop quickly, consistent with another study using a logistic model; Marceau et al., 2011). But, these findings need to be considered in light of the varying estimates of tempo provided by the different methods and the number of boys who had not completed development.

Pubertal timing was related to behavior in girls more than in boys, consistent with previous work. Given hypotheses about sex differences in the type of off-time development that matters (Graber, 2013) – with problems associated with early timing in girls and boys, and late timing in boys – it is necessary to assess pubertal development in boys beyond the ages of this study. Our results show that it is possible to obtain good estimates of pubertal development in boys from PDS data, encouraging further study of psychological consequences of their off-time development.

Strengths and Limitations

The study extends other work in several notable ways. We explicitly tested and compared multiple methods for estimating pubertal timing and tempo, and their links to behavior. Using self-report data, we confirmed that a logistic model fit the data better than a linear model (Marceau et al., 2011), and extended previous work to show that different estimates of pubertal timing relate to behavior in similar ways, and that estimates of pubertal tempo vary by method (Marceau et al., 2011; Mendle et al., 2010). We studied both sexes, showing that self-report longitudinal data can be used to examine the psychological significance of puberty in boys as well as girls.

The study had several methodological strengths. Yearly assessments were available, with fewer missing data than in other studies (Marceau et al., 2011; Mendle et al., 2010; Paus et al., 2010). The sample was large enough to confirm results in two replicates and to see relatively small effects. Multiple psychological domains were assessed and at periods not contemporaneous with puberty assessments, ensuring that links reflected pubertal timing and not pubertal status.

Nevertheless, there are several limitations that need to be considered in interpreting the results. Puberty was assessed with self report rather than physical exam by a health professional. Although there are concerns about self report (e.g., Huang et al., 2012), it remains the method of choice by most investigators because of ease of use, cost, and non intrusive nature. It is therefore important to emphasize that PDS data generated good trajectories for both sexes, and parameter estimates correlated with behavior in ways that were expected from other studies using physical exams. Further, physical exam is not useful for most behavioral studies because it is difficult to obtain those data; for example, in one national study, only about 70% of children had valid pubertal measurements on at least one of seven annual assessments, and girls had more valid data than boys at all ages (Susman et al., 2010).

The assessment schedule meant that complete development was not captured for all children, especially boys. Other studies had similar limitations (e.g., Marceau et al., 2011). This creates particular challenges for estimating tempo with a linear model, and means that correlations between pubertal timing and behavior primarily reflect effects of early puberty, because it was difficult to differentiate on-time from late puberty in this sample.

There is more confidence in our results for girls than for boys. This is primarily due to lower rates of completed development for boys than for girls. But, the sample of boys was large enough to detect correlations of the size observed in girls.

The sample is selected in several ways, consisting of twins and adoptees and their controls of middle social class and with limited representation of minority youth. It is unclear how our results can be generalized to individuals of other races, ethnicities, or social classes. Further, the relative privilege of the sample may relate to the relatively late ages of pubertal development.

Finally, there are some limitations to our behavioral measures, although all are widely-used and psychometrically sound. Of particular concern, the CES-D assesses depression only over the past week; CIDI-SAM relies upon recall of substance use; measures of undesirable behavior (CIDI-SAM, CD) may be subject to underreporting; individuals may not have completely passed through the risk period for some behaviors. In general, however, results are consistent with others linking puberty to psychological risk. Further, a key aspect of the study was to assess outcome in mid-to-late adolescence, in order to examine longer-term effects of pubertal development, and to separate pubertal timing from status.

Opportunities and Challenges in Measuring and Modeling Pubertal Development

Our results combine with others in demonstrating that pubertal timing – however it is measured – relates to psychological problems in girls. On the one hand, this suggests that simple measures, such as recalled age at menarche, might suffice to address some questions about consequences of variations in pubertal development. On the other hand, even complex models such as those used here are not sufficient to address questions about mechanisms linking pubertal timing to problems or understanding pubertal influences on normative development. This is seen in several ways.

First, neither linear nor logistic models enable *separate* estimation of timing of different stages of pubertal development (e.g., onset versus mid puberty). But this information is crucial for understanding how puberty matters for different aspects of psychological outcome. On the one hand, mid puberty is a time of risk because of hormone increases; for example, a surge in estradiol is thought to trigger depression and disordered eating in vulnerable girls (Angold et al., 1998; Klump et al., 2010). On the other hand, pubertal onset is a time of risk, particularly for girls who mature early and whose early physical changes elicit responses from the social environment (Ge et al., 2002; Ge, Conger, & Elder, 1996).

