



Neural correlates of self-evaluation in relation to age and pubertal development in early adolescent girls

Marjolein E.A. Barendse^{a,*}, Danielle Cosme^a, John C. Flounoy^{a,b}, Nandita Vijayakumar^{a,c}, Theresa W. Cheng^a, Nicholas B. Allen^{a,d}, Jennifer H. Pfeifer^a

^a Department of Psychology, University of Oregon, Eugene, OR, USA

^b Department of Psychology, Harvard University, Cambridge, MA, USA

^c School of Psychology, Deakin University, Burwood, VIC, Australia

^d Melbourne School of Psychological Sciences, The University of Melbourne, Parkville, VIC, Australia

ARTICLE INFO

Keywords:
Adolescence
Puberty
Hormones
Self-concept
fMRI
Activation

ABSTRACT

Early adolescence is marked by puberty, and is also a time of flux in self-perception. However, there is limited research on the neural correlates of self-evaluation in relation to pubertal development. The current study examined relationships between neural activation during self-evaluation of social traits and maturation (age and pubertal development) in a community sample of female adolescents. Participants ($N = 143$; age $M = 11.65$, range = 10.0–13.0) completed a functional MRI task in which they judged the self-descriptiveness of adjectives for prosocial, antisocial and social status-related traits. Pubertal development was based on self-report, and was also examined using morning salivary testosterone, dehydroepiandrosterone, and estradiol. Contrary to pre-registered hypotheses, neither age nor pubertal development were related to neural activation during self-evaluation. We further examined whether activation in two regions-of-interest, the ventromedial prefrontal cortex (vmPFC) and perigenual anterior cingulate (pgACC), was associated with trial-level self-evaluative behavior. In line with preregistered hypotheses, higher vmPFC and pgACC activation during self-evaluation were both associated with a higher probability of endorsing negative adjectives, and a lower probability of endorsing positive adjectives. Future studies should examine neural trajectories of self-evaluation longitudinally, and investigate the predictive value of the neural correlates of self-evaluation for adolescent mental health.

1. Introduction

Forming a clear and multifaceted concept of the self is an important goal in adolescence. Self-concepts are defined in the developmental literature as self-perceptions pertaining to specific domains, roles, or social contexts (Harter, 2012). Although there is already some specificity in self-concepts earlier in childhood (Measelle et al., 1998), these self-perceptions become increasingly complex and differentiated across adolescence (Byrne and Shavelson, 1996; Harter, 2012), presumably from both cognitive development and social changes experienced during this period. However, these classic approaches to self-concept development have generally not considered neural and other biological contributions to adolescent changes in self-perception, which is the focus of the current study.

1.1. Maturation of neural correlates of self-evaluation with age

Adolescence is also a time of substantial brain development, and recent studies have shown that the neural correlates of self-perception change during this developmental period (Pfeifer et al., 2007, 2009, 2013; van der Crujisen et al., 2018). Across the lifespan, evaluating whether traits are self-descriptive or not robustly activates midline brain regions, such as the medial prefrontal cortex (mPFC), perigenual anterior cingulate cortex (pgACC), medial posterior parietal cortex, and precuneus (Northoff et al., 2006; van der Meer et al., 2010). Other regions supporting social cognition and social reward, such as temporoparietal junction (TPJ) and ventral striatum (VS), are also frequently recruited during self-evaluation, particularly by adolescents or in the social domain (Jankowski et al., 2014; Pfeifer et al., 2013; Romund et al., 2017). However, the specific developmental trajectories in engagement of these regions is unclear. Several studies reported positive

* Corresponding author at: Department of Psychology, 1227 University of Oregon, Eugene, OR 97403, United States.

E-mail address: barendse@uoregon.edu (M.E.A. Barendse).

<https://doi.org/10.1016/j.dcn.2020.100799>

Received 19 September 2019; Received in revised form 13 April 2020; Accepted 15 May 2020

Available online 20 May 2020

1878-9293/© 2020 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

associations between age and activation in the ventral mPFC (vmPFC) and pgACC (Dégeilh et al., 2015; Pfeifer et al., 2007, 2009, 2013; van der Crujnsen et al., 2018), while others have reported null findings (Jankowski et al., 2014). Cross-sectional studies comparing adolescents to adults have reported stronger activation in the dorsal mPFC (dmPFC) and TPJ during self-evaluation for early adolescents (Pfeifer et al., 2009), but greater activation in VS for adults when making self-evaluations in the social domain (Jankowski et al., 2014).

1.2. Puberty and the neural correlates of self-evaluation

The studies described above focused on age-related developmental changes. However, in adolescence, especially in early adolescence, puberty might be a better indicator of maturation than age (Dahl et al., 2018). Physical maturation and hormonal changes occurring with puberty have been associated with many aspects of brain development, including structural development in regions and networks related to self-referential processing (for a review, see Vijayakumar et al., 2018). Despite these relationships, few studies have examined the relationship between pubertal development and the neural correlates of self-evaluation, including whether pubertal processes relate to self-evaluative neural processes over and above age. Studies examining these relationships have found that more advanced pubertal development was associated with increased activation in the vmPFC during self-evaluation of social traits (after correcting for age) (Pfeifer et al., 2013), and increased VS activation when evaluating one's social self-concept from the perspective of a close friend (Jankowski et al., 2014). However, both of these findings were based on small samples ($N = 27$ and 18 respectively) and relied on self-reported pubertal stage. Including both hormonal and self-report measures of pubertal development could help to distinguish pubertal effects on the neural correlates of self-evaluation that are suggestive of more direct biological effects (e.g., hormones triggering receptors or affecting gene expression of neurons in networks supporting self-evaluative processes), or more indirect social effects (e.g., altered physical appearance changing impressions by or interactions with other people, which in turn shape self-perceptions), or both. Therefore, we aimed to examine how age, self-reported pubertal development and pubertal hormone levels are associated with neural activation during self-evaluation.

1.3. The role of valence in self-evaluation

Activation of the vmPFC and pgACC during self-evaluation might also vary by valence of the evaluated trait. In adolescents and young adults, higher activation has been reported in the vmPFC (van der Crujnsen et al., 2018) and ACC (subgenual ACC in Moran et al., 2006, pgACC in van der Crujnsen et al., 2018) during self-evaluation of positive (or favorable) traits compared to negative (or unfavorable) traits. In addition, multi-voxel activation patterns in the vmPFC during self-evaluation are similar to those during viewing of positive affective stimuli (Chavez et al., 2017). However, these findings ignore participants' self-descriptiveness of the traits: endorsing a positive trait might elicit different brain activation compared to rejecting the same trait. Recent research from our lab with late adolescents suggests that the effect of valence depends on whether or not the trait is endorsed. This study demonstrated that greater vmPFC, and potentially pgACC, activation is associated with an increased likelihood of endorsing negative self-evaluative statements (Cosme et al., 2019). This finding was based on a trial-level analysis, thus including within-person variation, showing that when participants had relatively high vmPFC activation in response to a negative item, such as 'lonely' or 'anxious', they were more likely to endorse this item as self-descriptive compared to when they showed lower vmPFC activation. Activation in the pgACC showed a similar pattern, but was not a significant predictor when including both vmPFC and pgACC response in the same model. However, it is unclear whether this relationship is the same in early adolescence. If there are

developmental effects in these areas, and their activation is related to responses during self-evaluation, they might be candidate neural markers of developmental changes in self-concept. Therefore, another aim of the current study was to examine the degree to which vmPFC and pgACC activation was associated with self-descriptiveness of evaluated traits, and whether that differed as a function of the valence of the adjective.

1.4. Study goals and pre-registered hypotheses

The overarching goal of the current study was to investigate the neural correlates of self-evaluation in early adolescent females. We focused on female adolescents mainly because an important part of the analyses includes pubertal processes, which differ vastly between the sexes. The broader project this study is embedded in was limited to girls for the same reason, as well as because of its focus on mental health problems that girls become increasingly at risk for across adolescence (Barendse et al., 2020). The choice to examine only female adolescents also resulted in more power than a mixed sample in which sex is a moderator or both sexes are examined separately. Further, we emphasized the social self, since social relationships become increasingly important during adolescence, and understanding one's attributes and qualities in interactions with others is crucial to managing these relationships. In addition, previous studies have found that puberty is specifically related to the neural correlates of social self-evaluations (Jankowski et al., 2014; Pfeifer et al., 2013). However, previous research has often focused on one component of social behavior and functioning, such as prosocial behavior (van der Crujnsen et al., 2018) or popularity (Pfeifer et al., 2007, 2013). The current study aimed to capture a broader range of social traits, divided into three main adjective types: prosociality, antisociality/aggressiveness, and social status/sociability (examples are 'considerate', 'selfish', and 'outgoing', respectively; see *Methods* for details).

