National Norms and Psychometrics for the Pubertal Development Scale

Yinuo Liu[1], Andrea E. Bonny[1,2], Eric A. Youngstrom[1,2]

1. Nationwide Children's Hospital, Columbus OH, U.S.

2. The Ohio State University, Columbus OH, U.S.

Corresponding Author: Dr. Eric A. Youngstrom, Institute of Mental and Behavioral Health Research, Nationwide Children's Hospital, 444 Butterfly Garden Drive, Columbus, OH, 43215, USA. Email: Eric.Youngstrom@nationwidechildrens.org. Phone number: (+1) 2164107975. ORCID: https://orcid.org/0000-0003-2251-6860

Abstract

Introduction: The Pubertal Development Scale (PDS) is widely used for puberty assessment, yet its psychometric properties and norms are limited to research data. This study examined the psychometric properties of parent- and self-report PDS and established continuous norms in nationally representative samples.

Methods: We analyzed two deidentified survey samples: a parent-report sample of children aged 6-18 (*N*=2000, *M*age=11.37, 47.2% female, 74.9% White), and a youth self-report sample aged 12-18 (*N*=754, *M*age=14.33, 49.6% female, 75.3% White). Both samples were representative of the U.S. population on key demographics, and the self-report sample consisted entirely of children whose parents also participated in the parent sample, thus creating parent-child dyads. Internal consistency was evaluated using Cronbach's alpha and McDonald's Omega. Cross-informant agreement was assessed with Intraclass Correlation Coefficient (ICC; two-way model, absolute agreement, single unit) and Bland-Altman plots. Age-dependent norms of each sex were established with Generalized Additive Models for Location, Scale, and Shape (GAMLSS), with 5th-95th percentile curves and reference tables provided.

Results: Parent- and self-report PDS demonstrated acceptable-to-good internal consistency (Cronbach's α: 0.78-0.89; McDonald's ω: 0.79-0.90). Among the 754 parent-youth dyads, excellent cross-informant agreement was observed for both sexes (ICC(2,1)=0.88). Parents' and children's PDS total scores did not differ significantly for boys; for girls, parents rated pubertal development on average 0.13 points lower than children's self-report. Regardless of informants, PDS scores increased nonlinearly with age and exhibited sex-specific developmental patterns. Girls showed earlier pubertal onset, faster progression, and greater convergence toward pubertal completion by late adolescence.

Discussion: The PDS demonstrated strong psychometrics in national samples, supporting its utility in the general pediatric population. The national norms provide empirical benchmarks for PDS score interpretation, strengthening its value as a broad estimation of pubertal status and a pre-screening tool for identifying early or delayed puberty.

Keywords: Puberty, Assessment, Continuous Norm, Cross-Informant Agreement, Psychometrics

Introduction

Puberty, the process of physical development to sexual maturity, is characterized by significant biological, hormonal, and mental changes.[1] Early or delayed puberty could have important implications on adolescent well-being. Early pubertal timing is associated with increased mental and behavioral health problems in both sexes: boys who mature early showed elevated risks of substance use[2] and delinquency,[3] while girls who mature early experience more depression, anxiety, eating problems, and substance use[4–6]. Delayed puberty could signal underlying chronic illness or nutritional deficiencies, cause emotional distress, and impair physical growth.[7] Accurate assessment of puberty is therefore crucial for clinicians and researchers to identify atypical pubertal timing and examine adolescent outcomes.

The Pubertal Development Scale (PDS) is a self- and parent-report measure widely used in research and clinic settings to assess puberty.[8] It asks the respondent to estimate the level of completion of key pubertal events like height spurt, body hair growth, skin changes, and sex-specific indicators. Originally developed by Petersen et al. as a self-report instrument for adolescents, PDS was later adapted to include a parent-report version, with both versions demonstrating good internal consistency and test-retest reliability in research samples.[9] Compared to the gold standard clinician-administered Tanner staging, the PDS does not require trained specialists to conduct physical examination while still showing moderate to good agreement with clinician judgements,[10] further supporting its validity. The PDS has been translated into Chinese, Korean, and Portuguese, with reliability and validity maintained across these languages and cultural contexts.[11–13] As a low-cost, non-invasive instrument with validated psychometrics, the PDS has been adopted in large-scale studies such as the Adolescent Brain Cognitive Development (ABCD) Study.[14,15]

Despite its extensive usage in research, the PDS hasn't been thoroughly examined in a nationally representative sample. Its initial norms were established by Petersen et al.[8] with 355 adolescents aged 11-14 years, and subsequent U.S. studies have been constrained by small sample sizes[9,16] or incomplete demographic representation.[14] National psychometrics and norms for the PDS, currently absent from literature, will benefit research on puberty by validating PDS in the general population and providing reference points for typical versus atypical pubertal development.

Norming the PDS requires careful consideration of sex and age effects, as puberty is characterized by marked sex differences and peak development at certain ages. To capture the nonlinear trajectories of pubertal development, we adopted the continuous norming method [17] and modeled the PDS scores of each sex with Generalized Additive Models for Location, Scale, and Shape (GAMLSS).[18] The GAMLSS approach enables modeling of all distributional parameters, supports hundreds of score distributions, and generates smooth norms unconstrained by artificial age bins, making it increasingly applied in the norming of psychological tests.[19–21] In simulations, GAMLSS outperformed traditional and polynomial methods and achieved higher norming accuracy with smaller samples,[22] further demonstrating this approach's theoretical strengths. Its capacity to model score mean and variance as continuous functions of age is particularly well-suited for capturing pubertal development, where both population average and between-person heterogeneity, reflected by the mean and spread of PDS scores, are of substantive interests.