Second, there is no consensus on the measurement – or even the meaning – of pubertal tempo. Most hypotheses about psychological consequences of variations in tempo focus on a slice of development (e.g., whether adverse consequences of early breast development are mitigated by slowed subsequent development). But current methods do not provide corresponding measures: linear models assume constant tempo across development, whereas logistic models assess tempo that is instantaneous at midpuberty, and traditional methods focus on linear change between an early event and a late event.

Third, despite progress in recognizing the potential differential significance of adrenarche versus gonadarche, existing measures do not adequately separate those processes, and modeling cannot improve inadequate measures. Data from several studies using physical features of puberty (obtained from self report or exams by health professionals) show adrenarche to occur later than gonadarche (Marceau et al., 2011; Paus et al., 2010; Shirtcliff et al., 2009), see Footnote 1. But adrenal hormones rise earlier than gonadal hormones (reviewed in Styne & Grumbach, 2011), suggesting that physical features alone are

¹We also attempted to model adrenarche and gonadarche separately, using a conversion of PDS items (Shirtcliff et al., 2009). Results showed timing of adrenarche to occur later than gonadarche, consistent with results using physical features, but not the endocrine literature (reviewed in Styne & Grumbach, 2011).

insufficient to separate adrenarche from gonadarche, perhaps because adrenal hormone levels are insufficient to produce physical changes until gonadarche (Dorn et al., 2006; Wan, Deng, Archer, & Sun, 2012).

Fourth, issues in describing pubertal development are not restricted to self report data. Direct measurement of adrenal and gonadal hormones, particularly at a single point in time, does not guarantee better assessment of pubertal processes than does measurement of physical features (even by self report; Shirtcliff et al., 2009). Individual differences in hormone levels reflect more than pubertal development: Hormone levels are moderately heritable (Harris, Vernon, & Boomsma, 1998), but fluctuate across the day (and menstrual cycle for girls, and season for boys), and with diet, exercise, and behavior itself (Carré, 2009; Stanton, Mullette-Gillman, & Huettel, 2011); responses to hormones also depend on other hormones that are present, and sensitivity of hormone receptors (Styne & Grumbach, 2011). There are also concerns about the validity of salivary assays (van Anders, 2010), and methods for measuring sex steroid hormones (Handelsman & Wartofsky, 2013).

Fifth, all mixed-effects growth curve models (as used in studies of pubertal development) describe individual development with reference to the group mean. But, there may be important individual differences in the shape of development. It might be fruitful to pursue this topic using other analytic approaches (e.g., latent profile analysis).

Finally, model fit and parameter estimates depend on the number and timing of assessments. For example, logistic models require less information to obtain estimates of timing and tempo for individuals than do linear models. Furthermore, estimates of tempo are more straightforward in linear than in logistic models, but this advantage is offset by the need for data at more times, particularly pubertal onset or end, and midpuberty.

Conclusions

Self reports of pubertal development provide meaningful data for both boys and girls; PDS data produced good trajectories of development, and estimates of individuals' pubertal timing and tempo were obtained even with missing assessments. Pubertal development is described better at the group level by a logistic than by a linear model. The timing of midpuberty is estimated to occur earlier in a logistic than in a linear model, but individual differences in pubertal timing are estimated in similar ways by the models, and by a traditional measure (age at menarche), and all estimates correlate in similar ways with internalizing and externalizing problems measured in mid-to-late adolescence/early adulthood. The tempo of puberty is not estimated consistently across linear and logistic models and a traditional approach, because estimates depend on model assumptions and available data. Associations of tempo with timing are not method invariant, limiting comparison across studies. Advances in modeling have facilitated understanding of the nature of pubertal development, but many questions about psychological development in adolescence cannot easily be studied with current models.

Methodological advances in describing and measuring puberty are essential for developmental science. As we have noted elsewhere in the paper, pubertal development is important for psychological development in adolescence and beyond, serving as a source of

psychological risk, as a contributor to normative change, and as a trigger for psychopathology in vulnerable youth (see, e.g., Ge & Natsuaki, 2009; Graber, 2013; Gunnar et al., 2009; Negri & Susman, 2011; Trotman et al., 2013). In fact, much contemporary work is focused on understanding the ways in which pubertal changes are responsible for the neural and behavioral changes of adolescence, as seen in several recent journal special issues (Engle, 2013; Luciana, 2010; Sisk & Berenbaum, 2013). This work depends on defining the aspects of puberty that matter and measuring them well, and thus methodological advances are essential to understanding how and why puberty matters for children's psychological development.

Acknowledgments

The research reported here was supported by grants from the National Institutes of Health, HD010333, HD036773, and DA011015. We thank Sally-Ann Rhea for overseeing the data collection and management, Michael Stallings for his input on assessment of psychological outcomes, and Kelly Klump for helpful comments on an earlier version of the manuscript.

