Developmental Variation in Amygdala Volumes: Modeling Differences Across Time, Age, and Puberty

SUPPLEMENTAL INFORMATION

Structural Image Processing and Validation Pipeline and Procedures

Pre-processing

Scans were received unprocessed, in compressed, NIFTI-1 format. Several preprocessing steps were performed using MRTrix3 (Tournier et al., 2012) and included, 1) reversing the byte order for each file (little-endian to big-endian), 2) reorientation of the images to Right-Anterior-Superior format, 3) random evaluation of file integrity and completion.

Processing, Parcellation, and Segmentation

Subcortical segmentation were performed using the longitudinal processing pipeline available in FreeSurfer (v6.0), a widely-used and freely-distributed software package (surfer.nmr.mgh.harvard.edu) that performs automated processing and analyses of brain imaging data. The technical details underlying the multi-step technical details of these procedures are extensively described in prior publications, and a cursory review will be provided here (Dale et al., 1999; Dale & Sereno, 1993; Fischl et al., 2001, 2002; Fischl, Salat, et al., 2004; Fischl, Sereno, & Dale, 1999; Fischl, Sereno, Tootell, et al., 1999; Fischl, van der Kouwe, et al., 2004; Fischl & Dale, 2000; Han et al., 2006; Jovicich et al., 2006; Reuter et al., 2010, 2012; Segonne et al., 2004). Initial processing involves the creation of an unbiased within-subject template space and image, created using robust, inverse consistent registration of scans across time points. Several subsequent processing steps are initialized with common information from this template, significantly increasing reliability and statistical power. FreeSurfer's recon-all completes two sequential processing algorithms to extract anatomical estimates. In an initial 'cortical surface stream', grey matter thickness and curvature are extracted using a complex deformation process. A second processing stream is then initiated to extract the volume from subcortical structures. The original volume is first registered to the MNI 305 atlas using an affine transformation, followed by initial labeling of subcortical structures, and N3 intensity correction (Sled et al., 1998). The volume is

then non-linearly registered to the MNI305 atlas. Voxelwise labeling is performed using an iterative probabilistic classification technique (Fischl et al., 2002). Segmentation of the amygdala is separately performed using an atlas derived from ultra-high-resolution *ex vivo* scans of *n=10* tissue blocks containing the amygdala. Amygdala boundaries and sub-nuclei were segmented using an iterative Bayesian inference algorithm that incorporated manual delineations of *ex vivo* amygdalae, and learned using an *in vivo* data set previously processed using FreeSurfer's 'recon-all'. Results from the training process were extra

Results derived from the FreeSurfer pipeline used in this study were validated by comparing volumetric estimates to ranges reported in previous studies of developing neuroanatomy. Mean values for the left and right amygdala were 1576.07 and 1632.71mm^3, respectively, closely approximating results from similar studies reporting these estimates in typically developing youth (Goddings et al., 2014; Østby et al., 2009; Uematsu et al., 2012; van der Plas et al., 2010), as well as Albaugh et al. (2017) who extracted the same information from the data, albeit using different processing methods. An initial validation procedure attempted to validate the FreeSurfer-derived volumetric estimates of various brain regions by comparison against original values obtained by NIHPD. Data sets accompanying the raw scan files contained ANIMAL-based estimates of several prominent regions. These included the lateral divisions of the cerebral lobes (separate estimates for grey and white matter), caudate, cerebellum, globus pallidus, putamen, and thalamus. However, subsequent investigation revealed marked discrepancies across ANIMAL and FreeSurfer in the segmentation procedures used to define these regions. For example, the original, ANIMAL-based estimates appear to incorporate portions of the insular cortex into estimates of frontal and temporal lobe volume. Similarly, whereas the original values for the thalamus' volume seem to reflect the size of the entire structure, estimates

provided by FreeSurfer refer only to the thalamus proper (which does not include either the epi- or perithalamus). Ultimately, it was necessary to limit direct comparison to the left and right putamen, caudate, total intracranial volume, total grey matter, and total white matter.

Reliability of the values obtained was assessed by summarizing the similarity of measurements for these regions of interest using the intraclass correlation coefficient (ICC, two-way mixed average – consistency; Shrout & Fleiss, 1979; Strother & Churchill, 2017). Shou et al. (2013) developed the image intraclass correlation coefficient or I2C2, a variant on the classic ICC, extended to better capture the high-dimensional, multivariate nature of neuroimaging data. Values for the I2C2 and ICC have equivalent interpretations - both statistics are generally bounded between 0 and 1 with higher values indicative of better reliability. Overall, both statistics suggested excellent consistency across measurements made with the ANIMAL and FreeSurfer pipelines (> 0.85; see Table S1).

Missing Data

Meaningful patterns of missingness in the data were evaluated using Little's test (1988), which indicated that an assumption of "Missing Completely at Random" (MCAR) was not viable, $\chi^2(42) = 211.731$, $p < .001$. A series of separate variance *t*-tests were conducted across cases with missing and non-missing values to identify meaningful variation in the incomplete data. Results indicated that younger participants were more likely to be missing data in variables derived from the Pubertal Development Scale, with a mean difference of 3.3 years across cases with and without missing values ($p < .001$). Participants providing data at second or third assessments were also more likely to be missing data on annual family income ($p < .001$). No other meaningful patterns in missing data were detected. For these analyses, data were limited to those cases without missing

values for pubertal development or age, ensuring an equivalent sample size across models incorporating respective predictors.

