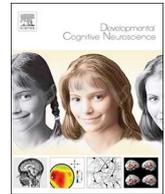




ELSEVIER

Contents lists available at ScienceDirect

## Developmental Cognitive Neuroscience

journal homepage: [www.elsevier.com/locate/dcn](http://www.elsevier.com/locate/dcn)

# Puberty and functional brain development in humans: Convergence in findings?

Junqiang Dai, K. Suzanne Scherf\*

Department of Psychology, United States



## ARTICLE INFO

## Keywords:

fMRI  
 Neuroimaging  
 Face processing  
 Reward processing  
 Social information processing  
 Cognitive processing

## ABSTRACT

Although there is a long history of studying the influence of pubertal hormones on brain function/structure in animals, this research in human adolescents is young but burgeoning. Here, we provide a comprehensive review of findings from neuroimaging studies investigating the relation between pubertal and *functional* brain development in humans. We quantified the findings from this literature in which statistics required for standard meta-analyses are often not provided (i.e., effect size in fMRI studies). To do so, we assessed *convergence* in findings within content domains (reward, facial emotion, social information, cognitive processing) in terms of the locus and directionality (i.e., positive/negative) of effects. Face processing is the only domain with convergence in the locus of effects in the amygdala. Social information processing is the only domain with convergence of positive effects; however, these effects are not consistently present in any brain region. There is no convergence of effects in either the reward or cognitive processing domains. This limited convergence in findings across domains is not the result of null findings or even due to the variety of experimental paradigms researchers employ. Instead, there are critical theoretical, methodological, and analytical issues that must be addressed in order to move the field forward.

## 1. Introduction

In 2004, Ronald Dahl gave the keynote address to the National Academy of Sciences on adolescent brain development. He began by providing an operational definition of adolescence describing it as, “that awkward period between sexual maturation and the attainment of adult roles and responsibilities” (p. 9; Dahl, 2004). This operational definition identifies adolescence as beginning with the onset of the physical and biological changes associated with puberty and concluding with the transition into an autonomous young adult who can take responsibility for their own behaviors (Dahl, 2004). This seminal talk identified many questions about the potential influence that pubertal development might have on the underlying neural circuitry supporting the emerging social behaviors that enable human adolescents to transition into autonomous adults.

At the very same time, results from a substantial body of literature were beginning to converge on the findings that pubertal hormones have both organizational and activational effects on brain structure, particularly in the neural circuitry that supports complex social behavior (for review see Sisk and Zehr, 2005a; 2005b). This work revealed that the brain is a target organ for steroid hormones and that there are time-sensitive, graded responses to hormones in the brain during adolescence (Sisk and Zehr, 2005a; 2005b). In fact, findings from the animal literature suggest that sex hormone-dependent organization of

neural circuits is a “fundamental feature of adolescent brain development” (Sisk and Zehr, 2005a; 2005b, p. 163). In other words, one cannot fully understand the mechanisms of adolescent brain development without understanding the influence of sex hormones on changes to brain structure and function. Addressing these mechanisms of brain development will be essential for understanding why and how novel complex social behaviors of adolescence emerge as do increasing vulnerabilities for developmental psychopathologies.

This call to arms led developmental neuroscientists to begin the work investigating how pubertal development might influence structural and functional brain development in human adolescents. There are many studies evaluating the relation between indices of pubertal and *structural brain development*, which have been critically reviewed (see review; Herting and Sowell, 2017; Goddings et al., 2019). However, work investigating associations between pubertal development and *functional brain development* is only just reaching a critical mass when it can be evaluated to assess convergence in findings. Here, we review this new body of work.

Specifically, we quantitatively evaluated the full set of neuroimaging studies, including fMRI and ERP studies, that investigated the relation between *functional brain development* and indices of pubertal development in human adolescents. Our broad goal was to understand whether there is convergence in the patterns of findings within functional domains (described below). More specifically, we had two

\* Corresponding author at: 113 Moore Building, Department of Psychology, Penn State University, University Park, PA, 16802, United States.  
 E-mail address: [suzyscherf@psu.edu](mailto:suzyscherf@psu.edu) (K.S. Scherf).

<https://doi.org/10.1016/j.dcn.2019.100690>

Received 21 December 2018; Received in revised form 28 July 2019; Accepted 1 August 2019

Available online 08 August 2019

1878-9293/© 2019 The Authors. 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/>).

organizing goals: first, we aimed to understand what the field has come to learn as a whole about the relation between pubertal and functional brain development. Second, we aimed to understand whether, and to what extent, inconsistencies in findings may be related to theoretical and/or methodological limitations. To the extent that these limitations exist, we provide recommendations and strategies to improve the work going forward.

The current paper is structured as follows. First, we describe the process of puberty and the operational ways that are typically used to assess pubertal development. Second, we review the 28 existing empirical studies investigating the relation between pubertal and functional brain development. We divided the papers into four separate functional domains based on the similarity of tasks and brain regions of interests; they include reward, facial emotion expression, social evaluation, and cognitive processing. We focused on determining the extent to which there are *converging findings* regarding the nature of the association (i.e., locus and directionality) between pubertal development and functional activation patterns in the brain within each domain. Finally, based on our review of the literature, we find consistent theoretical and methodological limitations across studies that limit the body of work as a whole and the convergence in findings. As a result, we provide a set of recommendations to address these limitations and hope that they will guide researchers working to understand the relation between pubertal and functional brain development in adolescents in future research.