References

- Achenbach, T.M. Manual for the Child Behavior Checklist: 4–18 and 1991 profile. Burlington, VT: University of Vermont Department of Psychiatry; 1991.
- Angold A, Costello EJ, Worthman CM. Puberty and depression: The roles of age, pubertal status and pubertal timing. *Psychological Medicine*. 1998; 28(1):51–61. [PubMed: 9483683]
- Biro FM, Huang B, Crawford PB, Lucky AW, Streigel-Moore R, Barton BA, et al. Pubertal correlates in Black and White girls. *Journal of Pediatrics*. 2006; 148(2):234–240. [PubMed: 16492435]
- Bleidorn W, Kandler C, Riemann R, Angleitner A, Spinath FM. Patterns and sources of adult personality development: Growth curve analyses of the NEO PI-R scales in a longitudinal twin study. *Journal of Personality and Social Psychology*. 2009; 97(1):142–155. [PubMed: 19586245]
- Bricker JB, Stallings MC, Corley RP, Wadsworth SJ, Bryan A, Timberlake DS, et al. Genetic and environmental influences on age at sexual initiation in the Colorado Adoption Project. *Behavior Genetics*. 2006; 36:820–832. [PubMed: 16710776]
- Button TMM, Hewitt JK, Rhee SH, Corley RP, Stallings MC. The moderating effect of religiosity on the genetic variance of problem alcohol use. *Alcoholism-Clinical and Experimental Research*. 2010; 34:1619–1624.
- Carré JM. No place like home: Testosterone responses to victory depend on game location. *American Journal of Human Biology*. 2009; 21(3):392–394. [PubMed: 19127526]
- Castellanos-Ryan N, Parent S, Vitaro F, Tremblay RE, Seguin JR. Pubertal development, personality, and substance use: A 10-year longitudinal study from childhood to adolescence. *Journal of Abnormal Psychology*. 2013; 122(3):782–796. [PubMed: 24016016]
- Coleman L, Coleman J. The measurement of puberty: A review. *Journal of Adolescence*. 2002; 25:535–550. [PubMed: 12234559]
- Cottler LB, Keating SK. Operationalization of alcohol and drug dependence criteria by means of a structured interview. *Recent Developments in Alcoholism*. 1990; 8:69–83. [PubMed: 2333396]
- Crowley TJ, Mikulich SK, Ehlers KM, Whitmore EA, MacDonald MJ. Validity of structured clinical evaluations in adolescents with conduct and substance problems. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2001; 40(3):265–273. [PubMed: 11288767]
- Dorn LD, Dahl RE, 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(1):30–56.
- Eaves L, Silberg J, Foley D, Bulik C, Maes H, Erkanli A, et al. Genetic and environmental influences on the relative timing of pubertal change. *Twin Research*. 2004; 7(5):471–481. [PubMed: 15527663]