In this study, we evaluate the psychometric properties of the PDS and establish its continuous norms in nationally representative samples. Both parent- and self-report PDS data are included, and cross-informant agreement is examined. We aim to provide reliable national norms for the PDS and strengthen the foundations of pubertal assessment across research and clinical contexts.

Methods

Table 1. Participant demographic characteristics

	Parent Sample (*N*=2000)	**Youth Sample (*N*=754)**	**2023 ACS[a]**
Child age	11.37 (3.39)	14.33 (1.66)	
6-7 years	17.4%	-	
8-9 years	16.6%	-	
10-11 years	16.0%	-	
12-13 years	18.4%	36.7%	
14-15 years	17.6%	34.5%	
16-18 years	14.0%	28.8%	
Child sex			
Female	47.2%	49.6%	50.5%
Male	52.8%	50.4%	49.5%
Race			
White	74.9%	75.3%	60.5%
Black or African American	9.0%	10.1%	12.1%
Asian	3.7%	1.7%	6.0%
American Indian or Alaska Native	1.8%	0.8%	1.0%
Native Hawaiian or Pacific Islander	0.2%	0.1%	0.2%
Some other race	4.8%	3.3%	7.4%
Multiple races selected	4.2%	8.5%	12.8%
Ethnicity			
Hispanic, Latino, or Spanish Origin	16.8%	14.0%	19.4%
Not Hispanic, Latino, or Spanish Origin	83.2%	86.0%	80.6%
Parent/guardian education			
High school graduate or higher	97.7%	98.0%	89.8%
Bachelor's degree or higher	47.4%	49.1%	36.2%
Median Household Income	$70,000-$79,999	$70,000-$79,999	$77,719

[a]Data from 2023 American Community Survey (ACS) 1-year estimates.

Participants

We conducted secondary analyses with two deidentified survey samples: a parent-report sample (*N*=2000) of children aged 6-18, and a youth self-report sample of children aged 12-18 (*N*=754). The self-report sample consisted entirely of children whose parents also participated in the parent sample, creating parent-child dyads. Both samples were representative of the U.S.

population on demographics (child sex, race, ethnicity) and aligned closely with 2023 American Community Survey (ACS) estimates. Key demographics were summarized along with 2023 ACS estimates in Table 1.

In the parent-report sample, children were on average 11.37 years old (*SD*=3.39) and evenly distributed across the 6-18 age range, with 47.2% of children being female. Based on parent report, 74.9% of children were White, 9% Black or African American, 3.7% Asian, and 16.8% Hispanic, Latino, or Spanish Origin. Nearly all parents (97.7%) had obtained a high school diploma, and 47.4% had obtained a bachelor's degree or higher. The median household income was between $70,000 and $79,999 annually.

In the youth self-report sample, children were on average 14.33 years old (*SD*=1.66) with 49.6% female. Based on youth self-identification, 75.3% were White, 10.1% Black, 1.7% Asian, and 14% Hispanic. Most parents (98%) had a high school diploma, and 49.1% of parents had a bachelor's degree or higher. The median household income was also between $70,000 and $79,999 annually.

Measure

The Petersen Pubertal Development Scale (PDS) is a five-item measure assessing pubertal status. The first three questions on height spurt, body hair growth, and skin changes are the same regardless of child sex. For the next two questions, boys are asked about voice deepening and facial hair growth, and girls are asked about breast development and menarche. All questions except menarche are rated on a 4-point Likert-type scale (1= not yet started, 2 = barely started, 3 = definitely started, 4 = seems complete), whereas the presence of menarche is rated dichotomously (1 = No, 4 = Yes). Prior studies commonly averaged the five items, resulting in continuous scores ranging from 1 to 4. To avoid the extra calculation and preserve the granularity of score changes, we summed all items to create a PDS total score ranging from 5 to 20 when modeling and presenting norms. PDS percentile curves and reference tables rescaled to 1-4 were also included in Supplemental Figure 1 and Supplemental Table 2.

Procedure

This study was approved by the Institutional Review Board at Nationwide Children's Hospital. We contracted with the survey agency YouGov to administer the PDS along with other scales to parents and children. YouGov is a research and data analytics company that has a massive global panel of over 29 million registered members with a significant portion in the U.S., which enables the company to conduct national surveys. YouGov collected and cleaned all responses in 2025, paired parent and child data to create dyads, and applied stratification based on key demographics (child sex, race, ethnicity, and parent education level) to align with the 2023 ACS 1-year estimates, resulting in nationally representative samples.

Analyses

As the PDS contains sex-specific items, all analyses were conducted separately for boys and girls. We evaluated the PDS's internal consistency using Cronbach's alpha and McDonald's omega. To assess the cross-informant agreement, we calculated Intraclass Correlation Coefficient (ICC) for parent-child dyads, using a two-way random effects model with absolute agreement and single unit of analysis, specified as ICC(2,1). ICC values range from 0 to 1, where 0 indicates no agreement and 1 indicates perfect agreement, with thresholds defined as <0.2 poor,

0.2-0.4 fair, 0.4-0.6 moderate, 0.6-0.8 good, 0.8-1.0 excellent.[23] To visualize the agreement between parent- and self-report PDS scores, we created Bland-Altman plots[24] depicting the difference between informants (y-axis) against their mean scores (x-axis). Mean difference and limits of agreement were examined.