The preregistered (<https://osf.io/n268d/registrations>) hypotheses were as follows:

1.4.1. Age associations

We aimed to determine how age is associated with neural activation during self-evaluation. Based on previous research (Dégeilh et al., 2015; Pfeifer et al., 2007, 2009, 2013; van der Crujnsen et al., 2018), we expected that activation in the ventromedial prefrontal cortex (vmPFC) and adjacent perigenual anterior cingulate cortex (pgACC) during self-evaluation (relative to the control condition) would increase with age. We also examined three other regions of interest (ROIs): the ventral striatum (VS), temporoparietal junction (TPJ), and dorsomedial prefrontal cortex (dmPFC), but without directional hypotheses due to the limited literature examining or reporting age-related activation in these areas (Jankowski et al., 2014; Pfeifer et al., 2009).

1.4.2. Puberty associations

We aimed to determine how pubertal development, as indexed by secondary sexual characteristics as well as hormone levels, is associated with neural activation during self-evaluation. Based on the literature described in Section 1.2 (Jankowski et al., 2014; Pfeifer et al., 2013), we hypothesized that self-reported pubertal development, as well as levels of DHEA, testosterone, and estradiol, would explain neural activation during self-evaluation in areas subserving self-referential, affective and reward processes (including the five *a priori* ROIs listed in hypothesis 1) over and above age (in other words, they will be a significant predictor after covarying for age).

1.4.3. Interactions between puberty and adjective type

The aim was to establish whether adjective type interacts with pubertal development, as indexed by secondary sexual characteristics as well as hormone levels, in its association with neural activation during self-evaluation. Given previous findings showing that

neurodevelopmental effects of self-evaluation differed as a function of trait type (Pfeifer et al., 2013; van der Crujisen et al., 2018), we hypothesized that adjective type would moderate the relationship between puberty and neural activation. Specifically, we expected that participants with more advanced self-reported pubertal development would show greater vmPFC activation during self-evaluation of adjectives related to both social status/sociability and antisociality/aggressiveness, but not prosociality. Although no study has examined neural activation during self-evaluation of prosocial traits in relation to pubertal development, previous research has failed to find associations between age and vmPFC activation while evaluating prosocial traits (van der Crujisen et al., 2018). We also expected that participants with higher testosterone would show increased vmPFC activation to antisocial/aggressive adjectives, because of reported links between testosterone levels and dominance and approach motivation (Montoya et al., 2012; Vermeersch et al., 2010).

1.4.4. Valence effects and association with endorsement

The final aims were to examine the main effect of valence on neural activation in the vmPFC, pgACC and VS, and to determine to which extent vmPFC and pgACC activation was associated with self-descriptiveness of evaluated traits, depending on valence of the adjective. Based on the literature described in Section 1.3 (Chavez et al., 2017; Moran et al., 2006; van der Crujisen et al., 2018), we hypothesized that activation in the vmPFC, pgACC and VS would be higher when evaluating positive compared to negative adjectives. The VS was included here because of its known role in processing of rewards and other positive outcomes (Bartra et al., 2013), as well as in self-evaluation (Jankowski et al., 2014; Pfeifer et al., 2013; Romund et al., 2017). Further, based on previous findings in late adolescents (Cosme et al., 2019), we predicted that vmPFC and pgACC activation would be differentially associated with endorsement according to valence, such that, at a trial-level, stronger engagement of vmPFC and pgACC would be associated with a higher probability of endorsing negatively valenced traits as self-descriptive.

2. Methods

2.1. Participants

Female adolescents aged 10.0–13.0 years were recruited from the community, together with one of their parents/guardians, to participate in a larger project called Transitions in Adolescent Girls¹ (Barendse et al., 2020). Families were recruited primarily through recruitment letters distributed by schools in the greater Eugene and Springfield area (Oregon, USA), and to a minimal extent through secure databases of people who registered their interest in our lab's or department's research, recruitment flyers posted around the community or disseminated at community events, and snowballing efforts. Participants were required to be fluent in English and to have normal or corrected-to-normal vision. Exclusion criteria included a diagnosis of a developmental disability, psychotic disorder, or behavioral disorder (including autism); current use of psychotropic medication; MRI contraindications; and the participant reporting or suspecting to be pregnant. Parents or guardians gave written informed consent and adolescents gave assent to participate. Ethics approval was granted by the University of Oregon Institutional Review Board.

In total, 174 participants were included in the study. See Barendse and colleagues (2020) for demographic information of the sample. Out of these 174, 14 did not complete the self-evaluation fMRI scan and 12

were excluded from all analyses because of motion artefacts in the fMRI data (see *Processing of fMRI data*), leading to a final sample of $N = 148$. Another five participants were excluded from the analyses of hypotheses 1, 2, and 3 because they had only one run of complete fMRI data, leading to a sample size of 143 for these analyses (age $M = 11.63$, $SD = 0.82$). These five participants could be included in analyses for hypothesis 4 because those analyses were done on a trial level.

2.2. Self-reported pubertal development

Participants completed the Pubertal Development Scale as a measure of physical pubertal development (Petersen et al., 1988). This scale contains six items, including questions on height growth, skin changes, breast development, body hair growth, and menarche. Following customary practices in the field, the question regarding social comparison of pubertal timing was not used in the current manuscript. Scores on the PDS were converted into Tanner stages, ranging from one (prepubertal) to five (postpubertal), according to the syntax developed by Shirtcliff et al. (2009).

2.3. Saliva collection and hormone measurement

Participants collected four 2 ml saliva samples at home. Each sample was provided through passive drool in the morning, directly after waking, one week apart. This was done to provide a more stable estimate of basal hormone levels, given that they fluctuate over the day and over the menstrual cycle. Participants were instructed not to eat or brush their teeth before collecting the sample. Families stored the samples in their home freezer until the day of the scan, when they brought it to the lab on ice in a cooler bag. Participants and their parents or guardians were trained on how to collect and store the samples. They recorded the time of day at collection and reported illnesses and medication use during the month of saliva collection. Saliva samples were stored in a -80°C freezer in the lab until they were shipped (overnight on dry ice) to the Stress Physiology Investigative Team at Iowa State University, where they were assayed in duplicate for dehydroepiandrosterone (DHEA), testosterone (T), and estradiol (E2) using Enzyme-Linked Immunosorbent Assay (ELISA) (www.salimetrics.com). The intra-assay coefficients of variation (CVs) were 10.48 % for DHEA, 1.80 % for T, and 7.76 % for E2. We processed the samples in two batches. The inter-assay CVs for Batch 1 were 20.62 % for DHEA, 10.23 % for T, and 11.53 % for E2, and for Batch 2 were 21.43 % for DHEA, 8.34 % for T, and 15.55 % for E2. All CVs reported are for the optical density wavelengths.

2.4. Processing of pubertal data

Any hormone samples collected >45 min after waking were excluded. Mean salivary hormone concentrations that were non-detectable and too low (i.e., left-censored) or too high (i.e., right censored) were substituted by the lower limit of sensitivity for that assay or the upper limit of the standards for that assay, respectively, if other samples of that participant showed similarly low or high levels. Otherwise, that concentration was excluded from further analyses. For exact hormone cleaning protocols, see Barendse and colleagues (2020). Remaining hormone levels were averaged across the samples. The levels were highly correlated across the different weeks (DHEA $r = .81-.86$; T $r = .59-.78$; E2 $r = .55-.72$). The averaged variables were log-transformed to achieve an approximately Gaussian distribution. Outliers on the hormone variables after transformation ($-3 > z > 3$) were winsorized to 1 % above the next highest or below the next lowest value ($n = 2$ for DHEA, $n = 0$ for T and $n = 3$ for E2).

Missing data on the averaged hormone levels ($n = 11$) as well as on Tanner stage ($n = 13$) could be assumed to be missing completely at random ($X^2(10) = 12.84$, $p = .23$) and were imputed using multiple imputation with the *mice* package in R version 3.4.2 (van Buuren and Karin, 2011), taking the averaged imputed value over 50 imputations.

¹ Although the study was advertised using the name Transitions in Adolescent Girls, during the first wave of data collection 1.7% of participants reported a non-binary gender identity. Thus while the title and abstract refer to 'girls' and 'female', we acknowledge that not all participants identify as such.

2.5. fMRI self-evaluation task

The self-evaluation fMRI paradigm was based on our previous research on self-evaluation in adolescents (Jankowski et al., 2014; Pfeifer et al., 2007, 2013). In this paradigm adolescents were presented with 50 individual trait adjectives that are relevant to interpersonal relationships (e.g., friendly, respectful, popular, awkward, selfish, and aggressive). In the self-evaluation condition, participants reported whether a given trait describes them (yes/no), and in the “change” condition participants reported whether they believe the trait is something that can change about people in general (i.e., is malleable; yes/no). Every trait adjective was presented for 4.7 s. Participants could respond at any time by pressing a button on a button box and reaction times were recorded. The task used a mixed event-related design with alternating “self” and “change” blocks of five adjectives. The task was split into two runs; adjectives that were presented in the “change” condition in run 1 were presented in the self-evaluation condition in run 2 and vice versa. The order of adjectives was randomized.