- Engle RW. Introduction to special issue on the teenage brain. *Current Directions in Psychological Science*. 2013; 22(2):79.
- Ge X, Brody GH, Conger RD, Simons RL, Murry VM. Contextual amplification of pubertal transition effects on deviant peer affiliation and externalizing behavior among African American children. *Developmental Psychology*. 2002; 38:42–54. [PubMed: 11806701]
- Ge X, Conger RD, Elder GH. Coming of age too early: Pubertal influences on girls' vulnerability to psychological distress. *Child Development*. 1996; 67:386–400.
- Ge XJ, Natsuaki MN. In search of explanations for early pubertal timing effects on developmental psychopathology. *Current Directions in Psychological Science*. 2009; 18(6):327–331.
- Graber JA. Pubertal timing and the development of psychopathology in adolescence and beyond. *Hormones and Behavior*. 2013; 64(2):262–269. [PubMed: 23998670]
- Grimm KJ, Ram N. Nonlinear growth models in Mplus and SAS. *Structural Equation Modeling: A Multidisciplinary Journal*. 2009; 16(4):676–701. [PubMed: 23882134]
- Gunnar MR, Wewerka S, Frenn K, Long JD, Griggs C. Developmental changes in hypothalamus-pituitary-adrenal activity over the transition to adolescence: Normative changes and associations with puberty. *Development and Psychopathology*. 2009; 21(1):69–85. [PubMed: 19144223]
- Handelsman DJ, Wartofsky L. Requirement for mass spectrometry sex steroid assays in the *Journal of Clinical Endocrinology and Metabolism*. *Journal of Clinical Endocrinology and Metabolism*. 2013; 98(10):3971–3973. [PubMed: 24098015]
- Harris JA, Vernon PA, Boomsma DI. The heritability of testosterone: A study of Dutch adolescent twins and their parents. *Behavior Genetics*. 1998; 28(3):165–171. [PubMed: 9670592]
- Huang B, Biro FM, Dorn LD. Determination of relative timing of pubertal maturation through ordinal logistic modeling: Evaluation of growth and timing parameters. *Journal of Adolescent Health*. 2009; 45(4):383–388. [PubMed: 19766943]
- Huang B, Hillman J, Biro FM, Ding L, Dorn LD, Susman EJ. Correspondence between gonadal steroid hormone concentrations and secondary sexual characteristics assessed by clinicians, adolescents, and parents. *Journal of Research on Adolescence*. 2012; 22(2):381–391. [PubMed: 23204809]
- Clump KL, Keel PK, Sisk C, Burt SA. Preliminary evidence that estradiol moderates genetic influences on disordered eating attitudes and behaviors during puberty. *Psychological Medicine*. 2010; 40(10):1745–1753. [PubMed: 20059800]
- Lenroot RK, Giedd JN. Sex differences in the adolescent brain. *Brain and Cognition*. 2010; 72(1):46–55. [PubMed: 19913969]
- Luciana M. Adolescent brain development: Current themes and future directions: Introduction to the special issue. *Brain and Cognition*. 2010; 72(1):1–5. [PubMed: 20006416]
- 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; 47(5):1389–1409. [PubMed: 21639623]
- Martí-Henneberg C, Vizmanos B. The duration of puberty in girls is related to the timing of its onset. *Journal of Pediatrics*. 1997; 131(4):618–621. [PubMed: 9386670]
- Mendle J. Beyond pubertal timing: New directions for studying individual differences in development. *Current Directions in Psychological Science*. 2014; 23(3):215–219.
- Mendle J, Ferrero J. Detrimental psychological outcomes associated with pubertal timing in adolescent boys. *Developmental Review*. 2012; 32(1):49–66.
- 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(5):1341–1353. [PubMed: 20822243]
- 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(3):717–746.
- Pantsiotou S, Papadimitriou A, Douros K, Priftis K, Nicolaidou P, Fretzayas A. Maturational tempo differences in relation to the timing of the onset of puberty in girls. *Acta Paediatrica*. 2008; 97(2): 217–220. [PubMed: 18177444]

- Paus T, Nawaz-Khan I, Leonard G, Perron M, Pike GB, Pitiot A, et al. Sexual dimorphism in the adolescent brain: Role of testosterone and androgen receptor in global and local volumes of grey and white matter. *Hormones and Behavior*. 2010; 57(1):63–75. [PubMed: 19703457]
- 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. [PubMed: 24277579]
- Plomin, R.; DeFries, JC. *Origins of individual differences in infancy: The Colorado Adoption Project*. New York: Academic Press; 1985.
- Radloff L. The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*. 1977; 7:385–401.
- Rhea S-A, Gross AA, Haberstick BC, Corley RP. Colorado Twin Registry. *Twin Research and Human Genetics*. 2006; 9:941–949. [PubMed: 17254434]
- Schmitz KE, Hovell MF, Nichols JF, Irvin VL, Keating K, Simon GM, et al. A validation study of early adolescents' pubertal self-assessments. *Journal of Early Adolescence*. 2004; 24(4):357–384.
- Shaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME. NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2000; 39(1):28–38. [PubMed: 10638065]
- Shirtcliff EA, Dahl RE, Pollak SD. Pubertal development: Correspondence between hormonal and physical development. *Child Development*. 2009; 80(2):327–337. [PubMed: 19466995]
- Sisk CL, Berenbaum SA. Editorial for the special issue of hormones and behavior on puberty and adolescence. *Hormones and Behavior*. 2013; 64(2):173–174. [PubMed: 23896205]
- Stallings MC, Corley RP, Hewitt JK, Krauter KS, Lessem JM, Mikulich SK, et al. A genome-wide search for quantitative trait loci influencing substance dependence vulnerability in adolescence. *Drug and Alcohol Dependence*. 2003; 70:295–307. [PubMed: 12757967]
- Stanton SJ, Mullette-Gillman ODA, Huettel SA. Seasonal variation of salivary testosterone in men, normally cycling women, and women using hormonal contraceptives. *Physiology & Behavior*. 2011; 104(5):804–808. [PubMed: 21802437]
- Styne, DM.; Grumbach, MM. Puberty: Ontogeny, neuroendocrinology, physiology, and disorders. In: Kronenberg, HM.; Shlomo, M.; Polonsky, KS.; Larsen, PR., editors. *Williams textbook of endocrinology*. 12th ed.. Philadelphia, PA: Saunders Elsevier; 2011.
- Susman, EJ.; Dorn, LD. Puberty: Its role in development. In: Lerner, RM.; Steinberg, L., editors. *Handbook of adolescent psychology*. 3rd ed.. Hoboken, NJ: Wiley; 2009. p. 116-151.
- Susman EJ, Houts RM, Steinberg L, Belsky J, Cauffman E, DeHart G, et al. Longitudinal development of secondary sexual characteristics in girls and boys between ages 9 1/2 and 15 1/2 years. *Archives of Pediatrics & Adolescent Medicine*. 2010; 164(2):166–173. [PubMed: 20124146]
- Tanner, JM. *Growth at adolescence*. 2nd ed.. Oxford: Blackwell Scientific; 1962.
- Trotman HD, Holtzman CW, Ryan AT, Shapiro DI, MacDonald AN, Goulding SM, et al. The development of psychotic disorders in adolescence: A potential role for hormones. *Hormones and Behavior*. 2013; 64:411–419. [PubMed: 23998682]
- van Anders SM. Chewing gum has large effects on salivary testosterone, estradiol, and secretory immunoglobulin A assays in women and men. *Psychoneuroendocrinology*. 2010; 35(2):305–309. [PubMed: 19615825]
- Wan W, Deng X, Archer KJ, Sun SS. Pubertal pathways and the relationship to anthropometric changes in childhood: The Fels longitudinal study. *Open Journal of Pediatrics*. 2012; 2:118–126.