Mixed Effects Modeling

Mixed effects modeling (also referred to as hierarchical or multi-level modeling), is an extension of the genera linear model to incorporate parameters that vary at more than one level of observation, and is an appropriate method of analyzing CS data (Hoffman, 2015). In the case of a two-level design (e.g., multiple observations nested within individuals), the mixed effects framework separates variance in the dependent variable attributable to the observations within individuals (within-subjects or 'Level 1' variance) or to the individuals themselves (between-subjects or 'Level 2' variance). This approach effectively accounts for the non-independence of nested data points. In longitudinal applications, mixed effects models allow researchers to simultaneously consider intra- and inter-individual change trajectories. Moreover, mixed effects models are known to be robust to uneven time spacing in collection, as well as the presence of missing data at various time points (Hoffman, 2015).

In the current analyses, mixed effects model parameters were estimated using the *lme4* package (v1.1.16; (Bates et al., 2018) available in *R* (v3.5.x). Parameter values were derived using maximum likelihood estimation as permissible for the sample size (Carey, 2013) and appropriate given the unbalanced nature of the data (Raudenbush, 1995). Wald tests were used to determine the significance of fixed effects, using the Kenward-Roger adjusted degrees of freedom (Halekoh & Højsgaard, 2014; Kenward & Roger, 1997). The significance of individual random effects was tested by examining the difference in fit (as -2-log-likelihood) of nested models with and without the term. The log likelihood distribution approximates the chi-square distribution; therefore, this

test is akin to a chi-square difference test. Model parameters were determined using maximum-likelihood estimation, which is appropriate for samples of this size (Hoffman, 2015). Where necessary, significant interactions were decomposed using the Johnson-Neyman technique (Johnson & Neyman, 1936), which provides the range of values along a continuous moderator (e.g., *Age@1stScan*), for which the association between the predictor and outcome is significant (Preacher et al., 2006).

Model Comparisons

As shown in Table S3, the *Combined Model*, which included effects of chronological age as well as pubertal development (rather than either effect alone) exhibited the best fit to right amygdala volume data. The ΔAIC of 10.01 well exceeds the suggested threshold of 5.9, indicating the superiority of this model in comparison to the *Pubertal Development* or *Chronological Age* models. The AIC weight of .99, reflects the probability that the observed data were created by the *Combined Model* in comparison to the alternatives. Results comparing model fit for left amygdala volumes were less conclusive. Though the *Pubertal Development* model had the lowest AIC value, this was only slightly greater than that of the *Combined Model* (ΔAIC = 0.66). Therefore, the superiority of any one model predicting change in left amygdala volumes could not be determined.

Combined Model

A *Combined Model* tested the effects of chronological age and pubertal development in a single model predicting amygdala volumes (mixed model formulas provided in Table S1). Fixed effects of time in study (*Time*), change in pubertal development during the study (*TSChange*), and

their interaction (*Time* x *TSChange*) were added to a random intercept model predicting right amygdala volumes. Each of these fixed effects terms (though none of the corresponding random effects) were significant in a model predicting right amygdala volumes (Table S2-A). Participants showed an increase of 10.03mm^3 ($p < .05$) for each year of participation in the study (*Time*), and an increase of 50.88mm^3 ($p < .01$) for each Tanner stage (*TSChange*) reached during participation in the study, controlling for puberty and age, respectively. The significant coefficient for the interaction term (*Time* x *TSChange*) indicates that the pace of right amygdala growth slows as youth progress through puberty, in that for each Tanner stage reached, the rate of change declines by −9.74mm^3/year. A final 'unconditional' model was retained that included only the significant fixed effects.

Between-subjects predictors age at first scan (*Age@1stScan*) and pubertal status at first scan (*TS@1stScan*) were added to the unconditional model. *Age@1stScan* significantly affected the intercept, such that for each year of youths' age at initial assessment, right amygdala volumes were 24.14mm^3 larger on average ($p < .001$; Table S2-A). Pubertal development at initial assessment (*TS@1stScan*) similarly predicted the size of right amygdalae, though the direction of the effect was reversed. For each Tanner stage reached by the time participants started the study, right amygdala volumes were −37.06mm^3 smaller on average ($p < .05$). Two-way cross-level interaction terms (*Time* x *Age@1stscan, Time* x *TS@1stScan, TSChange* x *Age@1stScan*, and *TSChange* x *TS@1stscan*) were added to the model next. Age at first scan (*Age@1stScan*) moderated the slope of growth across pubertal development (*TSChange*), in that older youth entering the study tended to show less change in right amygdala volumes as they passed through puberty. Specifically, for each year of *Age@1stScan*, the rate of change decreased by −9.42mm^3/Tanner stage ($p < .05$). No other cross-level interaction terms reached significance.

Next, separate three-way cross-level interaction terms were added to the model, reflecting 'moderated moderation' of age at first scan and pubertal development at first scan on the interaction between pubertal progress and time in study (e.g., *Age@1stScan* x *TSChange* x *Time; TS@1stScan* x *TSChange* x *Time*). Neither effect was significant, as shown in Table 5-A. Sex-specific analyses did not reveal meaningful variation as shown in Tables S10.