We established age- and sex-specific national norms for parent-report and self-report PDS with GAMLSS. Guided by Timmerman et al.,[25] we compared multiple candidate distributions and selected the best-fitting model based on the lowest Akaike Information Criterion (AIC) values (see Supplemental Table 1). The *t*-family distribution best fitted parent-report of both sexes and self-report of boys, while the BCPE distribution best fitted self-report of girls. All GAMLSS models used P-spline smoothers to model age-dependent changes in the mean (μ) and variance (σ), while holding skewness (ν) constant. Worm plots[26] confirmed adequate local fit. For each model, smooth percentile curves (5th, 10th, 25th, 50th, 75th, 90th, 95th) were estimated and plotted, and percentile tables were created for clinical reference.

Results

Internal Consistency

For parent-report PDS, both sexes demonstrated good internal consistency (Cronbach's α=0.89 for both sexes; McDonald's ω=0.89 for boys and 0.90 for girls; see Table 2). For self-report PDS, internal consistency was good for boys (α=0.84, ω=0.85) and acceptable for girls (α=0.78; ω=0.79). Both formats of PDS consistently measure the latent construct of puberty, but parent-report PDS showed higher consistency possibly due to parents' higher reading levels and better observation of pubertal changes.

Table 2. PDS's internal consistency (Cronbach's α and McDonald's ω) and cross-informant agreement (Intraclass Correlation Coefficient with 95% confidence interval)

Internal Consistency	Cronbach's α	McDonald's ω
Boy (parent-report)	.89	.89
Girl (parent-report)	.89	.90
Boy (self-report)	.84	.85
Girl (self-report)	.78	.79
Cross-Informant Agreement	ICC(2,1)	95% CI
Parent-Boy	.88	.86-.90
Parent-Girl	.88	.86-.91

Note. ICC based on two-way random effects model for absolute agreement and single unit of analyses, specified as ICC(2,1).

Cross-Informant Agreement

ICC analyses revealed excellent cross-informant agreement between parents and children regardless of child sex (boys: ICC=.88, 95% CI=.86-.90; girls: ICC=.88, 95% CI=.86-.91), meaning that parents and children were highly consistent in rating the child's pubertal development, with small random variation between informants.

Bland-Altman plots further supported these high ICC values. For boys (Figure 1, Panel A), parents showed a nonsignificant mean difference of -0.07 from their children (paired *t*-test, *p*=0.397), indicating that on average, parents and boys rated pubertal development almost identically. The 95% limits of agreement ranged from -3.04 to 2.90, meaning that for 95% of boy-parent dyads, scores differed by no more than roughly ±3 points. A linear regression of differences on means revealed a small but statistically significant proportional bias (β =.054, *p*=.037), indicating that parent–boy difference shifted slightly in the positive direction as PDS mean scores increased. In other words, at lower mean PDS levels, parents tended to rate boys slightly lower than boys rated themselves, whereas at higher levels ratings converged or became slightly higher. However, the magnitude of this effect was minimal, explaining only 1.2% of variance and suggesting limited practical impact. For girls (Figure 1, Panel B), parents showed a statistically significant mean difference of -0.13 (paired *t*-test, *p*<.05), but this small difference is clinically negligible given that the PDS total scores range from 5 to 20. The 95% limits of agreement were from -2.51 to 2.25, which is a narrower spread than observed for boys. No evidence of proportional bias was found (β=.013, *p*=.619), suggesting that parent–girl differences remained consistent across the PDS score range.

Together, ICC analyses and Bland-Altman plots indicated that for children aged 12-18, parent-report and self-report PDS scores demonstrate strong agreement with acceptable ranges of difference.

Figure 1. Bland-Altman plots depicting parent–youth agreement on PDS total scores, with mean difference, limits of agreement (LoA), and linear regression line of differences on means displayed. Panel A: For boys, the mean difference was not significant, with LoA ranging from −3.04 to 2.90. A small but significant proportional bias was observed (β=.054, *p*=.037). Panel B: For girls, the mean difference of -0.13 was significant, with LoA having a tighter range of −2.51 to 2.25. No evidence of proportional bias was found (β=.013, *p*=.619).