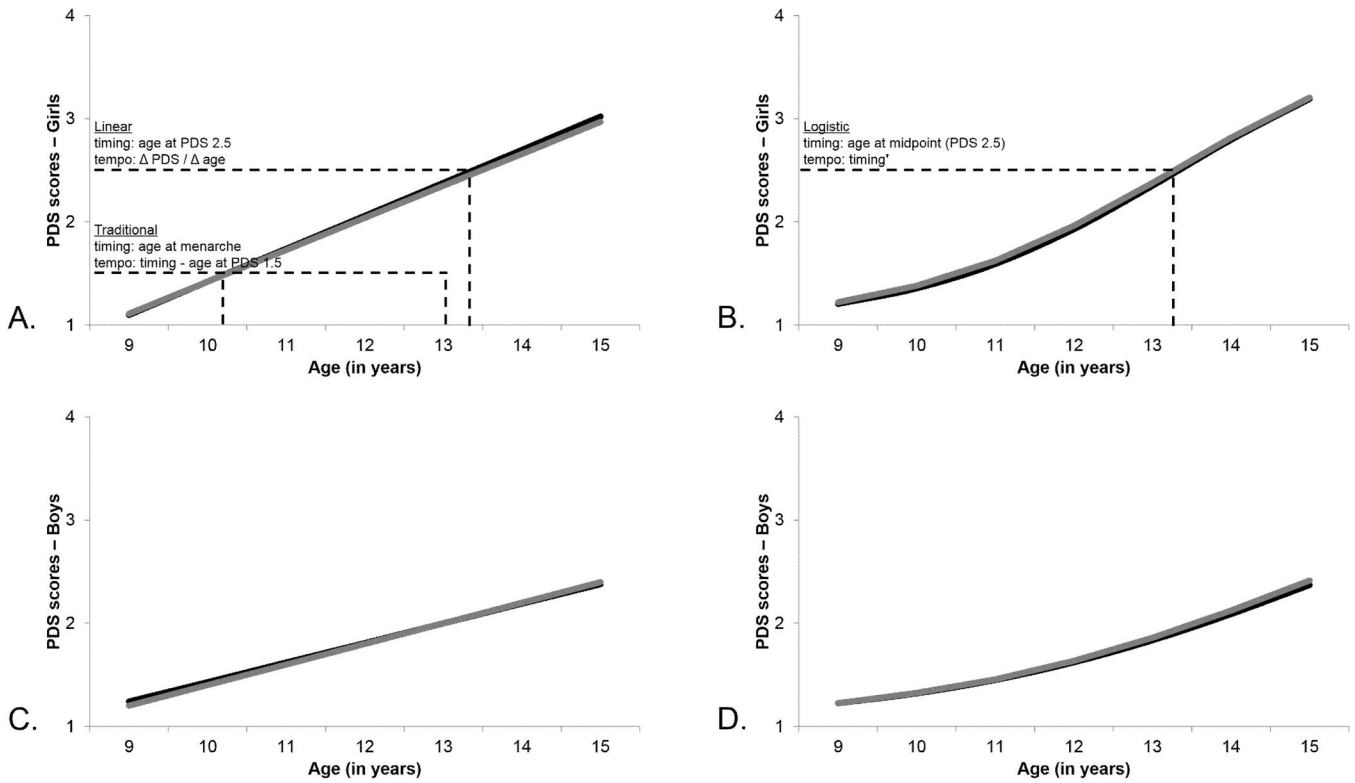


Figure 1.