2.5.1. Stimuli

The broad goal of the fMRI task was to engage adolescents in self-evaluation with respect to adjectives that are relevant to interpersonal relationships. The list of adjectives was generated as part of a search for adjectives that are associated with two types of popularity: prosocial (or sometimes called sociometric) popularity and populist (“everyone knows this person is popular, but they’re not necessarily well-liked”) popularity (de Bruyn and Cillessen, 2006). The final selection of adjectives was based on pilot data of 100 female participants (aged 18–25, native English speakers). Exploratory factor analysis on this data suggested a three-factor solution that was iteratively refined. The final factor solution comprised the following factors: ‘prosociality’ (e.g., nice, helpful; 18 items), ‘antisociality/aggressiveness’ (e.g., bossy, mean; 14 items), and ‘social status/sociability’ (e.g., popular, shy; 18 items). Although ‘prosociality’ and ‘antisociality/aggressiveness’ may sound like opposites along the same dimension, the factor analysis showed that these were better represented as separable dimensions or factors. The prosociality adjectives were all positively valenced, the anti-sociality/aggressiveness adjectives were negatively valenced, and the social status/sociability factor contained a mix of positive and negative adjectives. For more details on the item selection process, a visualisation of the task presentation, and the full list of items for each adjective type, see Barendse et al. (2020).

2.6. fMRI data acquisition

All scans were acquired using a Siemens Skyra 3.0 T scanner at the Lewis Center for Neuroimaging at the University of Oregon. Participants completed a mock scan prior to the MRI scan to familiarize them with the scanner. They received instructions for the task and practiced the task during the mock scan. Acquisition parameters were as follows: 2×180 volumes of 72 slices with 2 mm isometric voxels, TR = 2000 ms, TE = 25 ms, multiband acceleration factor = 3, in plane acceleration factor = 2, FOV = 208 mm, flip angle = 90° , duration = 6.5 min. per run. T1-weighted images were acquired as follows for co-registration: sagittal 3D MP-RAGE, 176 slices with 1 mm isometric voxels, FOV = 256 mm, TR = 2500 ms, TE = 3.41 ms, flip angle = 7° , TI = 1100 ms, matrix size = 256×256 , acceleration factor = 2.

2.7. Preprocessing of fMRI data

First, raw DICOM images were converted to NIFTI format with MRIConvert and organized into Brain Imaging Data Structure (BIDS) format (Gorgolewski et al., 2016). Then, fMRIPrep was used to preprocess the fMRI data (v1.0.0; <https://github.com/poldracklab/fmriprep>, (Esteban et al., 2019)). This included correcting the T1-weighted images for intensity non-uniformity using N4BiasFieldCorrection

(v2.1.0) and skull-stripping it using AntsBrainExtraction (v2.1.0). Brain surfaces were reconstructed from the T1-weighted volumes using recon-all (FreeSurfer v6.0.1) and this was used to refine the brain mask estimated in the previous step. FSL’s fast (FSL v5.0.9) was applied to perform brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) on the brain-extracted T1-weighted images. Finally, spatial normalization to MNI-space (ICBM 152 Nonlinear Asymmetrical template version 2009c) was performed through nonlinear registration with the antsRegistration tool of ANTs v2.1.0, using brain-extracted versions of both T1w volume and template.

Functional data were motion corrected with mcflirt (FSL v5.0.9). Distortion correction was performed using field maps processed with fugue (FSL v5.0.9). Next, co-registration to the corresponding T1-weighted image was performed using boundary-based registration with six degrees of freedom (bbrregister in FreeSurfer v6.0.1). Motion correcting transformations, field distortion correcting warp, BOLD-to-T1w transformation and T1w-to-MNI template warp were concatenated and applied in a single step using antsApplyTransforms (ANTs v2.1.0) with Lanczos interpolation. For statistical analyses in MNI space, preprocessed functional data were smoothed using a 6 mm full-width at half maximum (FWHM) smoothing kernel; for analyses in native space, images were smoothed with a 2 mm FWHM smoothing kernel.

Excessive motion was identified using an automated motion assessment tool (Cosme et al., 2018; <https://github.com/dcosme/auto-motion-fmriprep>). This tool is a trained classifier that utilizes the motion confound files generated by fMRIPrep and classifies whether or not fMRI volumes contain motion artefacts. The classifier was applied to each subject task run and returned a binary classification indicating the presence or absence of motion artefacts for each volume. Participants whose scans contained >20 % of volumes classified as containing a motion artefact across the two task runs, were excluded from further analyses (n = 12).

2.8. Defining ROIs

Bilateral ROIs were created of the perigenual anterior cingulate cortex (pgACC), ventromedial prefrontal cortex (vmPFC), ventral striatum (VS), temporoparietal junction (TPJ), and dorsomedial prefrontal cortex (dmPFC). The VS ROI was created by combining the nucleus accumbens and putamen segmentations of the FreeSurfer segmentation atlas (Fischl, 2012). All other ROIs were based on the HCP MMP 1.0 cortical parcellation atlas (Glasser et al., 2016) (p32_ROI, 10r_ROI, and 10v_ROI parcels for vmPFC; a24_ROI, 25_ROI, and s32_ROI parcels for pgACC; 10d_ROI, 9m_ROI, and d32_ROI for dmPFC; TPOJ1_ROI, and TPOJ2_ROI for TPJ). For hypothesis 4, the ROIs were mapped onto participants’ T1-weighted scans using FreeSurfer 6 and concatenated and binarized with fslmaths. Since analyses for hypotheses 1 and 2 were between subject comparisons, ROIs for these analyses were defined in MNI space, using The HCP MMP 1.0 parcellation in volumetric space (<https://neurovault.org/images/24150/>), see Supplementary Fig. 1 for a visualization.

2.9. Analyses of hypotheses 1–3

Subject-level models were estimated in SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK, <http://www.fil.ion.ucl.ac.uk/spm>). For scripts, see https://github.com/dsnlab/TAG_scripts/tree/53f56675e55790c02f614704328e7ccc9e10cde3. These event-related general linear models used a canonical hemodynamic response function, high-pass filtering of 100 s, and the FAST algorithm for auto-correlation modeling. The models for hypotheses 1–3 included two conditions (self and change) and three adjective types (prosociality, antisociality/aggressiveness, and social status/sociability). Trials with missing responses were modeled as a separate regressor of no interest. Motion parameters (Euclidean distance, Euclidean rotation, the first

derivatives of both, and the ‘trash’ regressor from the automated motion classifier described above) were added to these subject level models as regressors of no interest.

For hypothesis 1 and 2, a contrast of the self > change condition was computed and then correlated with age and pubertal development, respectively, using linear regression models in SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK, <http://www.fil.ion.ucl.ac.uk/spm>), thresholded at cluster family-wise error (FWE) corrected $p = .05$ (with a voxel-wise threshold of .001). Age was included as a covariate in the models for hypothesis 2. In addition, ROI analyses were conducted with the anatomically defined ROIs described above (pgACC, vmPFC, VS, TPJ, and dmPFC). Mean activation across the ROI for the self > change contrast was extracted with Marsbar 0.44 (Brett et al., 2002) in SPM12 and related to age, Tanner stage, or hormone levels in R v. 3.4.2 using the same models and covariate as the whole brain analyses. A Bonferroni correction was applied to correct for the number of ROIs examined, accounting for the mean correlation between the ROIs (<http://www.quantitativeskills.com/sisa/calculations/bonfer.htm>). The mean correlation between the five ROIs was 0.69, leading to an adjusted p-threshold of 0.0303.

For hypothesis 3, we conducted an ANCOVA analysis in AFNI 18.2.04 (Cox, 1996), with 2 within-subject factors, i.e., condition (self versus change) and adjective type (prosociality, antisociality/aggressiveness, and social status/sociability), and two between-subject covariates, i.e., pubertal development (either Tanner stage or hormone level) and age (as a covariate of no interest). We used AFNI 3dFWHMx and 3dClustSim in accordance with recent guidelines (Cox et al., 2017) to determine the statistical significance and cluster-forming threshold for cluster FWE correction. An FWE-corrected p-threshold of 0.05 was reached by applying a voxel-wise threshold of .001, and cluster extent $k > 66$ (bisided correction and $nn = 3$; average ACF parameters: 0.622255, 4.33855, 9.6616).

In post-hoc (i.e., not preregistered) analyses, we explored if any interactions between pubertal development and adjective type were better explained by valence. For this purpose, we repeated the ANCOVA analyses for hypothesis 3 with valence as the moderator. Methods and results of this analysis are described in the Supplementary Materials.

2.10. Trial-level analyses (hypothesis 4)

Analyses for hypothesis 4 were conducted using multilevel models on a trial-by-trial basis. The analyses for this hypothesis focused on valence rather than adjective type because the previous research it is based on demonstrated the relevance of valence (Chavez et al., 2017; Cosme et al., 2019; van der Cruisjes et al., 2018). Event-related subject-level models were computed in native space in SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK, <http://www.fil.ion.ucl.ac.uk/spm>), modeling each trial as a separate regressor. All other settings remained the same as described above for the other hypotheses, and the same realignment parameters were added as regressors of no interest. Then, mean parameter estimates were extracted from the vmPFC, pgACC and VS ROIs for each trial for each participant using 3dmaskave in AFNI 18.2.04 (Cox, 1996). Only parameter estimates of trials in the self-evaluation condition were used in the regression models. Outliers (>3 standard deviations from the median of the ROI) on these were excluded.