Corresponding analyses were conducted to evaluate a *Combined Model* predicting left amygdala volumes. Results from these analyses are provided in Table S2-B. None of the fixed effect terms significantly predicted growth in left amygdala volumes, though random effects of *Time* and *TSChange* suggest meaningful between-subjects differences in slopes. Between-subjects predictors age at first scan (*Age@1stScan*) and pubertal status at first scan (*TS@1stScan*) were added next. As in the right amygdala, *Age@1stScan* significantly influenced the intercept, in that for each year of a participant's age at initial assessment, left amygdala volumes were 21.95mm^3 larger on average ($p < .001$; Table S2-B). Pubertal development at study entry (*TS@1stScan*) also predicted left amygdala volumes at initial scan, though as in the right amygdala, the effect was reversed. For each Tanner stage at initial assessment (*TS@1stScan*), participants showed a -45.65mm^3 reduction in left amygdala volumes ($p < .01$). Interaction terms were added next, though none reached significance in the full sample. However, in analyses specific to boys, a three-way interaction between *Time, TSChange,* and *TS@1stScan* reached significance, such that among pubertally advanced youth, pubertal maturation over the course of the study predicted a decline in amygdala volume (Table S11).

Supplemental References

Albaugh, M. D., Nguyen, T.-V., Ducharme, S., Collins, D. L., Botteron, K. N., D'Alberto, N., Evans, A. C., Karama, S., Hudziak, J. J., & Group, B. D. C. (2017). Age-related volumetric change of limbic structures and subclinical anxious/depressed symptomatology in typically developing children and adolescents. *Biological Psychology*, *124*, 133–140.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2018). *lme4: Linear Mixed-Effects Models Using "Eigen" and S4* (1.1.16) [Computer software]. http://cran.r-project.org/package=lme4

Carey, G. (2013). *Quantitative methods in neuroscience*. Author. http://psych.colorado.edu/~carey/qmin/QMIN_2013_03_17.pdf

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage*, *9*(2), 179–194.

Dale, A. M., & Sereno, M. I. (1993). Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience*, *5*(2), 162–176. https://doi.org/10.1162/jocn.1993.5.2.162

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, *97*(20), 11050–11055. https://doi.org/10.1073/pnas.200033797

Fischl, B., Liu, A., & Dale, A. M. (2001). Automated manifold surgery: Constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Transactions on Medical Imaging*, *20*(1), 70–80. https://doi.org/10.1109/42.906426

Fischl, B., Salat, D. H., Busa, E., Albert, M. S., Dieterich, M., Haselgrove, C., van der Kouwe, A. J., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., & Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*, 341–355.

Fischl, B., Salat, D. H., van der Kouwe, A., Makris, N., Ségonne, F., Quinn, B. T., & Dale, A. M. (2004). Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, *23*(Supplement 1), S69–S84. https://doi.org/DOI: 10.1016/j.neuroimage.2004.07.016

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis: II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, *9*(2), 195–207.

Fischl, B., Sereno, M. I., Tootell, R. B. H., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, *8*(4), 272–284. https://doi.org/10.1002/(SICI)1097-0193(1999)8:4<272::AID-HBM10>3.0.CO;2-4

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J. M., & Kennedy, D. N. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex*, *14*(1), 11–22.

Goddings, A.-L., Mills, K. L., Clasen, L. S., Giedd, J. N., Viner, R. M., & Blakemore, S.-J. (2014). The influence of puberty on subcortical brain development. *NeuroImage*, *88*, 242–251.

Halekoh, U., & Højsgaard, S. (2014). A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models–The R package pbkrtest. *Journal of Statistical Software*, *59*(9), 1–30.

Han, X., Jovicich, J., Salat, D. H., van der Kouwe, A., Quinn, B. T., Czanner, S., Busa, E., Pacheco, J., Albert, M. S., Killiany, R. J., Maguire, R. P., Rosas, H. D., Makris, N., Dale, A. M., Dickerson, B. C., & Fischl, B. (2006). Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *NeuroImage*, *32*(1), 180–194.

Hoffman, L. (2015). *Longitudinal analysis: Modeling within-person fluctuation and change*. Routledge.

Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. *Statistical Research Memoirs*, *1*, 57–93.

Jovicich, J., Czanner, S., Greve, D. N., Haley, E., van der Kouwe, A., Gollub, R. L., Kennedy, D. N., Schmitt, F., Brown, G., MacFall, J. R., Fischl, B., & Dale, A. M. (2006). Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *NeuroImage*, *30*(2), 436–443. https://doi.org/DOI: 10.1016/j.neuroimage.2005.09.046

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 983–997.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Østby, Y., Tamnes, C. K., Fjell, A. M., Westlye, L. T., Due-Tønnessen, P., & Walhovd, K. B. (2009). Heterogeneity in subcortical brain development: A structural magnetic resonance imaging study of brain maturation from 8 to 30 years. *Journal of Neuroscience*, *29*(38), 11772–11782.

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*(4), 437–448.