Mean pubertal development trajectories (according to PDS scores) for girls and boys, with descriptions of how timing and tempo parameters were estimated; black lines are replicate 1, and gray lines are replicate 2. Figure 1A shows linear results for girls, providing a description of how linear timing and tempo parameters were estimated for both sexes; it also provides a description of how traditional timing and tempo parameters were estimated for girls. Figure 1B shows logistic results for girls, providing a description of how logistic timing and tempo parameters were estimated for both sexes, where timing' (timing prime) is the first derivative of timing. Figure 1C shows linear results for boys. Figure 1D shows logistic results for boys.

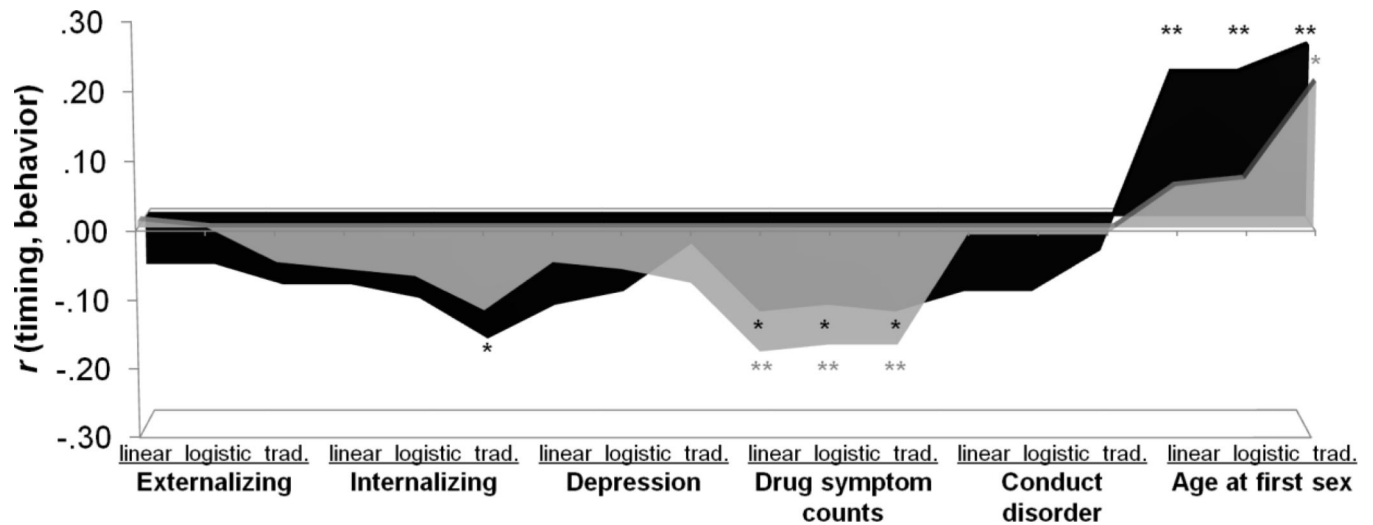


Figure 2.

Area graph of pubertal timing correlations with behavior for girls by method of assessment/point in puberty (linear model mid puberty, logistic model mid puberty, age at menarche), for each replicate. The black plot and asterisks are replicate 1 (range of N 's: 174–282); the gray plot and asterisks are replicate 2 (range of N 's: 153–277). Each peak reflects a discrete timing-behavior correlation; correlations significantly different from 0, * $p < .05$, ** $p < .01$; trad.: traditional.