2.10.1. Main valence effects

To examine whether activation in *a priori* ROIs (the vmPFC, pgACC and VS) differed as a function of valence, we computed a series of mixed effects regression models using the lme4 package in R (Bates et al., 2015). For each ROI, we regressed ROI activation on the fixed effect of valence and compare this model to a null model including the intercept only using chi-square difference tests. In all models, participant intercepts were modeled as random effects.

2.10.2. Interactions between valence and ROI activation in relation to endorsement

To assess the degree to which endorsement differed as a function of valence and activation in vmPFC and pgACC, we fit a series of multilevel logistic regression models using the lme4 package in R version 3.4.2 (Bates et al., 2015). Prior to modeling, parameter estimates were standardized within ROI and participant to correct for differences in variability between participants and ROIs. We then compared a set of three nested models with trial-level responses (yes or no) as the outcome variable. The base model included the fixed effect of valence (positive or negative) as the only regressor. The second model included additional regressors for the fixed effects of vmPFC and pgACC activation, and the third included the interactions between each ROI and valence. In all models, participant intercepts were modeled as random effects.

Since activation in the vmPFC and pgACC were highly correlated (repeated measures adjusted correlation $r = .84$), we examined potential multicollinearity in the models by calculating variance inflation factors (VIFs) for mixed effects models (<https://github.com/aufrank/R-hacks/blob/master/mer-utils.R>). Zuur et al. (2010) recommend dropping predictors or changing the model when the maximum VIF is above three. Because the maximum VIF in the third model (interaction model) was 7.7, we modeled each ROI separately, which resolved the multicollinearity issue (all VIFs dropped below three). For each ROI, we computed the three nested models described above and compared model fit to the previous, simpler model using chi-square difference tests. Model parameters from the best fitting models are presented.

2.10.3. Post-hoc analyses

In post-hoc analyses (i.e. analyses that were not pre-registered and therefore were not described in the hypotheses), we examined if age or pubertal development was associated with vmPFC and pgACC activation at the trial-level. To do this, we compared series of linear mixed effect models using lme4 in R 3.4.2 (Bates et al., 2015) with either trial-level vmPFC or pgACC activation as the outcome variable. The base model included a fixed effect of valence; in subsequent models we added, first, fixed effects of age and pubertal development (Tanner stage or hormone level) and, second, interactions between valence and age, and valence and puberty. All models included random intercepts nested within participants. Chi-square difference tests were used to compare model fit.

Additional post-hoc trial-level analyses were conducted to examine whether the effects of trial-level vmPFC or pgACC activation on responses could be explained by reaction time (RT). To test this, we added RT*valence fixed effects regressors to the vmPFC and pgACC models described above, to examine if this would change the significance of the interactive effect of ROI and valence on responses. We also examined whether RT was associated with responses and if this effect was moderated by valence. We ran logistic regression models using the lme4 package in R version 3.4.2 (Bates et al., 2015), and compared model fit between (1) a base model including a fixed effect of valence and a random intercept nested within participant, (2) a main effect model, which added a fixed effect for RT, and (3) an interaction model, which added the interaction between valence and RT.

Finally, to disentangle the effects of valence and adjective type on endorsement, we repeated the main valence by ROI activation analysis with adjective type as the moderator. Methods and results of this analysis can be found in the Supplementary Materials.

3. Results

3.1. Descriptive information and behavior

Table 1 shows descriptive information of the sample and of the main predictor variables for hypothesis 1 and 2. Average endorsement was higher for prosociality adjectives (94 %, SD = 10 %), compared to social status/sociability adjectives (39 %, SD = 12 %) and antisociality/aggressiveness adjectives (20 %, SD = 16 %). Separating the adjectives

Table 1
Descriptive statistics of maturational indices (N = 148).

	Mean	SD	Range
Age	11.63	0.82	10.03 – 13.17
Tanner stage	2.97	1.00	1 – 5
Testosterone (pg/ml)	42.16	22.88	11.19 – 150.80
DHEA (pg/ml)	110.60	125.57	1.13 – 998.32
Estradiol (pg/ml)	0.91	0.46	0.10 – 3.11

Note: DHEA = dehydroepiandrosterone, PDS = pubertal development scale, SD = standard deviation.

by valence showed that, on average, 83 % of positive adjectives were endorsed and 21 % of negative adjectives. Additionally, about 80 % of the traits were considered malleable, and this was similar across the three adjective types.

3.2. Main effects of the fMRI task

Self-evaluation (compared to the “change” condition) elicited activation in a number of mainly midline brain areas, including the medial PFC extending into ACC, the precuneus and medial parietal cortex, insula, thalamus, caudate and parts of the visual cortex (see Fig. 1). The opposite contrast (change > self) showed activation in lateral regions including the left and right lateral PFC, the left and right lateral OFC, the inferior parietal cortex bilaterally, and the middle temporal lobe (see Fig. 1).

3.3. Associations between maturational indices and brain activation during self-evaluation (hypotheses 1 and 2)

In order to test the first and second hypotheses, we related age, Tanner stage, DHEA levels, testosterone levels, and estradiol levels to whole brain activation during self-evaluation. Thresholded ($p_{fwe} = .05$) whole-brain analyses showed **no** associations between age and neural activation during self-evaluation. In addition, there were **no** significant

associations between Tanner stage, DHEA, testosterone, or estradiol (controlling for age) and neural activation during self-evaluation. Unthresholded maps are available online: <https://neurovault.org/collections/5703/>.

We complemented these whole-brain analyses by conducting ROI analyses in the vmPFC, pgACC, VS, TPJ, and dmPFC. The ROI analyses also showed no associations between age and activation in any of the ROIs. Similarly, there were no significant associations between Tanner stage, DHEA, testosterone, or estradiol (controlling for age) and neural activation in any of the ROIs. See Table 2 for correlations between age, pubertal variables and ROI activation; see Table 3 for detailed output of the linear regression models; and see Supplementary Fig. 4 for scatterplots of the correlations between Tanner stage and activation in each ROI.

3.4. Interactions between adjective type and pubertal development (hypothesis 3)

Whole-brain analyses examined if any associations between pubertal development (Tanner stage and levels of DHEA, testosterone, and estradiol) and neural activation during self-evaluation were moderated by adjective type. These analyses showed no interactions between Tanner stage or estradiol levels and adjective type on neural activation. However, participants with higher testosterone levels had stronger activation in the ventral part of the left and right precentral gyri when evaluating prosocial adjectives versus both other types of adjective (peak xyz = -66, -8, 18, $k_e = 95$, $F = 11.52$, $p_{fwe} < .05$; peak xyz = 56, -2, 22, $k_e = 367$, $F = 11.69$, $p_{fwe} < .05$; see Fig. 2). For DHEA, a similar interaction was found (peak xyz = -66, -18, 16, $k_e = 89$, $F = 11.12$, $p_{fwe} < .05$; peak xyz = 68, -12, 22, $k_e = 367$, $F = 12.99$, $p_{fwe} < .05$; see Fig. 2), but higher DHEA levels were related to more activation during prosocial versus antisocial/aggressive adjectives. These effects remained statistically significant after removing one participant with outlying values ($-3 > z > 3$; top left of scatterplots in Fig. 2).

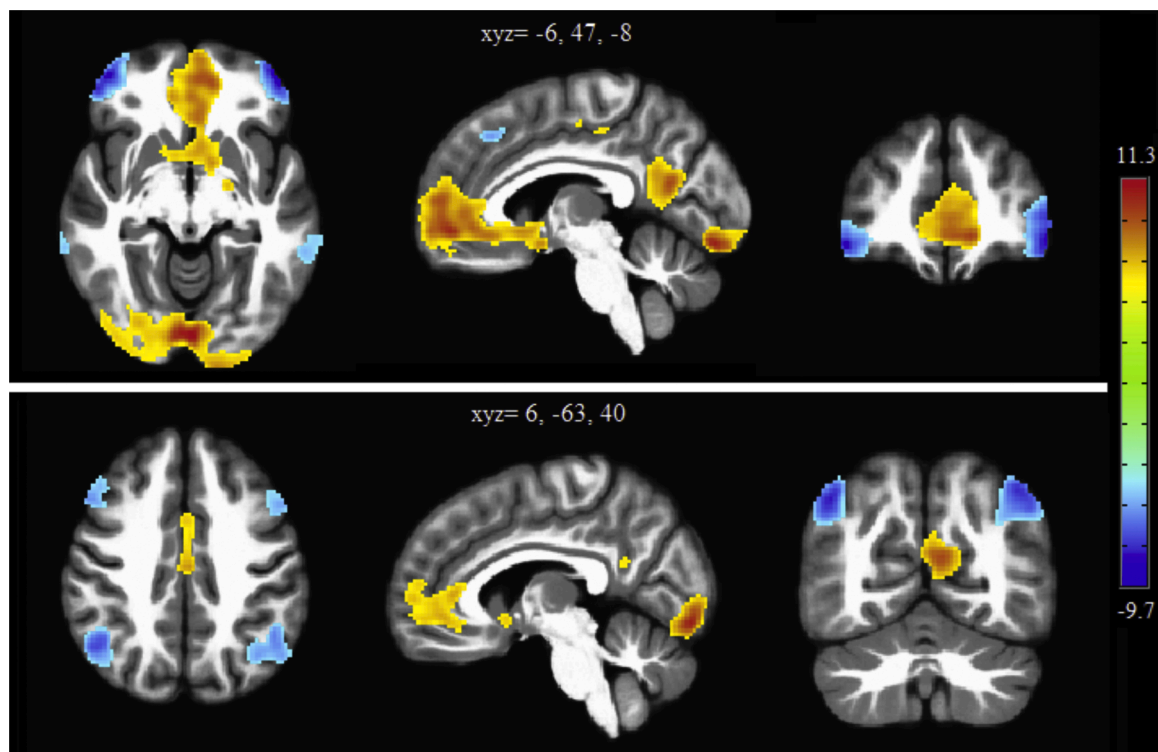


Fig. 1. Main effects of task, thresholded at voxelwise $p_{fwe} < .01$ to illustrate the core self-evaluation related areas. Positive (yellow-red) clusters are self > change and negative (blue) clusters are change > self. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 2
Correlations between age, pubertal development, and ROI activation.