Raudenbush, S. W. (1995). Maximum likelihood estimation for unbalanced multilevel covariance structure models via the EM algorithm. *British Journal of Mathematical and Statistical Psychology*, *48*(2), 359–370. https://doi.org/10.1111/j.2044-8317.1995.tb01068.x

Reuter, M., Rosas, H. D., & Fischl, B. (2010). Highly accurate inverse consistent registration: A robust approach. *NeuroImage*, *53*(4), 1181–1196. https://doi.org/10.1016/j.neuroimage.2010.07.020

Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, *61*(4), 1402–1418. https://doi.org/10.1016/j.neuroimage.2012.02.084

Segonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D. H., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, *22*(3), 1060–1075. https://doi.org/DOI: 10.1016/j.neuroimage.2004.03.032

Shou, H., Eloyan, A., Lee, S., Zipunnikov, V., Crainiceanu, A. N., Nebel, M. B., Caffo, B., Lindquist, M. A., & Crainiceanu, C. M. (2013). Quantifying the reliability of image replication studies: The image intra-class correlation coefficient (I2C2). *Cognitive, Affective & Behavioral Neuroscience*, *13*(4), 714–724. https://doi.org/10.3758/s13415-013-0196-0

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428.

Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *Medical Imaging, IEEE Transactions On*, *17*(1), 87–97.

Strother, S. C., & Churchill, N. (2017). Neuroimage preprocessing. In H. Ombao, M. Lindquist, W. Thompson, & J. Aston (Eds.), *Handbook of neuroimaging data analysis* (pp. 264–308). Taylor & Francis.

Tournier, J.-D., Calamante, F., & Connelly, A. (2012). MRtrix: Diffusion tractography in crossing fiber regions. *International Journal of Imaging Systems and Technology*, *22*(1), 53–66. https://doi.org/10.1002/ima.22005

Uematsu, A., Matsui, M., Tanaka, C., Takahashi, T., Noguchi, K., Suzuki, M., & Nishijo, H. (2012). Developmental trajectories of amygdala and hippocampus from infancy to early adulthood in healthy individuals. *PloS One*, *7*(10), e46970.

van der Plas, E. A., Boes, A. D., Wemmie, J. A., Tranel, D., & Nopoulos, P. (2010). Amygdala volume correlates positively with fearfulness in normal healthy girls. *Social Cognitive and Affective Neuroscience*, *5*(4), 424–431.

Table S1

Consistency of volumetric estimates derived from NIHPD ANIMAL and FreeSurfer pipelines

	ICC[a]	I2C2
Caudate		
Right	.92	.94
Left	.90	.94
Putamen		
Right	.91	.91
Left	.92	.91
Intracranial Volume	.97	.96
Intracranial Grey Matter	.94	.92
Cerebral White Matter	.93	.95

[a] $df_{1,2}$ = 633, 633

Note. ICC = Intraclass Correlation Coefficient (two-way mixed average measure, consistency). I2C2 = Image Intraclass Correlation Coefficient.

Table S2-A

Change in Right Amygdala Volumes as a Function of Chronological Time and Pubertal Maturation

	Right Amygdale Volume (mm^3)									
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters										
Intercept	1635.98***	10.69	1631.55***	12.16	1435.06***	48.45	1389.41***	49.40	1390.70***	49.52
Time			10.03*	4.24	10.47*	4.26	36.56*	15.44	35.52*	16.27
TSChange			50.88**	16.30	48.96**	16.31	131.25**	45.53	91.77	84.37
Time x TSChange			-9.74*	4.82	-9.71*	4.82	-10.97*	5.07	1.97	25.00
Age@1stScan					24.14***	6.43	29.11***	6.53	29.08***	6.55
TS@1stScan					-37.06*	16.52	-42.66*	16.69	-43.16*	16.73
Time x Age@1stScan							-1.72	2.27	-1.67	2.48
Time x TS@1stScan							-1.83	6.32	-1.46	6.82
TSChange x Age@1stScan							-9.42*	4.75	-8.39	8.59
TSChange x TS@1stScan							15.52	13.02	30.97	24.45
Time x TSChange x Age1stScan									-0.33	2.65
Time x TSChange x TS@1stScan									-5.28	7.47
Random Effects Parameters										
σ^2	7862.962		6704.878		6696.349		6129.389		6104.549	
τ_{00}	32926.234		34093.613		31834.560		32054.573		32055.209	
n	330		274		274		274		274	
Scans	637		513		513		513		513	
Deviance	8226.705		6599.369		6582.388		6560.180		6559.093	

Note. Time = Age at current scan – Age at first scan; *TSChange* = Tanner stage at current scan – Tanner stage at first scan; *Age@1stScan* = chronological age at initial assessment/study entry; *TS@1stScan* = pubertal development (Tanner stage) at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Table S2-B

Change in Left Amygdala Volumes as a Function of Chronological Time and Pubertal Maturation

	Left Amygdala Volume (mm^3)									
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters										
Intercept	1580.11***	9.81	1579.21***	11.30	1430.70***	44.99	1412.56***	45.60	1411.38***	45.71
Time			1.15	3.99	1.42	3.97	11.66	14.85	14.18	15.62
TSChange			12.76	15.34	7.67	15.46	74.96	47.97	83.88	79.81
Time x TSChange			0.64	4.54	2.88	4.45	3.48	4.82	-0.28	23.57
Age@1stScan					21.95***	5.97	24.26***	6.03	24.56***	6.05
TS@1stScan					-45.65**	15.29	-49.40**	15.41	-50.49**	15.44
Time x Age@1stScan							-1.43	2.21	-2.06	2.42
Time x TS@1stScan							1.91	6.16	3.96	6.63
TSChange x Age@1stScan							-8.39	5.17	-12.43	8.22
TSChange x TS@1stScan							17.07	14.24	39.72	23.50
Time x TSChange x Age1stScan									1.57	2.50
Time x TSChange x TS@1stScan									-8.43	7.03
Random Effects Parameters										
σ^2	6302.988		5943.494		4693.055		4663.495		4650.129	
τ_{00}	27944.557		29340.910		27860.193		27851.561		27877.427	
n	330		274		274		274		274	
Scans	637		513		513		513		513	
Deviance	8102.344		6530.337		6508.588		6500.777		6499.371	