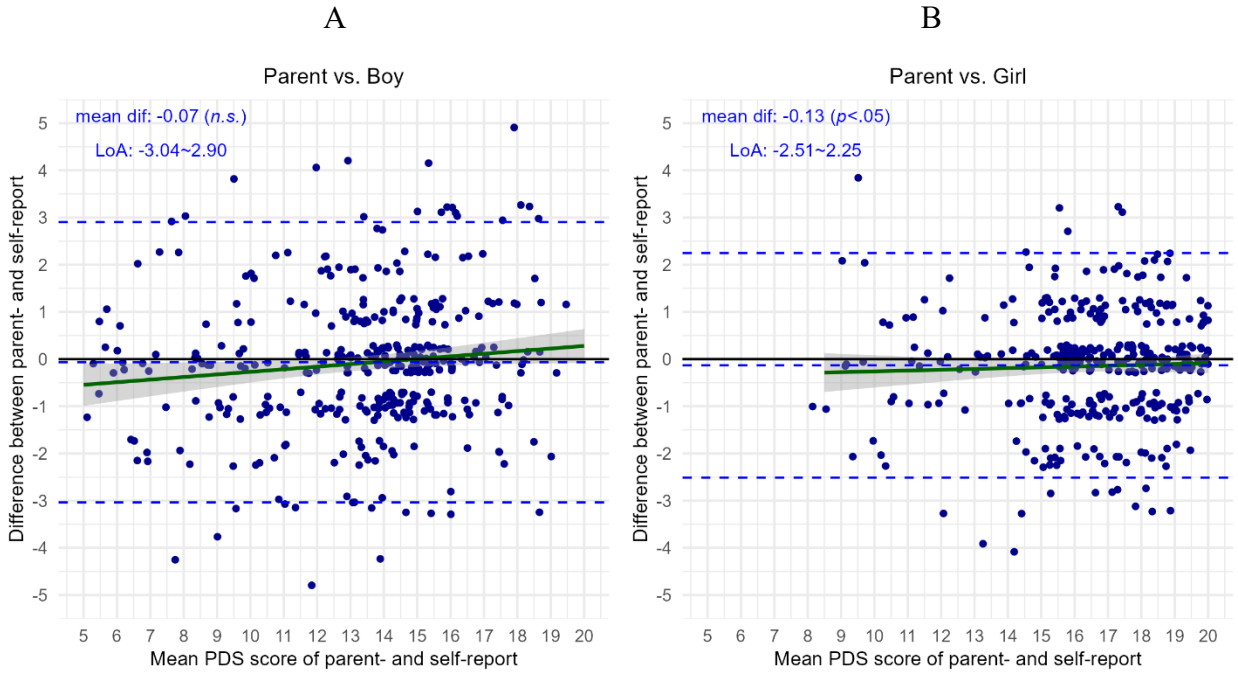
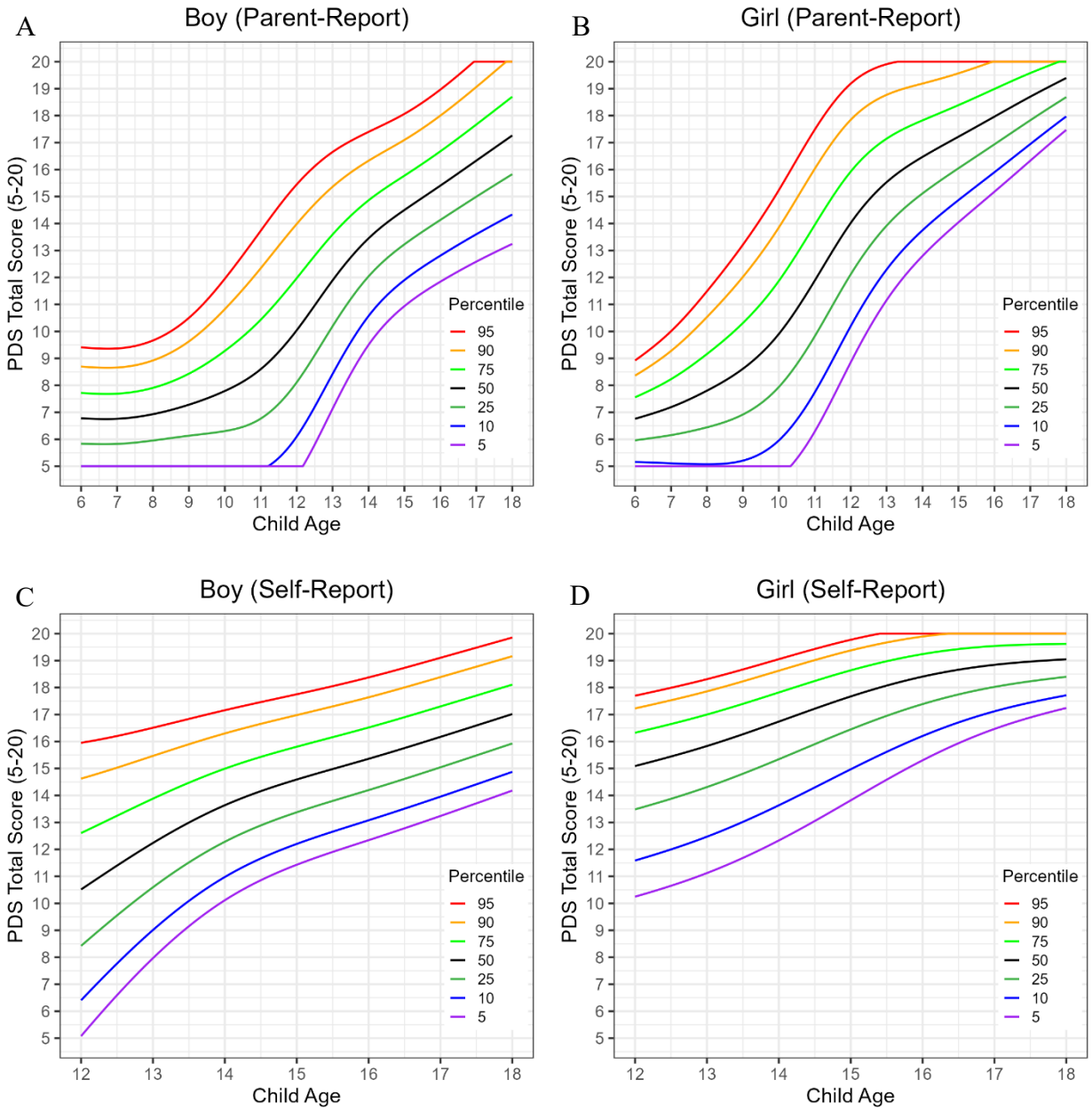


Figure 2. GAMLSS-estimated percentile curves (5th-95th) for parent-report PDS total scores (range: 5-20) across ages 6-18 years and self-report PDS total score across ages 12-18 years. Panel (A) parent report about boys, (B) parent report about girls, (C) self-report for boys, (D) self-report for girls.