	Age	Tanner stage	DHEA	Testosterone	Estradiol	vmPFC	pgACC	VS	TPJ
PDSS	0.40*								
DHEA	0.44*	0.47*							
Testosterone	0.45*	0.45*	0.80*						
Estradiol	0.41*	0.41*	0.54*	0.60*					
vmPFC	0.00	-0.16	-0.12	-0.07	-0.08				
pgACC	-0.02	-0.13	-0.08	-0.06	-0.09	0.91*			
VS	-0.05	-0.10	-0.12	-0.06	-0.11	0.68*	0.67*		
TPJ	0.05	0.07	0.02	0.03	-0.04	0.43*	0.45*	0.73*	
dmPFC	0.02	-0.14	-0.03	0.01	-0.05	0.91*	0.87*	0.72*	0.52*

Note: * $p < .01$; DHEA = dehydroepiandrosterone, dmPFC = dorsomedial prefrontal cortex, pgACC = perigenual anterior cingulate cortex, TPJ = temporoparietal junction, vmPFC = ventromedial prefrontal cortex, VS = ventral striatum.

3.5. Trial-level analyses (hypothesis 4)

To test the first part of hypothesis 4, trial-level analyses assessed the main effect of valence of the adjective on activation in the vmPFC, pgACC, and VS. For all three regions, the model including valence did *not* outperform the null model (vmPFC: $\chi^2(1) = 0.22$, $p = 0.64$; pgACC: $\chi^2(1) = 0.14$, $p = 0.71$; VS: $\chi^2(1) = 0.05$, $p = 0.82$).

To test the second component of this hypothesis, we assessed the degree to which endorsement differed as a function of adjective valence and activation in vmPFC and pgACC. For both the vmPFC and pgACC, the best fitting model included an interactive effect of valence and ROI activation on endorsement (see Tables 4 and 5, and Fig. 3). These models showed that higher activation in the vmPFC or pgACC was associated with a higher probability of endorsing negatively valenced adjectives, but a lower probability of endorsing positive adjectives.

3.5.1. Post-hoc analyses of age and puberty with trial-level data

In the analyses assessing the relationship between vmPFC and pgACC activation at the trial-level and age or pubertal development, none of the models outperformed the base model (including only valence as predictor). Thus, consistent with the whole-brain analysis described above, neither age nor pubertal development were associated with vmPFC or pgACC activation at the trial-level.

3.5.2. Post-hoc analyses of reaction time

The average RT for positive adjectives was 1.64 s (SD = 0.30) if they were endorsed, and 2.06 s (SD = 0.48) if they were rejected. The average RT for negative adjectives was 1.95 s (SD = 0.48) if they were endorsed, and 1.89 s (SD = 0.33) if they were rejected. The best fitting RT model showed that RT moderated the relationship between valence and endorsement. Participants were less likely to endorse a positive adjective when they took longer to respond. Further, adding interactions between RT and valence to the best fitting ROI models (described in 3.5 Trial-level analysis (hypothesis 4)) improved model fit (see Table 4). In this model, both the interactions between RT and valence, and ROI activation and valence were independently associated with response tendencies (see Table 6). Thus, adding RT did not attenuate valence interactions with the vmPFC or pgACC.

4. Discussion

4.1. Neural correlates of self-evaluation in relation to age and puberty

The current study examined how age and pubertal development are associated with neural activation during self-evaluation in a community sample of girls. In contrast to previous research (Jankowski et al., 2014; Pfeifer et al., 2007, 2009, 2013), we found no associations between either age or any measure of pubertal development and brain activation during self-evaluation in our hypothesized brain regions (including the vmPFC and pgACC, see preregistration: <https://osf.io/n268d/registrations>). The current study focused on self-evaluation of social traits,

including adjectives related to prosociality, social status/sociability, and antisociality/aggressiveness. These null findings are in line with van der Crujjsen et al. (2018), which reported no age effects (spanning ages 11–21) when self-evaluation was focused on prosocial traits. However, they are in contrast to Pfeifer et al. (2013), which reported age and puberty effects (at age 10–13) on vmPFC activation during self-evaluation of social traits related to status. We did not find an association between pubertal development and vmPFC activation, neither across all trials nor in interaction with adjective type. One potential explanation is that Pfeifer et al. (2013) was a longitudinal study and therefore assessed intra-individual changes in pubertal development and neural activation over time, rather than the cross-sectional associations examined in this study.

Although lack of power or other study-specific methodological factors can never be completely ruled out as reasons for null findings, the sample size of the current study was larger than that of most previous studies and variation in pubertal development was substantial across a relatively limited age window. Furthermore, the results were consistent across whole-brain, ROI and trial-level analyses, as well as across self-reported and biological measures of pubertal development, and the task showed robust effects on brain activation in regions associated with self-evaluation. Given the sample size of the current study (143) and the significance threshold used in ROI analyses (.0303), we had 95 % power to detect an f^2 of 0.10 or larger ($R^2 = 0.093$; for age) and 0.12 or larger ($R^2 = 0.108$; for any pubertal development variable), based on post-hoc power analyses for linear models with the pwr package v1.2-2 in R. Both of these are considered small effects (Cohen, 1988). Considering that our false-negative rate is controlled at the same level as our false-positive rate for small effects, and the low likelihood that other methodological factors obfuscated such associations, we conclude that there are no substantial associations with age and puberty in early adolescent girls.

Another possible explanation is that associations with age and/or pubertal development are stronger in boys, and are therefore more often found in a mixed sex sample. This could be the case for both main effects of pubertal development and interactions with adjective type. Pfeifer et al. (2013) reported no significant sex difference in the correlation between pubertal development and change in vmPFC activation, although in absolute terms the correlation was weaker in girls ($r = .45$ versus $.74$). Other studies did not examine whether associations with age or pubertal development differed by sex (Jankowski et al., 2014; Pfeifer et al., 2007; van der Crujjsen et al., 2018), thus this remains an open question.

With respect to the interaction between pubertal development and adjective type (hypothesis 3), we found that adjective type moderated the associations of DHEA and testosterone levels with neural activation in the bilateral ventral motor cortex. Girls with relatively high DHEA or testosterone levels activated this area more strongly when judging prosocial adjectives, whereas girls with lower hormone levels exhibited greater activation to antisocial and social status/sociability focused adjectives. This part of the motor cortex, which controls the mouth and tongue and is typically associated with speech production (based on

Table 3
Results of analyses relating age and pubertal development to ROI activation during self-evaluation.

Predictor	Outcome	F (df), p	t (SE), p	Adjusted R ²
Age	vmPFC	0.00 (1,141), 0.99	-0.02 (0.16), 0.99	-0.007
	pgACC	0.07 (1,141), 0.79	-0.26 (0.18), 0.79	-0.007
	VS	0.29 (1,141), 0.59	-0.54 (0.12), 0.59	-0.005
	TPJ	0.43 (1,141), 0.51	0.65 (0.11), 0.51	-0.004
	dmPFC	0.04 (1,141), 0.84	0.20 (0.16), 0.84	-0.007
Tanner stage	vmPFC	2.15 (2,140), 0.12	-2.07 (0.15), 0.04	0.016
	pgACC	1.21 (2,140), 0.30	-1.54 (0.17), 0.13	0.003
	VS	0.73 (2,140), 0.49	-1.08 (0.11), 0.28	-0.004
	TPJ	0.41 (2,140), 0.66	0.63 (0.10), 0.53	-0.008
	dmPFC	1.75 (2,140), 0.18	-1.86 (0.15), 0.06	0.011
DHEA	vmPFC	1.23 (2,140), 0.30	-1.57 (0.13), 0.12	0.003
	pgACC	0.52 (2,140), 0.60	-0.99 (0.14), 0.33	-0.007
	VS	0.98 (2,140), 0.38	-1.29 (0.10), 0.20	0.000
	TPJ	0.22 (2,140), 0.81	-0.09 (0.09), 0.93	-0.011
	dmPFC	0.15 (2,140), 0.86	-0.51 (0.13), 0.61	-0.012
Testosterone	vmPFC	0.46 (2,140), 0.63	-0.96 (0.31), 0.34	-0.008
	pgACC	0.29 (2,140), 0.75	-0.72 (0.34), 0.47	-0.010
	VS	0.31 (2,140), 0.73	-0.58 (0.23), 0.57	-0.010
	TPJ	0.22 (2,140), 0.80	0.11 (0.21), 0.91	-0.011
	dmPFC	0.02 (2,140), 0.98	0.06 (0.31), 0.95	-0.013
Estradiol	vmPFC	0.56 (2,140), 0.57	-1.05 (0.30), 0.29	-0.006
	pgACC	0.56 (2,140), 0.57	-1.02 (0.33), 0.31	-0.006
	VS	0.88 (2,140), 0.42	-1.21 (0.22), 0.23	-0.002
	TPJ	0.50 (2,140), 0.45	-0.76 (0.21), 0.45	-0.007
	dmPFC	0.26 (2,140), 0.77	-0.69 (0.30), 0.49	0.004