Note. Time = Age at current scan – Age at first scan. *TSChange* = Tanner stage at current scan – Tanner stage at first scan. *Age@1stScan* = chronological age at initial assessment/study entry. *TS@1stScan* = pubertal development (Tanner stage) at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Table S3

Comparative Fit of Models Predicting Change in Right and Left Amygdala Volumes

	k	AIC	ΔAIC	AIC Weight
Right Amygdala				
1. Combined	14	5432.24	—	0.99
2. Pubertal Development	6	5442.25	10.01	0.01
3. Chronological Age	6	6647.69	1215.45	0.00
Left Amygdala				
1. Pubertal Development	6	5376.83	—	0.58
2. Combined	14	5377.49	0.66	0.42
3. Chronological Age	6	6591.40	1214.57	0.00

Note. k = # of model parameters. AIC = Akaike Information Criterion.

Table S4-A

Change in Right Amygdala Volumes as a Function of Time and Age at Initial Scan among Boys

	Right Amygdala Volume (mm^3)							
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters								
Intercept	1741.49***	13.94	1726.71***	14.51	1542.42***	45.30	1513.61***	46.10
Time			16.91***	4.03	18.12***	4.52	70.13***	12.99
Age@1stscan					15.94***	3.74	18.40***	3.81
Time x Age@1stScan							-4.54***	1.07
Random Effects Parameters								
σ^2	7815.889		6686.744		4735.929		5198.186	
τ_{00}	23328.475		24567.899		22543.395		22265.877	
n	146		146		146		146	
Scans	266		266		266		266	
Deviance	3401.317		3385.191		3364.323		3350.306	

Note. Time = Age at current scan – Age at first scan. *Age@1stScan* = chronological age at initial assessment/study entry.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$. σ^2 = Random effects variance. τ_{00} = Between-subjects variance.

Table S4-B

Johnson-Neyman Significance Regions for the Conditional Effect of Age at First Scan on Change in Right Amygdala Volumes across Time in Study in Boys

Age at 1st Scan	Slope	SE	LCI	UCI	*t*	*p*
20.00	-21.08	9.87	-40.41	-1.74	-2.14	.034
19.00	-16.51	8.89	-33.94	0.92	-1.86	.065
18.00	-11.95	7.95	-27.52	3.63	-1.50	.135
17.00	-7.39	7.04	-21.18	6.40	-1.05	.296
16.00	-2.82	6.18	-14.93	9.29	-0.46	.648
15.00	1.74	5.40	-8.83	12.32	0.32	.747
14.00	6.31	4.73	-2.96	15.57	1.33	.184
13.00	10.87	4.23	2.57	19.17	2.57	.011
12.00	15.43	3.97	7.65	23.22	3.89	< .001
11.00	20.00	3.99	12.18	27.82	5.01	< .001
10.00	24.56	4.28	16.16	32.96	5.73	< .001
9.00	29.12	4.81	19.70	38.54	6.06	< .001
8.00	33.69	5.49	22.93	44.45	6.14	< .001
7.00	38.25	6.28	25.94	50.57	6.09	< .001
6.00	42.82	7.15	28.80	56.83	5.99	< .001
5.00	47.38	8.07	31.57	63.19	5.87	< .001

Note. LCI = 95% lower confidence interval. UCI = 95% upper confidence interval.

Table S5-A

Change in Right Amygdala Volumes as a Function of Time and Age at Initial Scan among Girls

	Right Amygdala Volume (mm^3)							
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters								
Intercept	1553.60***	12.68	1533.87***	13.09	1403.00***	39.44	1381.02***	40.29
Time			18.21***	3.23	21.44***	3.97	65.70***	11.84
Age@1stscan					11.46***	3.32	13.45***	3.40
Time x Age@1stScan							-3.98***	1.00
Random Effects Parameters								
σ^2	7864.698		6784.542		4143.717		4368.976	
τ_{00}	25062.644		25380.051		24061.928		24396.493	
n	184		184		184		184	
Scans	371		371		371		371	
Deviance	4738.711		4709.187		4694.625		4679.824	

Note. Time = Age at current scan – Age at first scan. *Age@1stScan* = chronological age at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Table S5-B

Johnson-Neyman Significance Regions for the Conditional Effect of Age at First Scan on Change in Right Amygdala Volumes across Time in Study in Girls

Age at 1st Scan	Slope	SE	LCI	UCI	*t*	*p*
20.00	-13.74	8.17	-29.76	2.27	-1.68	.094
19.00	-10.02	7.38	-24.48	4.43	-1.36	.176
18.00	-6.30	6.60	-19.24	6.63	-0.95	.341
17.00	-2.58	5.85	-14.05	8.89	-0.44	.660
16.00	1.14	5.14	-8.93	11.21	0.22	.825
15.00	4.86	4.48	-3.92	13.65	1.08	.279
14.00	8.58	3.91	0.92	16.25	2.20	.029
13.00	12.30	3.46	5.52	19.09	3.55	< .001
12.00	16.03	3.20	9.76	22.29	5.01	< .001
11.00	19.75	3.15	13.56	25.93	6.26	< .001
10.00	23.47	3.35	16.91	30.03	7.01	< .001
9.00	27.19	3.74	19.87	34.51	7.28	< .001
8.00	30.91	4.27	22.54	39.28	7.24	< .001
7.00	34.63	4.90	25.03	44.23	7.07	< .001
6.00	38.35	5.59	27.39	49.32	6.86	< .001
5.00	42.07	6.33	29.67	54.48	6.65	< .001

Note. LCI = 95% lower confidence interval. UCI = 95% upper confidence interval.