Parent-Report Norms

Based on GAMLSS percentile curves (see Figure 2), parent-report PDS scores increased nonlinearly with age and exhibited sex-specific variance patterns. Among boys (Figure 2 Panel A), the median score increased from 7 at age 6 to 17 at age 18, with the steepest growth occurring between ages 11-15. The 5th-95th percentile range was narrow during prepuberty (5 points at ages 6-10), expanded during peak development (7-10 points at ages 11-15), and remained wide during through adolescence (7 points at ages 16-18). Even at age 18, 90% of boys

still scored between 13 and 20, suggesting substantial heterogeneity in boy's pubertal timing. In contrast, girls exhibited earlier pubertal onset and faster progression (Figure 2 Panel B). The median score rose from 7 at age 6 to 19 at age 18, and the peak growth occurred between ages 9-13, showing a two-year advancement compared to boys that is consistent with established biological evidence [27]. The 5th-95th percentile began widening as early as age 7, reached maximum dispersion of 11 points around age 11, then gradually narrowed after age 13. By age 18, 90% of girls scored between 17 and 20, reflecting more uniform pubertal completion among girls. Table 3 presents key percentile estimates rounded to integers of each age-sex combination.

Table 3. Key percentiles of parent-report and self-report PDS total score (range: 5-20) for boys and girls across 6-18 years

Boy Parent-Report

Age	P95	P90	P75	P50	P25	P10	P5
6	9	9	8	7	6	5	-
7	9	9	8	7	6	5	-
8	10	9	8	7	6	5	-
9	10	10	8	7	6	5	-
10	12	11	9	8	6	5	-
11	14	12	10	9	7	5	-
12	15	14	12	10	8	6	5
13	17	15	14	12	10	8	7
14	17	16	15	13	12	11	10
15	18	17	16	14	13	12	11
16	19	18	17	15	14	13	12
17	20	19	18	16	15	14	13
18	-	20	19	17	16	14	13

Girl Parent-Report

Age	P95	P90	P75	P50	P25	P10	P5
6	9	8	8	7	6	5	-
7	10	9	8	7	6	5	-
8	11	11	9	8	6	5	-
9	13	12	10	9	7	5	-
10	15	14	12	10	8	6	5
11	17	16	14	12	10	8	6
12	19	18	16	14	12	10	9
13	20	19	17	16	14	12	11
14	20	19	18	16	15	14	13
15	-	20	18	17	16	15	14
16	-	20	19	18	17	16	15
17	-	-	20	19	18	17	16
18	-	-	20	19	19	18	17

Boy Self-Report

Age	P95	P90	P75	P50	P25	P10	P5
12	16	15	13	11	8	6	5
13	17	15	14	12	11	9	8
14	17	16	15	14	12	11	10
15	18	17	16	15	13	12	11
16	18	18	17	15	14	13	12
17	19	18	17	16	15	14	13
18	20	19	18	17	16	15	14

Girl Self-Report

Age	P95	P90	P75	P50	P25	P10	P5
12	18	17	16	15	13	12	10
13	18	18	17	16	14	12	11
14	19	19	18	17	15	14	12
15	20	19	19	18	16	15	14
16	-	20	19	18	17	16	15
17	-	-	20	19	18	17	16
18	-	-	20	19	18	18	17

Youth Self-Report Norms

Youth self-report PDS scores demonstrated developmental patterns similar to parent-report across the overlapping age range of 12-18 years (see Figure 2). For boys (Figure 2 Panel C), the median score rose from 11 at age 12 to 17 at age 18. Dispersion, reflected by the 5th-95th percentile gap, was widest at age 12 (11 points) and narrowed to 6 points by age 18. For girls (Figure 2 Panel D), the median score rose from 15 at age 12 to 19 at age 18, and the dispersion

decreased from 8 points to 3 points across this age range. Girls again demonstrated greater convergence in pubertal completion than boys by age 18, regardless of informant type (parent-report or youth self-report). Key percentile estimates are presented in Table 3.

Discussion

In nationally representative samples of U.S. children, the PDS demonstrated acceptable to good internal consistency (Cronbach's α: 0.78-0.89) for both parent-report and youth self-report forms, along with excellent cross-informant agreement (ICC=0.88), supporting its utility in general population settings. Using the GAMLSS approach, we established sex-specific continuous norms for the PDS spanning ages 6–18 for parent-report and 12–18 for self-report, presenting both percentile curves and reference tables.

The national norms for the PDS could serve as important benchmarks to advance the interpretation of PDS scores and to aid in the identification of atypical pubertal timing. Pubertal development exhibited sex differences, age-dependent changes, and inter-individual heterogeneity, reflected in the nonlinear trajectories and varying spread of PDS scores across age within each sex. The GAMLSS norms effectively modeled these sources of variations and estimated 5th-95th percentiles to contextualize the PDS. Researchers and clinicians could refer to the median curve (50th percentile) as the expected pubertal trajectory and use the 90th and 10th percentiles to signal early or delayed puberty. Taken together, the GAMLSS norms positioned the PDS within a dynamic interpretative framework that examines more than a score's face value but also its developmental meaning given a child's age and sex.