Note: Tanner stage and hormone models included age as a covariate. Bonferroni-adjusted p-threshold (see *Methods*) was .0303. df = degrees of freedom, DHEA = dehydroepiandrosterone, dmPFC = dorsomedial prefrontal cortex, pgACC = perigenual anterior cingulate cortex, SE = standard error, TPJ = temporoparietal junction, vmPFC = ventromedial prefrontal cortex, VS = ventral striatum.

posterior probabilities for these locations on NeuroSynth (Yarkoni et al., 2011), see Supplementary Table 2), has not consistently been associated with self-evaluation (van der Meer et al., 2010). Post-hoc analyses examining the interaction between valence and pubertal development, showed associations with DHEA and testosterone in clusters overlapping those found in the analysis with adjective type as a moderator, showing that the findings might, at least partially, be explained by valence. Unfortunately, our task was not designed to fully distinguish between adjective type and valence, thus this should be examined further in future research. Moreover, the results indicate the involvement of adrenal hormonal processes specifically, as associations with DHEA and testosterone, but not estradiol, were identified. Why adrenal hormonal processes would be associated with activation in this specific motor area is

unclear and warrants replication before any firm conclusions can be drawn.

4.2. Valence and associations with endorsement

On a trial-by-trial basis, we demonstrated no main effect of valence on vmPFC and pgACC activation during self-evaluation. However, our interaction analysis showed that it is important to take self-descriptiveness into account in this association. Specifically, we found that stronger vmPFC and pgACC activation were related to a higher probability of endorsing negative adjectives and a lower probability of endorsing positive adjectives. Post-hoc analyses showed that this interactive association held for negative status/sociability and antisociality/aggressiveness adjectives, as well as for positive status/sociability adjectives, but not for adjectives related to prosociality. This finding is in line with our fourth hypothesis that stronger engagement of vmPFC and pgACC would be associated with a higher probability of endorsing negatively valenced traits as self-descriptive. It is also largely consistent with previous findings from our lab in late adolescents (Cosme et al., 2019), which demonstrated a positive association between vmPFC activation and the probability of endorsing of negative statements related to ill-being. However, this finding partly contradicts the theory that the vmPFC does not support the formation of self-evaluations, but that its activation reflects personal significance or value to the self (D'Argembeau, 2013). Although we did not measure value to the self, our results show that the strength of vmPFC activation is associated with the self-evaluation that is formed (endorse or reject). In post-hoc reaction time (RT) analyses, we considered the possibility that increased vmPFC and pgACC activation were reflective of more thorough consideration of the adjective in relation to oneself, resulting in longer RT, and thereby leading to a decreased tendency to respond in a 'reflexive' or socially desirable manner. This would be in line with the theory that the vmPFC represents self-relatedness, and is activated more strongly when one draws on personal history and memories (Flagan and Beer, 2013). However, variation in RT (in interaction with valence) did not explain the association between vmPFC by valence or pgACC by valence and endorsement. Instead, interactions between RT and valence and between vmPFC or pgACC activation and valence were both independently associated with response tendencies. It is also notable that there was an overwhelming tendency to endorse positive traits and reject negative ones, which could partly arise from the use of a non-clinical sample and partly because of the self-positivity bias (Mezulis et al., 2004; Taylor and Brown, 1994). Therefore, an alternative explanation is that vmPFC and pgACC activation may help you identify what distinguishes you from others, with stronger activation making you more likely to 'swim against the current' (i.e., in our sample, rejecting something positive or endorsing something negative). For example, the positive word that all participants endorsed (nice) had the lowest mean vmPFC activation, while the word with the highest mean vmPFC activation (popular) was endorsed only 38 % of the time (see Supplementary Fig. 5 for parameter estimates and percentage endorsement for each word). However, this interpretation needs to be tested directly in future research.

4.3. Limitations and future directions

The current study had several strengths, including a relatively large sample of pubertal girls, and multiple ways of measuring pubertal development. However, the findings should be considered in light of several limitations. First, the current study was cross-sectional in nature. The only longitudinal study relating pubertal development and neural correlates of self-evaluation showed that increases in activation in the vmPFC were correlated with increases in self-reported pubertal development (Pfeifer et al., 2013), albeit in a small sample (N = 27). The discrepancy with the findings of the current study warrants further examination in a larger longitudinal dataset. Second, valence and adjective

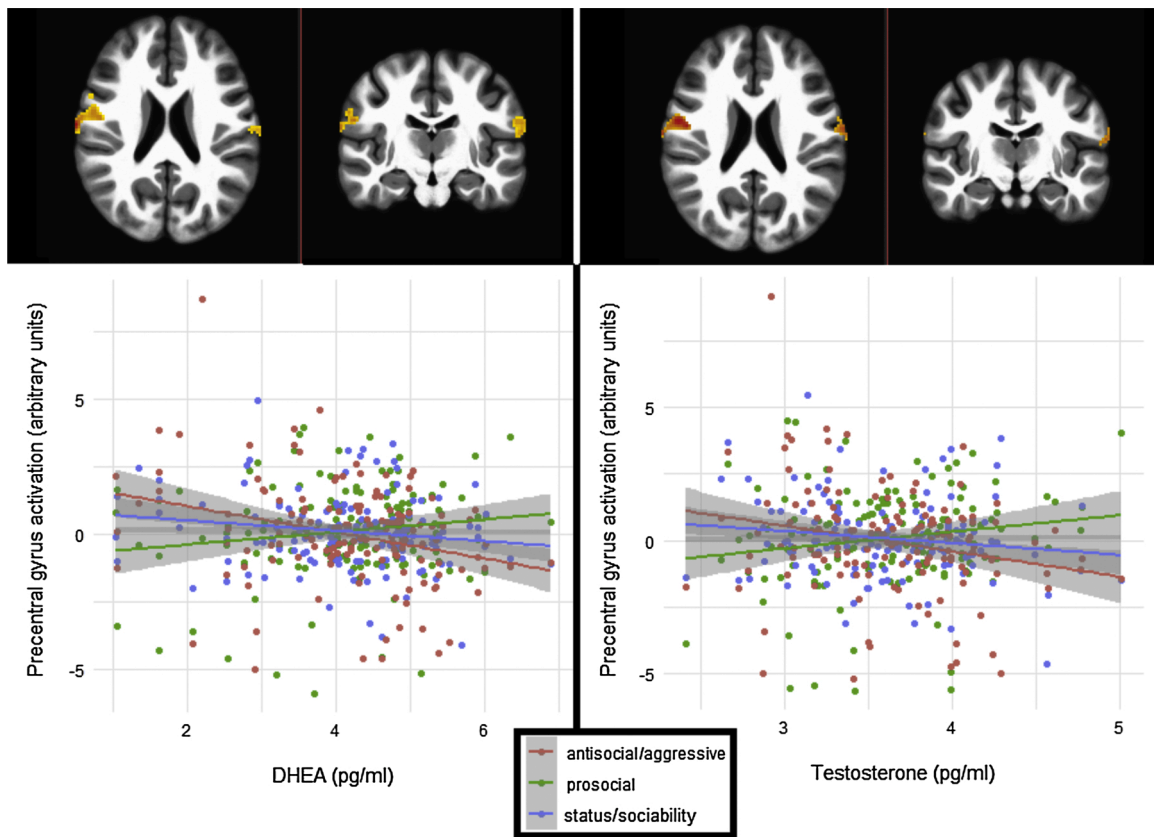


Fig. 2. Interactions between adjective type and DHEA (left) and testosterone (right). Top panels show clusters significant at $p_{fwe} < .05$. Bottom panels are individual parameter estimates averaged across clusters for each adjective type and plotted against hormone level.

Table 4
Comparison of trial-level models predicting endorsement.