Table S6-A

Change in Left Amygdala Volumes as a Function of Time and Age at Initial Scan among Boys

	Right Amygdala Volume (mm^3)							
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters								
Intercept	1664.49***	13.89	1654.31***	14.33	1555.04***	45.59	1541.66***	46.14
Time			11.70**	3.71	12.79**	4.28	38.84**	13.56
Age@1stscan					8.57*	3.77	9.71*	3.82
Time x Age@1stScan							-2.28	1.13
Random Effects Parameters								
σ^2	6163.180		5625.818		3797.357		3868.289	
τ_{00}	24153.246		24793.151		23520.456		23506.584	
n	146		146		146		146	
Scans	266		266		266		266	
Deviance	3371.547		3362.099		3346.928		3343.062	

Note. Time = Age at current scan – Age at first scan. *Age@1stScan* = chronological age at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table S6-B

Johnson-Neyman Significance Regions for the Conditional Effect of Age at First Scan on Change in Left Amygdala Volumes across Time in Study in Boys

Age at 1st Scan	Slope	SE	LCI	UCI	*t*	*p*
20.00	-8.05	10.30	-28.24	12.14	-0.78	.436
19.00	-5.66	9.29	-23.87	12.55	-0.61	.544
18.00	-3.27	8.31	-19.55	13.02	-0.39	.695
17.00	-0.88	7.36	-15.30	13.55	-0.12	.905
16.00	1.51	6.46	-11.16	14.18	0.23	.815
15.00	3.90	5.65	-7.17	14.97	0.69	.491
14.00	6.29	4.94	-3.40	15.98	1.27	.205
13.00	8.68	4.41	0.03	17.33	1.97	.051
12.00	11.07	4.12	3.00	19.14	2.69	.008
11.00	13.46	4.11	5.40	21.52	3.27	.001
10.00	15.85	4.40	7.23	24.47	3.60	< .001
9.00	18.24	4.93	8.59	27.89	3.70	< .001
8.00	20.63	5.62	9.61	31.65	3.67	< .001
7.00	23.02	6.44	10.40	35.64	3.57	< .001
6.00	25.41	7.33	11.04	39.78	3.47	< .001
5.00	27.80	8.28	11.57	44.03	3.36	.001

Note. LCI = 95% lower confidence interval. UCI = 95% upper confidence interval.

Table S7-A

Change in Left Amygdala Volumes as a Function of Time and Age at Initial Scan among Girls

	Right Amygdala Volume (mm^3)							
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters								
Intercept	1514.17***	11.60	1505.63***	12.07	1472.88***	37.63	1444.31***	38.93
Time			7.89*	3.08	7.95*	3.11	35.67***	10.03
Age@1stscan					2.90	3.16	5.43	3.28
Time x Age@1stScan							-2.45**	0.84
Random Effects Parameters								
σ^2	6371.956		6161.550		6057.982		5749.073	
τ_{00}	21114.662		21209.119		21081.373		21324.370	
n	184		184		184		184	
Scans	371		371		371		371	
Deviance	4666.733		4660.258		4659.428		4651.250	

Note. Time = Age at current scan – Age at first scan. *Age@1stScan* = chronological age at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Table S7-B

Johnson-Neyman Significance Regions for the Conditional Effect of Age at First Scan on Change in Left Amygdala Volumes across Time in Study in Girls

Age at 1st Scan	Slope	SE	LCI	UCI	*t*	*p*
20.00	-13.02	7.83	-28.36	2.32	-1.66	.098
19.00	-10.61	7.06	-24.45	3.24	-1.50	.135
18.00	-8.20	6.32	-20.59	4.19	-1.30	.196
17.00	-5.79	5.60	-16.77	5.19	-1.03	.303
16.00	-3.38	4.92	-13.02	6.26	-0.69	.493
15.00	-0.97	4.29	-9.38	7.44	-0.23	.822
14.00	1.44	3.74	-5.90	8.78	0.38	.701
13.00	3.85	3.32	-2.65	10.35	1.16	.247
12.00	6.26	3.06	0.26	12.26	2.05	.042
11.00	8.67	3.02	2.75	14.59	2.87	.005
10.00	11.08	3.20	4.80	17.36	3.46	.001
9.00	13.49	3.58	6.48	20.50	3.77	< .001
8.00	15.90	4.09	7.89	23.91	3.89	< .001
7.00	18.31	4.69	9.11	27.51	3.90	< .001
6.00	20.72	5.36	10.22	31.22	3.87	< .001
5.00	23.13	6.07	11.24	35.02	3.81	< .001

Note. LCI = 95% lower confidence interval. UCI = 95% upper confidence interval.