As parent- and self-report PDS showed excellent cross-informant agreement at ages 12-18, researchers and clinicians could select informants based on their clinical questions and contextual needs. Parent-report PDS, covering a broader age range of 6-18, is preferable when the study population includes younger children for whom the self-assessment is less accurate[28] and normative references are unavailable. Parent-report is also the most viable option for detecting precocious puberty, defined as pubertal onset before age 8 in girls and age 9 in boys,[29] since children at this age range might lack the bodily awareness and vocabulary to report on pubertal changes. Self-report PDS works well with adolescents capable of reliable self-assessment, particularly in contexts that prioritize patient privacy or direct clinician-adolescent interaction. The PDS, with parent- and self-report formats, is a flexible measure that can be adapted to diverse research and clinical settings. As a short and text-based scale, it could be incorporated into the online patient portals (e.g. MyChart) and administered electronically before each visit, providing clinically useful information to healthcare professionals without burdening the clients and families.

This study has two major limitations. First, the PDS psychometrics and norms were derived from population survey data without cross-validation against the gold-standard Tanner staging. Accordingly, the PDS and its normative benchmarks should not be used diagnostically, but rather as a broad estimation of pubertal development in research contexts or as a pre-screener to warrant further assessment in pediatric settings. Second, although the 6–18 age range of the current study offers broader coverage than prior studies, it may not fully capture the complete pubertal development period for all youths. Notably, boys continued to show substantial heterogeneity in PDS scores at age 18, suggesting that pubertal completion could extend beyond the upper age limit of the current norms. Future studies should consider expanding the age range to better represent late-developing adolescents.

In conclusion, this study establishes national psychometrics and continuous norms for the PDS, filling a gap in the standardization of this widely used puberty scale. The PDS showed good internal consistency and high cross-informant agreement between parent- and self-report formats. The GAMLSS norms, modeling on child age and sex, provide empirical benchmarks for score interpretation. These findings together strengthen the psychometric foundation of the PDS and advance its utility in both research and pediatric practice.

References

1. Breehl L, Caban O. Physiology, puberty. Published online 2018.

2. Patton GC, McMorris BJ, Toumbourou JW, Hemphill SA, Donath S, Catalano RF. Puberty and the onset of substance use and abuse. *Pediatrics*. 2004;114(3):e300-e306. doi:10.1542/peds.2003-0626-F

3. Williams JM, Dunlop LC. Pubertal timing and self-reported delinquency among male adolescents. *J Adolesc*. 1999;22(1):157-171. doi:10.1006/jado.1998.0208

4. Angold A, Costello EJ, Worthman CM. Puberty and depression: the roles of age, pubertal status and pubertal timing. *Psychol Med*. 1998;28(1):51-61. doi:10.1017/S003329179700593X

5. Knight JA, Kehm RD, Schwartz L, et al. Prepubertal internalizing symptoms and timing of puberty onset in girls. *Am J Epidemiol*. 2021;190(3):431-438. doi:10.1093/aje/kwaa223

6. Kaltiala-Heino R, Marttunen M, Rantanen P, Rimpelä M. Early puberty is associated with mental health problems in middle adolescence. *Soc Sci Med*. 2003;57(6):1055-1064. doi:10.1016/S0277-9536(02)00480-X

7. Palmert MR, Dunkel L. Delayed Puberty. *N Engl J Med*. 2012;366(5):443-453. doi:10.1056/NEJMcp1109290

8. Petersen AC, Crockett L, Richards M, Boxer A. A self-report measure of pubertal status: Reliability, validity, and initial norms. *J Youth Adolescence*. 1988;17(2):117-133. doi:10.1007/BF01537962

9. Koopman-Verhoeff ME, Gredvig-Ardito C, Barker DH, Saletin JM, Carskadon MA. Classifying pubertal development using child and parent report: comparing the pubertal development scales to Tanner Staging. *J Adolesc Health*. 2020;66(5):597-602. doi:10.1016/j.jadohealth.2019.11.308

10. Pan D, Lin C, Fu S. A systematic review and meta-analysis of the self-reported Pubertal Development Scale's applicability to children. *J Pediatr Endocrinol Metab*. 2025;38(7):718-724. doi:10.1515/jpem-2024-0407

11. Chan NPT, Sung RYT, Nelson EAS, So HK, Tse YK, Kong APS. Measurement of pubertal status with a Chinese self-report Pubertal Development Scale. *Matern Child Health J*. 2010;14(3):466-473. doi:10.1007/s10995-009-0481-2

12. Kim J, Kim D, Pyun H, et al. Validation of the Korean version of the Pubertal Development Scale (PDS-K): a non-invasive self-report tool for epidemiological use. *Epidemiol Health*. 2025;47:e2025059. doi:10.4178/epih.e2025059

13. Pompéia S, Zanini GDAV, Freitas RSD, et al. Adapted version of the Pubertal Development Scale for use in Brazil. *Rev saúde pública*. 2019;53:56. doi:10.11606/s1518-8787.2019053000915

14. Cheng TW, Magis-Weinberg L, Guazzelli Williamson V, et al. A researcher's guide to the measurement and modeling of puberty in the ABCD Study® at baseline. *Front Endocrinol*. 2021;12:608575. doi:10.31234/osf.io/4fv3k

15. Beltz AM, Pham H, Smith T, et al. Research Review: On the (mis)use of puberty data in the ABCD Study® – a systematic review, problem illustration, and path forward. *J Child Psychol*. 2026;67(1):138-151. doi:10.1111/jcpp.70035

16. Brooks-Gunn J, Warren MP, Rosso J, Gargiulo J. Validity of self-report measures of girls' pubertal status. *Child Development*. Published online 1987:829-841.