ROI	Model	AIC	BIC	Chi-square	Degrees of freedom	p
vmPFC	Base model	6625	6645	-	-	-
	Main effect ROI model	6626	6654	0.40	1	.52
	Interaction model (ROI*valence)	6610	6644	18.13	1	<.001
	Interaction model with RT*valence (post hoc)	6447	6494	167.50	2	<.001
	Base model	6625	6645	-	-	-
pgACC	Main effect ROI model	6626	6653	0.50	1	0.48
	Interaction model (ROI x valence)	6609	6643	19.31	1	<.001
	Interaction model with RT x valence (post hoc)	6449	6497	163.56	2	<.001

Note: AIC = akaike information criterion, BIC = bayesian information criterion, pgACC = perigenual anterior cingulate cortex, ROI = region of interest, RT = reaction time, vmPFC = ventromedial prefrontal cortex.

type were not evenly balanced in this design (only the status/sociability factor had both positive and negative adjectives). Therefore, the findings regarding the interaction between pubertal development and adjective type have to be interpreted with caution. Third, our conclusions only hold for adolescent girls. Since sex differences in self-esteem have been reported specifically in early adolescence (Helwig and Ruprecht, 2017), it would be important to examine similar associations in boys or directly compare the sexes. Fourth, it would be relevant to repeat age and

Table 5
Parameter estimates of the ROI*valence interaction models.

ROI	Parameter	b	Odds ratio	SE	z	p
vmPFC	Intercept	0.13	1.14	0.04	3.55	<.001
	vmPFC	-0.04	1.00	0.03	-0.14	.89
	Valence	-1.49	0.23	0.03	-46.36	<.001
pgACC	vmPFC x Valence	0.14	1.15	0.03	4.25	<.001
	Intercept	0.13	1.14	0.04	3.53	<.001
	pgACC	-0.01	0.99	0.03	-0.22	.83
	Valence	-1.49	0.23	0.03	-46.36	<.001
	pgACC x Valence	0.14	1.15	0.03	4.38	<.001

Note: parameter estimates (b) are log-odds. pgACC = perigenual anterior cingulate cortex, ROI = region of interest, SE = standard error, vmPFC = ventromedial prefrontal cortex.

puberty analyses using a different control condition in the fMRI task to examine robustness or replicability of the findings. Future research should also examine how individual variation in neural correlates of self-evaluation relate to well-being and mental health in adolescence. This is relevant because of the increasing risk for internalizing disorders in adolescence (Costello et al., 2011) and the involvement of self-referential processing in these disorders (Lemogne et al., 2012). The vmPFC and pgACC would be candidate areas because of their association with self-descriptiveness judgements.

To conclude, we found no clear association between age or pubertal development and neural activation during self-evaluation in early adolescent girls, in contrast to our preregistered hypotheses which were based on prior research with small samples. Activation in the vmPFC and pgACC was associated with self-evaluative behavior in a valence-dependent manner. Specifically, increased trial-level activity in these regions was associated with greater probability of endorsing a negative trait as self-descriptive, but lower probability of endorsing a positive

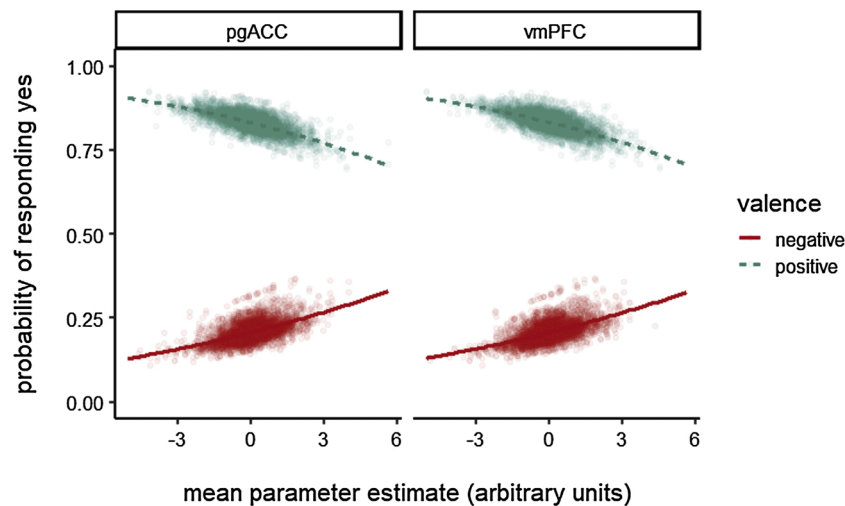


Fig. 3. Predicted response probabilities by valence of the adjective, plotted against the mean parameter estimates in the perigenual anterior cingulate cortex (pgACC) and in the ventromedial prefrontal cortex (vmPFC).

Table 6

Parameter estimates of the models including interactions between ROI and valence, and reaction time and valence.

ROI	Parameter	b	Odds ratio	SE	z	p
vmPFC	Intercept	0.72	2.05	0.09	7.84	<.001
	Valence	-2.29	0.10	0.09	-26.18	<.001
	vmPFC	-0.03	0.97	0.03	-0.95	0.34
	Reaction time	-0.32	0.73	0.04	-7.44	<.001
	vmPFC x Valence	0.17	1.19	0.03	5.11	<.001
	Reaction time x Valence	0.43	1.54	0.04	10.40	<.001
pgACC	Intercept	0.71	2.03	0.09	7.83	<.001
	Valence	-2.27	0.10	0.09	-26.10	<.001
	pgACC	-0.03	0.97	0.03	-0.77	0.44
	Reaction time	-0.32	0.73	0.04	-7.43	<.001
	pgACC x Valence	0.16	1.17	0.03	4.89	<.001
	Reaction time x Valence	0.42	1.57	0.04	10.22	<.001

Note: pgACC = perigenual anterior cingulate cortex, ROI = region of interest, SE = standard error, vmPFC = ventromedial prefrontal cortex.

trait. One possible explanation is that vmPFC and pgACC activation may help you identify what distinguishes you from others. Future studies should attempt to confirm or deny associations between puberty and neural activity related to self-evaluation in longitudinal designs, as well as test the predictive value of the neural correlates of self-evaluation for mental health in adolescence.

Funding

This work was supported by the National Institute of Mental Health (R01MH107418). TWC was also supported by the National Center for Advancing Translational Sciences (TL1TR002371); and DC was supported by the National Institutes of Health (F31CA232357). The funding agencies had no role in the design of the study or the collection, analysis, and interpretation of data or in writing the manuscript, apart from their financial contribution; the content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Declaration of Competing Interest

The authors declare that they have no competing interests.

Acknowledgements

We would like to thank all participants and their families for their involvement in the study. We would also like to thank the Lewis Center for Neuroimaging at the University of Oregon for their MRI-related support and the use of their facilities.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.dcn.2020.100799>.

References

- Barendse, M.E.A., Vijayakumar, N., Byrne, M., Flannery, J., Cheng, T., Flournoy, J., Nelson, B., Cosme, D., Mobasser, A., Chavez, S., Hval, L., Brady, B., Nadel, H., Helzer, A., Shirlcliff, E.A., Allen, N., Pfeifer, J.H., 2020. Study protocol: transitions in adolescent girls (TAG). *Front. Psychiatry* 10. <https://doi.org/10.3389/fpsy.2019.01018>.
- Bartra, O., McGuire, J.T., Kable, J.W., 2013. The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage* 76, 412–427. <https://doi.org/10.1016/j.neuroimage.2013.02.063>.
- Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67 (1), 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Brett, M., Anton, J.-L., Valabregue, R., Poline, J.-B., 2002. Region of interest analysis using an SPM toolbox. In: 8th International Conference on Functional Mapping of the Human Brain. Sendai, Japan. <http://marsbar.sourceforge.net/index.html>.
- Byrne, B.M., Shavelson, R.J., 1996. On the structure of social self-concept for pre-, early, and late adolescents: a test of the Shavelson, Hubner, and Stanton (1976) model. *J. Pers. Soc. Psychol.* 70 (3), 599–613. <https://doi.org/10.1037/0022-3514.70.3.599>.
- Chavez, R.S., Heatherton, T.F., Wagner, D.D., 2017. Neural population decoding reveals the intrinsic positivity of the self. *Cereb. Cortex* 27 (11), 5222–5229. <https://doi.org/10.1093/cercor/bhw302>.
- Cohen, J., 1988. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Routledge.
- Cosme, D., Flournoy, J.C., Vijayakumar, N., 2018. Auto-motion-fmriprep: a tool for Automated assessment of motion artifacts. Zenodo. <https://doi.org/10.5281/zenodo.1412131>.
- Cosme, D., Mobasser, A., Ross, G., Pfeifer, J.H., 2019. If You're Happy and You Know It: Neural Correlates of Self-Evaluated Psychological Health and Well-Being [Preprint]. Manuscript submitted for publication. <https://doi.org/10.31234/osf.io/86n3b>.
- Costello, E.J., Copeland, W., Angold, A., 2011. Trends in psychopathology across the adolescent years: what changes when children become adolescents, and when adolescents become adults? *J. Child Psychol. Psychiatry* 52 (10), 1015–1025. <https://doi.org/10.1111/j.1469-7610.2011.02446.x>.
- Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29 (3), 162–173. <https://doi.org/10.1006/cbmr.1996.0014>.
- Cox, R.W., Chen, G., Glen, D.R., Reynolds, R.C., Taylor, P.A., 2017. FMRI clustering in AFNI: false-positive rates redux. *Brain Connect.* 7 (3), 152–171. <https://doi.org/10.1089/brain.2016.0475>.