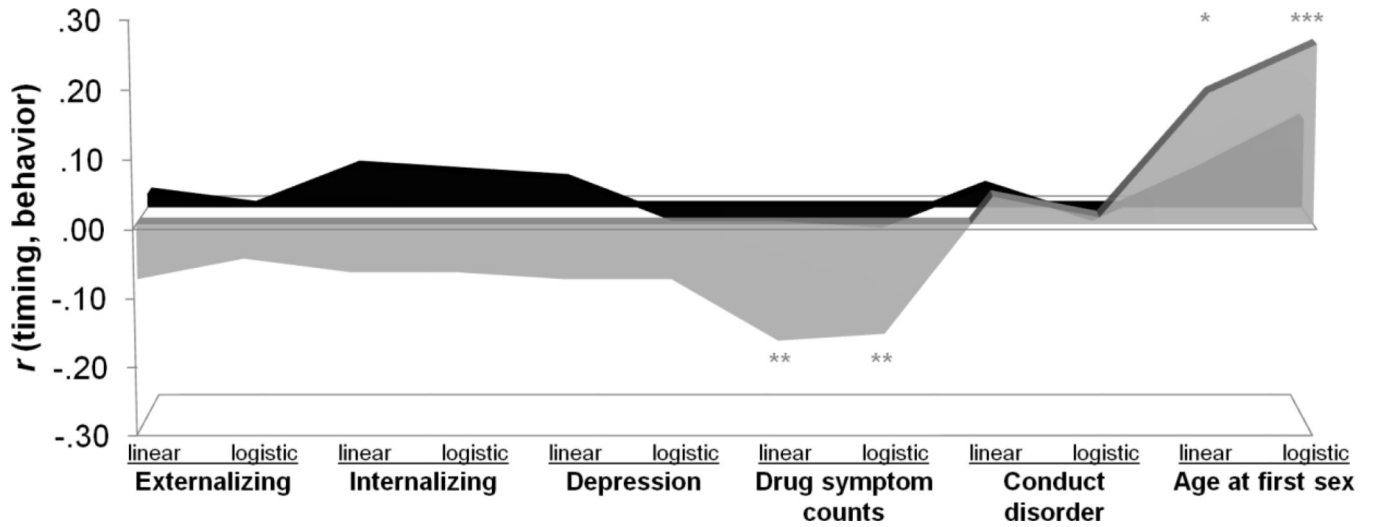


Figure 3.

Area graph of mid-pubertal timing correlations with behavior for boys by method of assessment (linear model, logistic model), for each replicate. The black plot is replicate 1 (range of N 's: 181–314); the gray plot and asterisks are replicate 2 (range of N 's: 179–310). Each peak reflects a single timing-behavior correlation; correlations significantly different from 0, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 1

Descriptive Data for Behavior Problems by Sex

Replicate 1	Girls			Boys			Sex difference effect size <i>d</i>
	Mean (SD)	Range	N	Mean (SD)	Range	N	
Externalizing problems	5.88 (7.09)	0–59.00	290	7.05 (8.03)	0–44.00	289	–.15 [†]
Internalizing problems	6.30 (6.75)	0–40.00	290	5.07 (5.56)	0–25.00	289	.20*
Depression	1.51 (.44)	1.00–3.35	250	1.42 (.34)	1.00–2.85	244	.23**
Drug symptom counts	.06 (1.10)	–.93–6.26	315	.03 (1.15)	–1.14–4.91	307	.03
Conduct disorder	1.73 (2.48)	0–18.00	320	3.18 (3.55)	0–18.00	321	–.47***
Age at first sex	16.60 (2.45)	11.00–26.00	192	16.57 (2.02)	11.00–23.00	190	.01

Replicate 2	Girls			Boys			Sex difference effect size <i>d</i>
	Mean (SD)	Range	N	Mean (SD)	Range	N	
Externalizing problems	6.60 (7.14)	0–44.00	284	8.24 (8.95)	0–64.00	287	–.20*
Internalizing problems	6.64 (6.35)	0–29.00	284	5.32 (5.98)	0–36.00	287	.21*
Depression	1.50 (.38)	1.00–3.15	246	1.48 (.38)	1.00–2.75	225	.05
Drug symptom counts	.07 (1.15)	–.94–4.78	307	.14 (1.20)	–1.04–5.68	296	–.06
Conduct disorder	1.69 (2.24)	0–13.00	313	3.07 (3.45)	0–21.00	320	–.47***
Age at first sex	16.58 (2.19)	12.00–24.00	171	16.35 (1.99)	11.00–24.00	191	.11

Note. Sex differences tested by independent samples *t*-test.

[†] *p* < .10.

* *p* < .05.

** *p* < .01.

*** *p* < .001.

Table 2

Model Fit

Girls	Linear		Logistic	
	Replicate 1 N=359	Replicate 2 N=350	Replicate 1 N=359	Replicate 2 N=350
AIC	2346	2282	1913	1857
BIC	2369	2305	1936	1880

Boys	Linear		Logistic	
	Replicate 1 N=374	Replicate 2 N=377	Replicate 1 N=374	Replicate 2 N=377
AIC	2018	2024	1808	1808
BIC	2042	2047	1831	1831

Note. Lower values indicate better fit.

Table 3

Estimates of Pubertal Timing: Descriptive Data, Parameter Differences, and Parameter Correlations Across Method by Sex