Table S8-A

Change in Right Amygdala Volumes as a Function of Pubertal Development and Tanner Stage at Initial Scan in Boys

	Right Amygdala Volume (mm^3)							
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters								
Intercept	1741.49***	13.94	1743.06***	15.91	1655.02***	30.69	1657.95***	30.92
TSChange			33.84***	8.16	35.21***	8.13	18.15	18.07
TS@1stScan					35.76**	10.82	34.15**	10.97
TSChange x TS@1stScan							10.48	9.91
Random Effects Parameters								
σ^2	7815.889		5914.187		5908.070		5774.728	
τ_{00}	23328.475		25261.374		22816.602		23100.023	
n	146		119		119		119	
Scans	266		204		204		204	
Deviance	3401.317		2594.028		2583.564		2582.470	

Note. TSChange = Tanner stage at current scan – Tanner stage at first scan. *TS@1stScan* = pubertal development (Tanner Stage) at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Table S8-B

Change in Right Amygdala Volumes as a Function of Pubertal Development and Tanner Stage at Initial Scan in Girls

	Right Amygdala Volume (mm^3)							
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters								
Intercept	1553.60***	12.68	1556.58***	14.23	1535.29***	31.43	1530.27***	31.68
TSChange			36.33***	9.02	36.99***	9.06	69.18**	20.72
TS@1stScan					8.15	10.73	10.75	10.89
TSChange x TS@1stScan							-18.71	10.83
Random Effects Parameters								
σ^2	7864.698		7489.697		7495.301		7290.921	
τ_{00}	25062.644		25669.875		25533.029		25913.427	
n	184		155		155		155	
Scans	371		309		309		309	
Deviance	4738.711		3942.342		3941.767		3938.850	

Note. TSChange = Tanner stage at current scan – Tanner stage at first scan. *TS@1stScan* = pubertal development (Tanner Stage) at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Table S9-A

Change in Left Amygdala Volumes as a Function of Pubertal Development and Tanner Stage at Initial Scan in Boys

	Left Amygdala Volume (mm^3)							
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters								
Intercept	1664.49***	13.89	1664.98***	16.28	1637.06***	32.40	1638.04***	32.50
TSChange			19.69*	8.17	25.91*	11.17	16.42	24.66
TS@1stScan					11.14	11.47	10.59	11.54
TSChange x TS@1stScan							5.18	11.66
Random Effects Parameters								
σ^2	6163.180		5911.114		4383.390		4320.865	
τ_{00}	24153.246		26705.686		27059.893		27149.292	
n	146		119		119		119	
Scans	266		204		204		204	
Deviance	3371.547		2599.696		2592.273		2592.086	

Note. TSChange = Tanner stage at current scan – Tanner stage at first scan. *TS@1stScan* = pubertal development (Tanner Stage) at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance.
*$p < .05$. **$p < .01$. ***$p < .001$

Table S9-B

Change in Left Amygdala Volumes as a Function of Pubertal Development and Tanner Stage at Initial Scan in Girls

	Left Amygdala Volume (mm^3)							
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters								
Intercept	1514.17***	11.60	1515.40***	13.00	1514.61***	28.43	1513.20***	28.57
TSChange			14.40	8.03	18.28	9.53	29.55	22.02
TS@1stScan					0.05	9.73	0.75	9.82
TSChange x TS@1stScan							-6.06	10.76
Random Effects Parameters								
σ^2	6371.956		5912.222		5309.455		5256.545	
τ_{00}	21114.662		21683.511		21450.353		21548.189	
n	184		155		155		155	
Scans	371		309		309		309	
Deviance	4666.733		3878.266		3875.318		3875.010	

Note. TSChange = Tanner stage at current scan – Tanner stage at first scan. *TS@1stScan* = pubertal development (Tanner Stage) at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Table S10-A

Change in Right Amygdala Volumes as a Function of Chronological Time and Pubertal Maturation in Boys

	Right Amygdale Volume (mm^3)									
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters										
Intercept	1741.49***	13.94	1741.53***	16.15	1587.09***	69.52	1551.05***	70.57	1558.84***	70.34
Time			2.96	6.78	1.75	6.81	13.18	29.69	-4.06	31.17
TSChange			45.24*	20.61	46.79*	20.65	177.19*	76.41	-41.24	132.99
Time x TSChange			-4.76	6.13	-4.63	6.12	-3.56	6.28	76.64	41.28
Age@1stScan					9.40	8.75	13.91	8.85	13.10	8.83
TS@1stScan					16.23	21.17	8.43	21.37	9.17	21.31
Time x Age@1stScan							-0.72	4.17	2.07	4.50
Time x TS@1stScan							-3.81	9.84	-9.38	10.57
TSChange x Age@1stScan							-17.20*	7.70	0.10	12.63
TSChange x TS@1stScan							49.87**	17.36	54.07	30.63
Time x TSChange x Age1stScan									-6.81	4.02
Time x TSChange x TS@1stScan									0.34	9.10
Random Effects Parameters										
σ^2	7815.889		5849.899		5850.789		4911.845		4740.756	
τ_{00}	23328.475		25376.219		22658.876		23209.086		23066.451	
n	146		119		119		119		119	
Scans	266		204		204		204		204	
Deviance	3401.317		2593.385		2581.860		2566.562		2562.377	

Note. Time = Age at current scan – Age at first scan; *TSChange* = Tanner stage at current scan – Tanner stage at first scan; *Age@1stScan* = chronological age at initial assessment/study entry; *TS@1stScan* = pubertal development (Tanner stage) at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Table S10-B