- D'Argembeau, A., 2013. On the role of the ventromedial prefrontal cortex in Self-Processing: the valuation hypothesis. *Front. Hum. Neurosci.* 7 <https://doi.org/10.3389/fnhum.2013.00372>.
- Dahl, R.E., Allen, N.B., Wilbrecht, L., Suleiman, A.B., 2018. Importance of investing in adolescence from a developmental science perspective. *Nature* 554 (7693), 441–450. <https://doi.org/10.1038/nature25770>.
- de Bruyn, E.H., Cillessen, A.H.N., 2006. Popularity in early adolescence: prosocial and antisocial subtypes. *J. Adolesc. Res.* 21 (6), 607–627. <https://doi.org/10.1177/0743558406293966>.
- Dégeilh, F., Guillery-Girard, B., Dayan, J., Gaubert, M., Chételat, G., Egler, P.-J., Baleyte, J.-M., Eustache, F., Viard, A., 2015. Neural correlates of self and its interaction with memory in healthy adolescents. *Child Dev.* 86 (6), 1966–1983. <https://doi.org/10.1111/cdev.12440>.
- Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S.S., Wright, J., Durnez, J., Poldrack, R.A., Gorgolewski, K.J., 2019. fMRIprep: a robust preprocessing pipeline for functional MRI. *Nat. Methods* 16 (1), 111–116. <https://doi.org/10.1038/s41592-018-0235-4>.
- Fischl, B., 2012. FreeSurfer. *NeuroImage* 62 (2), 774–781. <https://doi.org/10.1016/j.neuroimage.2012.01.021>.
- Flagan, T., Beer, J.S., 2013. Three ways in which midline regions contribute to self-evaluation. *Front. Hum. Neurosci.* 7 <https://doi.org/10.3389/fnhum.2013.00450>.
- Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., Smith, S.M., Van Essen, D. C., 2016. A multi-modal parcellation of human cerebral cortex. *Nature* 536 (7615), 171–178. <https://doi.org/10.1038/nature18933>.
- Gorgolewski, K.J., Auer, T., Calhoun, V.D., Craddock, R.C., Das, S., Duff, E.P., Flandin, G., Ghosh, S.S., Glatard, T., Halchenko, Y.O., Handwerker, D.A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B.N., Nichols, T.E., Pellman, J., et al., 2016. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci. Data* 3 (1), 1–9. <https://doi.org/10.1038/sdata.2016.44>.
- Harter, S., 2012. *The Construction of the Self: Developmental and Sociocultural Foundations*, 2nd ed. Guilford Publications.
- Helwig, N.E., Ruprecht, M.R., 2017. Age, gender, and self-esteem: a sociocultural look through a nonparametric lens. *Arch. Sci. Psychol.* 5 (1), 19–31. <https://doi.org/10.1037/arc0000032>.
- Jankowski, K.F., Moore, W.E., Merchant, J.S., Kahn, L.E., Pfeifer, J.H., 2014. But do you think I'm cool? Developmental differences in striatal recruitment during direct and reflected social self-evaluations. *Dev. Cogn. Neurosci.* 8, 40–54. <https://doi.org/10.1016/j.dcn.2014.01.003>.
- Lemogne, C., Delaveau, P., Preton, M., Guionnet, S., Fossati, P., 2012. Medial prefrontal cortex and the self in major depression. *J. Affect. Disord.* 136 (1), e1–e11. <https://doi.org/10.1016/j.jad.2010.11.034>.
- Measelle, J.R., Ablow, J.C., Cowan, P.A., Cowan, C.P., 1998. Assessing young children's views of their academic, social, and emotional lives: an evaluation of the self-perception scales of the berkeley puppet interview. *Child Dev.* 69 (6), 1556–1576. <https://doi.org/10.1111/j.1467-8624.1998.tb06177.x>.
- Mezulis, A.H., Abramson, L.Y., Hyde, J.S., Hankin, B.L., 2004. Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychol. Bull.* 130 (5), 711–747. <https://doi.org/10.1037/0033-2909.130.5.711>.
- Montoya, E.R., Terburg, D., Bos, P.A., van Honk, J., 2012. Testosterone, cortisol, and serotonin as key regulators of social aggression: a review and theoretical perspective. *Motiv. Emot.* 36 (1), 65–73. <https://doi.org/10.1007/s11031-011-9264-3>.
- Moran, J.M., Macrae, C.N., Heatherton, T.F., Wyland, C.L., Kelley, W.M., 2006. Neuroanatomical evidence for distinct cognitive and affective components of self. *J. Cogn. Neurosci.* 18 (9), 1586–1594. <https://doi.org/10.1162/jocn.2006.18.9.1586>.
- Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., Panksepp, J., 2006. Self-referential processing in our brain—a meta-analysis of imaging studies on the self. *NeuroImage* 31 (1), 440–457. <https://doi.org/10.1016/j.neuroimage.2005.12.002>.
- Petersen, A.C., Crockett, L., Richards, M., Boxer, A., 1988. A self-report measure of pubertal status: reliability, validity, and initial norms. *J. Youth Adolesc.* 17 (2), 117–133. <https://doi.org/10.1007/BF01537962>.
- Pfeifer, J.H., Lieberman, M.D., Dapretto, M., 2007. “I know you are but what am I?”: neural bases of self- and social knowledge retrieval in children and adults. *J. Cogn. Neurosci.* 19 (8), 1323–1337. <https://doi.org/10.1162/jocn.2007.19.8.1323>.
- Pfeifer, J.H., Masten, C.L., Borofsky, L.A., Dapretto, M., Fuligni, A.J., Lieberman, M.D., 2009. Neural correlates of direct and reflected self-appraisals in adolescents and adults: when social perspective-taking informs self-perception. *Child Dev.* 80 (4), 1016–1038. <https://doi.org/10.1111/j.1467-8624.2009.01314.x>.
- Pfeifer, J.H., Kahn, L.E., Merchant, J.S., Peake, S.J., Veroude, K., Masten, C.L., Lieberman, M.D., Mazziotta, J.C., Dapretto, M., 2013. Longitudinal change in the neural bases of adolescent social self-evaluations: effects of age and pubertal development. *J. Neurosci.* 33 (17), 7415–7419. <https://doi.org/10.1523/JNEUROSCI.4074-12.2013>.
- Romund, L., Golde, S., Lorenz, R.C., Raufelder, D., Pelz, P., Gleich, T., Heinz, A., Beck, A., 2017. Neural correlates of the self-concept in adolescence—a focus on the significance of friends. *Hum. Brain Mapp.* 38 (2), 987–996. <https://doi.org/10.1002/hbm.23433>.
- Shirtcliff, E.A., Dahl, R.E., Pollak, S.D., 2009. Pubertal development: correspondence between hormonal and physical development. *Child Dev.* 80 (2), 327–337. <https://doi.org/10.1111/j.1467-8624.2009.01263.x>.
- Taylor, S.E., Brown, J.D., 1994. Positive illusions and well-being revisited: separating fact from fiction. *Psychol. Bull.* 116 (1), 21–27. <https://doi.org/10.1037/0033-2909.116.1.21> discussion 28.
- van Buuren, S., Karin, G.-O., 2011. Mice: multivariate imputation by chained equations in r. *J. Stat. Softw.* 045 (i03). https://econpapers.repec.org/article/jssjstsof/v_3a045_3ai03.htm.
- van der Crujisen, R., Peters, S., van der Aar, L.P.E., Crone, E.A., 2018. The neural signature of self-concept development in adolescence: the role of domain and valence distinctions. *Dev. Cogn. Neurosci.* 30, 1–12. <https://doi.org/10.1016/j.dcn.2017.11.005>.
- van der Meer, L., Costafreda, S., Aleman, A., David, A.S., 2010. Self-reflection and the brain: a theoretical review and meta-analysis of neuroimaging studies with implications for schizophrenia. *Neurosci. Biobehav. Rev.* 34 (6), 935–946. <https://doi.org/10.1016/j.neubiorev.2009.12.004>.
- Vermeersch, H., T'sjoen, G., Kaufman, J.M., Vincke, J., Houtte, M.V., 2010. Gender ideology, same-sex peer group affiliation and the relationship between testosterone and dominance in adolescent boys and girls. *J. Biosoc. Sci.* 42 (4), 463–475. <https://doi.org/10.1017/S0021932010000106>.
- Vijayakumar, N., Op de Macks, Z., Shirtcliff, E.A., Pfeifer, J.H., 2018. Puberty and the human brain: insights into adolescent development. *Neurosci. Biobehav. Rev.* 92, 417–436. <https://doi.org/10.1016/j.neubiorev.2018.06.004>.
- Yarkoni, T., Poldrack, R.A., Nichols, T.E., Essen, D.C.V., Wager, T.D., 2011. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* 8 (8), 665–670. <https://doi.org/10.1038/nmeth.1635>.
- Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evol.* 1 (1), 3–14. <https://doi.org/10.1111/j.2041-210X.2009.00001.x>.