17. Lenhard A, Lenhard W, Suggate S, Segerer R. A Continuous Solution to the Norming Problem. *Assessment*. 2018;25(1):112-125. doi:10.1177/1073191116656437

18. Rigby RA, Stasinopoulos DM. Generalized Additive Models for Location, Scale and Shape. *J R Stat Soc, Ser C, Appl Stat*. 2005;54(3):507-554. doi:10.1111/j.1467-9876.2005.00510.x

19. Kenny RA, Coen RF, Frewen J, Donoghue OA, Cronin H, Savva GM. Normative values of cognitive and physical function in older adults: findings from the Irish Longitudinal Study on Ageing. *J Am Geriatr Soc*. 2013;61(s2). doi:10.1111/jgs.12195

20. Liu W, Wang J, Zhang H, et al. Toward digital assessment of developmental dyslexia in mainland China: establishing nationwide norms with a GAMLSS approach. *Assessment*. Published online 2025.

21. Rubio-Guerra S, Sánchez-Saudinós MB, Sala I, et al. Development of amyloid-negative neuropsychological norms using GAMLSS. *Alz & Dem Diag Ass & Dis Mo*. 2025;17(4):e70224. doi:10.1002/dad2.70224

22. Frazier T, Youngstrom E, Knaebel B, Uljarevic M. Comparing Traditional and Five Continuous Norming Methods for Neurobehavioral Assessments: A Simulation Study. *Available at SSRN 6270328*.

23. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. Published online 1977:363-374.

24. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*. 1986;327(8476):307-310.

25. Timmerman ME, Voncken L, Albers CJ. A tutorial on regression-based norming of psychological tests with GAMLSS. *Psychol Methods*. 2021;26(3):357-373. doi:10.1037/met0000348

26. Buuren SV, Fredriks M. Worm plot: a simple diagnostic device for modelling growth reference curves. *Stat Med*. 2001;20(8):1259-1277. doi:10.1002/sim.746

27. Brix N, Ernst A, Lauridsen LLB, et al. Timing of puberty in boys and girls: A population-based study. *Paediatric Perinatal Epid*. 2019;33(1):70-78. doi:10.1111/ppe.12507

28. Peng X, Peng Y, Li Y, et al. Validity of web-based self-assessment of pubertal development against pediatrician assessments. *Pediatr Investig*. 2018;2(3):141-148. doi:10.1002/ped4.12050

29. Bradley SH, Lawrence N, Steele C, Mohamed Z. Precocious puberty. *Bmj*. 2020;368.

Authors' Contributions: EAY & YL conceptualized the project; YL ran the analyses and prepared the full manuscript; EAY& AB reviewed the manuscript.

Ethics Committee Approval: The IRB of Nationwide Children's Hospital provided an NHSR determination for analyses of deidentified survey data.

Data Access Statement: Code and sufficient data for replication available upon request from the corresponding author.

Disclosures: Eric A. Youngstrom is the co-founder and Executive Director of Helping Give Away Psychological Science, a 501c3; he has consulted about psychological assessment with Signant Health and received royalties from the American Psychological Association and Guilford Press, and he holds equity in Joe Startup Technologies and held equity in Autism Intervention Measures.

Supplemental Table 1. GAMLSS Model fit indices of candidate distributions

Model Distribution	AIC			
	Boy (Parent)	Girl (Parent)	Boy (Self)	Girl (Self)
BCPE	Not Converge	Not Converge	Not Converge	**1490**
BCCG	4750	4253	Not Converge	Not Converge
Normal	4801	4258	1713	1553
T-Family	**4733**	**4218**	**1707**	1538

Supplemental Table 2. Key Percentiles of PDS average scores (range: 1-4)

Boy Parent-Report

Age	P95	P90	P75	P50	P25	P10	P5
6	1.8	1.8	1.6	1.4	1.2	1.0	-
7	1.8	1.8	1.6	1.4	1.2	1.0	-
8	2.0	1.8	1.6	1.4	1.2	1.0	-
9	2.0	2.0	1.6	1.4	1.2	1.0	-
10	2.4	2.2	1.8	1.6	1.2	1.0	-
11	2.8	2.4	2.0	1.8	1.4	1.0	-
12	3.0	2.8	2.4	2.0	1.6	1.2	1.0
13	3.4	3.0	2.8	2.4	2.0	1.6	1.4
14	3.4	3.2	3.0	2.6	2.4	2.2	2.0
15	3.6	3.4	3.2	2.8	2.6	2.4	2.2
16	3.8	3.6	3.4	3.0	2.8	2.6	2.4
17	4.0	3.8	3.6	3.2	3.0	2.8	2.6
18	-	4.0	3.8	3.4	3.2	2.8	2.6