Change in Right Amygdala Volumes as a Function of Chronological Time and Pubertal Maturation in Girls

	Right Amygdale Volume (mm^3)									
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters										
Intercept	1553.60***	12.68	1547.10***	14.51	1413.98***	54.67	1357.85***	56.47	1359.01***	56.66
Time			16.32**	5.40	16.92**	5.43	45.18*	17.56	44.88*	18.65
TSChange			58.46*	24.53	56.02*	24.55	176.59**	57.80	148.03	101.32
Time x TSChange			-15.64*	7.27	-15.67*	7.27	-24.17**	7.54	-15.10	30.06
Age@1stScan					20.48**	7.71	25.93**	7.92	25.94**	7.95
TS@1stScan					-42.45*	20.79	-45.91*	21.16	-46.51*	21.22
Time x Age@1stScan							0.35	2.79	0.21	3.17
Time x TS@1stScan							-9.44	8.51	-8.56	9.52
TSChange x Age@1stScan							-11.70	6.80	-12.39	11.68
TSChange x TS@1stScan							5.55	20.56	24.81	37.89
Time x TSChange x Age1stScan									0.30	3.71
Time x TSChange x TS@1stScan									-6.75	11.92
Random Effects Parameters										
σ^2	7864.698		6994.688		6985.479		6006.507		5985.740	
τ_{00}	25062.644		25966.162		24657.478		25368.308		25368.717	
n	184		155		155		155		155	
Scans	371		309		309		309		309	
Deviance	4738.711		3931.837		3924.711		3902.175		3901.578	

Note. Time = Age at current scan – Age at first scan; *TSChange* = Tanner stage at current scan – Tanner stage at first scan; *Age@1stScan* = chronological age at initial assessment/study entry; *TS@1stScan* = pubertal development (Tanner stage) at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Table S11-A

Change in Left Amygdala Volumes as a Function of Chronological Time and Pubertal Maturation in Boys

	Right Amygdale Volume (mm^3)									
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters										
Intercept	1664.49***	13.89	1663.31***	16.51	1510.18***	72.07	1493.20***	72.63	1491.02***	73.01
Time			4.38	6.81	4.54	6.92	39.14	31.68	43.78	31.27
TSChange			16.64	20.68	12.67	19.90	46.76	88.54	4.03	135.73
Time x TSChange			-0.85	6.14	1.75	5.73	5.03	6.23	19.65	41.94
Age@1stScan					17.65	9.06	19.84*	9.11	20.44*	9.16
TS@1stScan					-25.60	21.88	-29.57	22.00	-32.08	22.13
Time x Age@1stScan							-5.02	4.50	-6.00	4.58
Time x TS@1stScan							8.14	10.60	11.75	10.76
TSChange x Age@1stScan							-5.28	9.01	-8.70	13.15
TSChange x TS@1stScan							19.65	20.75	72.52*	30.95
Time x TSChange x Age1stScan									1.40	4.08
Time x TSChange x TS@1stScan									-19.66*	8.86
Random Effects Parameters										
σ^2	6163.180		5874.788		3816.389		3621.445		3545.318	
τ_{00}	24153.246		26758.551		25895.853		26052.961		26361.082	
n	146		119		119		119		119	
Scans	266		204		204		204		204	
Deviance	3371.547		2599.279		2586.656		2581.305		2575.373	

Note. Time = Age at current scan – Age at first scan; *TSChange* = Tanner stage at current scan – Tanner stage at first scan; *Age@1stScan* = chronological age at initial assessment/study entry; *TS@1stScan* = pubertal development (Tanner stage) at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Table S11-B

Change in Left Amygdala Volumes as a Function of Chronological Time and Pubertal Maturation in Girls

	Right Amygdale Volume (mm^3)									
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects Parameters										
Intercept	1514.17***	11.60	1514.70***	13.29	1434.67***	50.16	1413.87***	51.72	1413.71***	51.94
Time			1.68	4.96	0.90	4.85	1.30	16.59	0.79	17.70
TSChange			6.86	22.52	4.37	22.92	108.89	58.57	116.82	97.79
Time x TSChange			1.39	6.68	3.64	6.72	1.24	7.30	-1.03	28.75
Age@1stScan					13.64	7.07	16.32*	7.26	16.25*	7.29
TS@1stScan					-31.58	19.02	-35.84	19.38	-35.36	19.46
Time x Age@1stScan							0.87	2.64	1.12	3.02
Time x TS@1stScan							-4.18	8.09	-5.17	9.10
TSChange x Age@1stScan							-14.69	6.96	-13.25	11.29
TSChange x TS@1stScan							32.94	20.87	19.96	36.89
Time x TSChange x Age1stScan									-0.59	3.55
Time x TSChange x TS@1stScan									4.75	11.52
Random Effects Parameters										
σ^2	6371.956		5895.303		5202.258		5247.934		5251.515	
τ_{00}	21114.662		21729.970		20941.317		21080.319		21108.822	
N	184		155		155		155		155	
Scans	371		309		309		309		309	
Deviance	4666.733		3878.049		3871.317		3864.423		3864.207	

Note. Time = Age at current scan – Age at first scan; *TSChange* = Tanner stage at current scan – Tanner stage at first scan; *Age@1stScan* = chronological age at initial assessment/study entry; *TS@1stScan* = pubertal development (Tanner stage) at initial assessment/study entry. σ^2 = Random effects variance. τ_{00} = Between-subjects variance. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$