Girls	Replicate 1			Replicate 2		
	Age at mid puberty	Age at menarche	Age at menarche	Age at mid puberty	Age at menarche	Age at menarche
	Linear	Logistic	Traditional	Linear	Logistic	Traditional
Descriptives						
Mean (SD)	13.40 (.91)	13.30 (.89)	12.95 (.99)	13.37 (1.03)	13.24 (.96)	12.87 (1.03)
Range	11.11–16.18	10.65–15.64	10.49–15.67	11.02–16.58	10.59–15.63	10.46–15.50
N	358	359	308	348	349	304
Method Difference Effect Size (<i>d</i>)						
Vs. Logistic	.61***			.70***		
Vs. Traditional	.47***	.36***		.54***	.42***	
Correlations						
With Logistic	.98***			.98***		
With Traditional	.73***	.75***		.75***	.78***	
Boys	Replicate 1			Replicate 2		
	Age at mid puberty	Age at mid puberty	Age at mid puberty	Age at mid puberty	Age at mid puberty	Age at mid puberty
	Linear	Logistic	Linear	Linear	Logistic	Linear
Descriptives						
Mean (SD)	16.03 (1.93)	15.29 (1.15)	15.94 (1.70)	15.17 (.91)		
Range	12.87–23.04	12.43–18.15	13.11–21.89	12.85–17.74		
N	344	374	363	377		
Method Difference Effect Size (<i>d</i>)						
Vs. Logistic	1.80***		1.79***			
Correlations						
With Logistic	.94***		.93***			

Note. All mean sex differences significant, $p < .001$. All correlations significantly different from 0.

 $p < .001$. All mean differences between models significant, $p < .001$.

Table 4

Estimates of Pubertal Tempo: Descriptive Data and Correlations Across Method by Sex

Girls	Replicate 1		Replicate 2	
	At mid puberty	Onset to menarche	At mid puberty	Onset to menarche
	Linear	Logistic	Linear	Logistic
Descriptives				
Mean (SD)	.32 (.03)	.60 (.09)	2.77 (.69)	.59 (.09)
Range	.23-.40	.33-.79	.93-4.68	.32-.81
N	359	359	307	349
Correlations				
With Logistic	.15**		.03	
With Traditional	-.30***	-.10	-.21***	-.03

Boys	Replicate 1		Replicate 2	
	At mid puberty	At mid puberty	Linear	Logistic
Descriptives				
Mean (SD)	.20 (.05)	.39 (.07)	.20 (.04)	.41 (.06)
Range	.10-.35	.19-.56	.10-.34	.23-.58
N	344	372	363	374
Correlations				
With Logistic	.56***		.41***	

Note. For linear and logistic tempo estimates, high scores reflect fast tempo, but for traditional tempo estimates, high scores reflect slow tempo. All mean sex differences significant, $p < .001$. Correlations significantly different from 0.

** $p < .01$.

*** $p < .001$.

Table 5

Correlations Between Pubertal Timing and Tempo within Method by Sex

Girls	Replicate 1		Replicate 2	
	<i>r</i>	<i>N</i>	<i>r</i>	<i>N</i>
Linear	-.94***	358	-.92***	348
Logistic	.16**	359	.32***	348
Traditional	.83***	307	.77***	302

Boys	Replicate 1		Replicate 2	
	<i>r</i>	<i>N</i>	<i>r</i>	<i>N</i>
Linear	-.91***	344	-.95***	363
Logistic	-.34***	372	-.33***	374

Note. For linear and logistic tempo estimates, high scores reflect fast tempo, but for traditional tempo estimates, high scores reflect slow tempo. Correlations significantly different from 0,

**
 $p < .01$,

 $p < .001$.

Table 6

Correlations between Pubertal Tempo and Behavior Problems by Method, by Sex

Girls	Replicate 1						Replicate 2					
	Linear		Logistic		Traditional		Linear		Logistic		Traditional	
	r	N	r	N	r	N	r	N	r	N	r	N
Externalizing problems	.05	261	-.06	259	-.07	259	-.01	261	.02	260	-.07	259
Internalizing problems	.08	261	-.01	259	-.17**	259	.00	261	-.13*	260	-.09	259
Depression	.10	214	-.09	212	.08	212	-.01	208	-.14	207	-.06	207
Drug symptom counts	.11	278	-.05	276	-.07	276	.18**	271	-.03	270	-.09	269
Conduct disorder	.06	282	-.10	280	.04	280	-.02	277	-.03	276	.01	275
Age at first sex	-.17*	174	.11	173	.18*	173	.00	153	.12	153	.25**	151

Boys	Replicate 1			Replicate 2			
	Linear		Logistic	Linear		Logistic	
	r	N	r	r	N	r	
Externalizing problems	-.01	259	.05	.08	266	.08	276
Internalizing problems	-.05	259	-.01	.06	266	.04	276
Depression	.00	211	.00	.08	208	-.10	215
Drug symptom counts	.02	274	.03	.16**	275	.11	284
Conduct disorder	-.07	287	-.09	-.06	298	-.05	308
Age at first sex	-.05	181	-.03	-.20**	179	.03	182

Note. For linear and logistic tempo estimates, high scores reflect fast tempo, but for traditional tempo estimates, high scores reflect slow tempo. Correlations significantly different from 0,

* $p < .05$,

** $p < .01$.