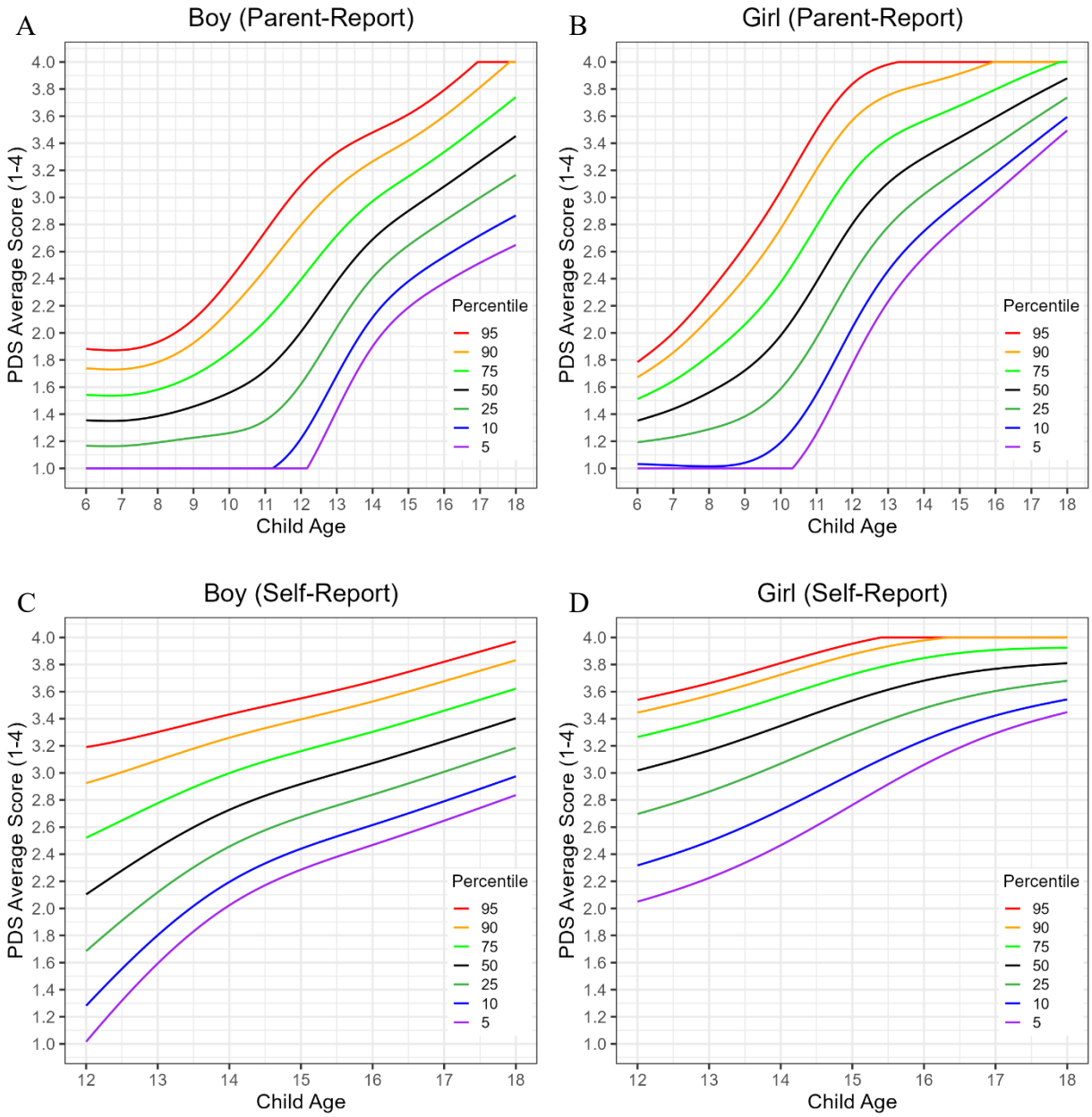
Girl Parent-Report

Age	P95	P90	P75	P50	P25	P10	P5
6	1.8	1.6	1.6	1.4	1.2	1.0	-
7	2.0	1.8	1.6	1.4	1.2	1.0	-
8	2.2	2.2	1.8	1.6	1.2	1.0	-
9	2.6	2.4	2.0	1.8	1.4	1.0	-
10	3.0	2.8	2.4	2.0	1.6	1.2	1.0
11	3.4	3.2	2.8	2.4	2.0	1.6	1.2
12	3.8	3.6	3.2	2.8	2.4	2.0	1.8
13	4.0	3.8	3.4	3.2	2.8	2.4	2.2
14	4.0	3.8	3.6	3.2	3.0	2.8	2.6
15	-	4.0	3.6	3.4	3.2	3.0	2.8
16	-	4.0	3.8	3.6	3.4	3.2	3.0
17	-	-	4.0	3.8	3.6	3.4	3.2
18	-	-	4.0	3.8	3.8	3.6	3.4

Boy Self-Report

Age	P95	P90	P75	P50	P25	P10	P5
12	3.2	3.0	2.6	2.2	1.6	1.2	1.0
13	3.4	3.0	2.8	2.4	2.2	1.8	1.6
14	3.4	3.2	3.0	2.8	2.4	2.2	2.0
15	3.6	3.4	3.2	3.0	2.6	2.4	2.2
16	3.6	3.6	3.4	3.0	2.8	2.6	2.4
17	3.8	3.6	3.4	3.2	3.0	2.8	2.6
18	4.0	3.8	3.6	3.4	3.2	3.0	2.8

Girl Self-Report

Age	P95	P90	P75	P50	P25	P10	P5
12	3.6	3.4	3.2	3.0	2.6	2.4	2.0
13	3.6	3.6	3.4	3.2	2.8	2.4	2.2
14	3.8	3.8	3.6	3.4	3.0	2.8	2.4
15	4.0	3.8	3.8	3.6	3.2	3.0	2.8
16	-	4.0	3.8	3.6	3.4	3.2	3.0
17	-	-	4.0	3.8	3.6	3.4	3.2
18	-	-	4.0	3.8	3.6	3.6	3.4

Supplemental Figure 1. GAMLSS-estimated percentile curves (5th-95th) for parent-report PDS average scores (range: 1-4) across ages 6-18 and self-report PDS total score across ages 12-18. C. Self-report of boys. D. Self-report of girls.