## 2. Puberty as a developmental process and its measurement

It is important to understand that puberty is a biological, developmental process, with huge social implications, that unfolds over the course of approximately 8–10 years. Here, we provide a brief overview of the multiple mechanisms of pubertal development and the measures that are typically used to assess this developmental process.

Puberty is a process that results from a series of coordinated neuroendocrine events that lead to internal and external physical changes in secondary sexual characteristics and, eventually, enable one to reach reproductive maturity. It includes two independent but temporally overlapping processes: adrenarche and gonadarche. Adrenarche is the awakening of the adrenal glands and begins as early as age 6. It involves the increase in adrenal androgens, including dehydroepiandrosterone (i.e., DHEA), its sulfate (i.e., DHEAS), and androstenedione (Grumbach and Styne, 1992). These androgens are responsible for the emergence of axillary and pubic hair, body odor, and skin changes. The second process, gonadarche, begins with the reactivation of gonadotropin-releasing hormone neurons that stimulate the secretion of sex steroid hormones, testosterone and estrogens, from the gonads (Plant, 2002). Gonadarche is responsible for the emergence of the secondary sex characteristics such as breast development and menstruation in females, and phallic and teste development in males.

Researchers usually measure pubertal development via pubertal staging or hormonal assay<sup>1</sup>. Pubertal staging can be attained by physical examination, self-report, or parent-report based on sex-specific questions using Tanner criteria (Tanner, 1962), which allows researchers to categorize participants into 1 of 5 stages based on their physical appearance. The criteria are different for male and female adolescents. Pubic hair and breast development are evaluated for females, while pubic hair and genitalia development are evaluated for males. In so doing, the staging captures the core elements of both adrenarche and gonadarche.

The gold standard for measuring pubertal staging is physical examination by research-trained physician or nurse practitioner. However, physical examination is not always feasible due to the

expense, restrictions in environmental settings, and/or participants' reluctance to submit to the exam. As a result, researchers often rely on self-reported or parent-reported surveys to assess pubertal staging. One of the most widely used surveys is the Pubertal Development Scale (PDS, Petersen et al., 1988), which assesses adrenarche similarly in females and males, but gonadarche with sex-specific questions. The adrenarche questions in the PDS ask about pubic/body hair and skin changes. For gonadarche in females, there are questions about growth spurt, breast development, and the onset of menstruation. In males, the questions probing gonadarche are related to growth spurt, voice changes, and facial hair growth. Another survey, that can be completed as either a self- or parent-reported measure is the Sexual Maturation Scale (SMS, Morris and Udry, 1980). It measures adrenarche by asking participants to evaluate which of 5 images best captures the extent of pubic hair development in the adolescent. With regard to the gonadarche, participants identify 1 of 5 images that best represents breast development in females or phallic and teste development in males.

Hormonal assay is another way to measure pubertal development, which is usually measured via blood serum or saliva. Hormones of adrenarche and gonadarche can both be assayed. For example, DHEA is the most commonly measured hormone of adrenarche and can be measured in male and female adolescents. The hormones of gonadarche are more complicated. For example, testosterone is measurable in both male and female adolescents, but the base levels are six times higher in male adolescents, and the mechanistic role of testosterone in female adolescents is less clear. In contrast, estrogen is very difficult to measure even in female adolescents because of the monthly cyclic nature of this hormone and the relatively low levels in peri-pubertal girls (Dorn and Biro, 2011).

Throughout the literature review, it is important to keep in mind how these measures represent slightly different, but overlapping, aspects of pubertal development. This is critical for thinking about the hypotheses relating the mechanistic processes of pubertal development to functional brain development. It is also critical to be aware of the difficulty in capturing the nature of a time-sensitive developmental process with a single measure that is collected at one point in time.

## 3. Approach

### 3.1. Selection criteria

We limited our literature search to studies that used neuroimaging methods, including functional magnetic resonance imaging (fMRI) and electroencephalography (EEG). To identify the studies, we input the search terms 'adolescent,' 'puberty,' 'testosterone,' 'estrogen,' 'PDS,' 'Tanner,' 'fMRI,' 'brain,' 'ERP,' 'EEG' into the PubMed and Google scholar databases. Studies included in this review had to conform to the following selection criteria: (a) published in an English peer-reviewed journal; (b) empirically assessed pubertal development (i.e., via hormonal assay, self- or parent-reported secondary sex characteristics, physical examination); and (c) used neuroimaging methods (fMRI or EEG/ERP) to examine functional brain development in adolescents. This resulted in a total of 28 studies.

Given our focus on functional brain development, we organized the studies into four domains based on the similarity in process and underlying neural circuitry that is elicited by the tasks in which participants are typically engaged (see also Goddings et al., 2019). This included 9 reward, 7 facial emotion, 7 social evaluation, and 5 cognitive processing studies. Table 1 lists the studies in each domain as well as the demographic characteristics of the study sample and the method(s) of pubertal assessment. Fig. 1 identifies the set of neural regions that were investigated across studies in each functional domain.

### 3.2. Meta-analytic approach

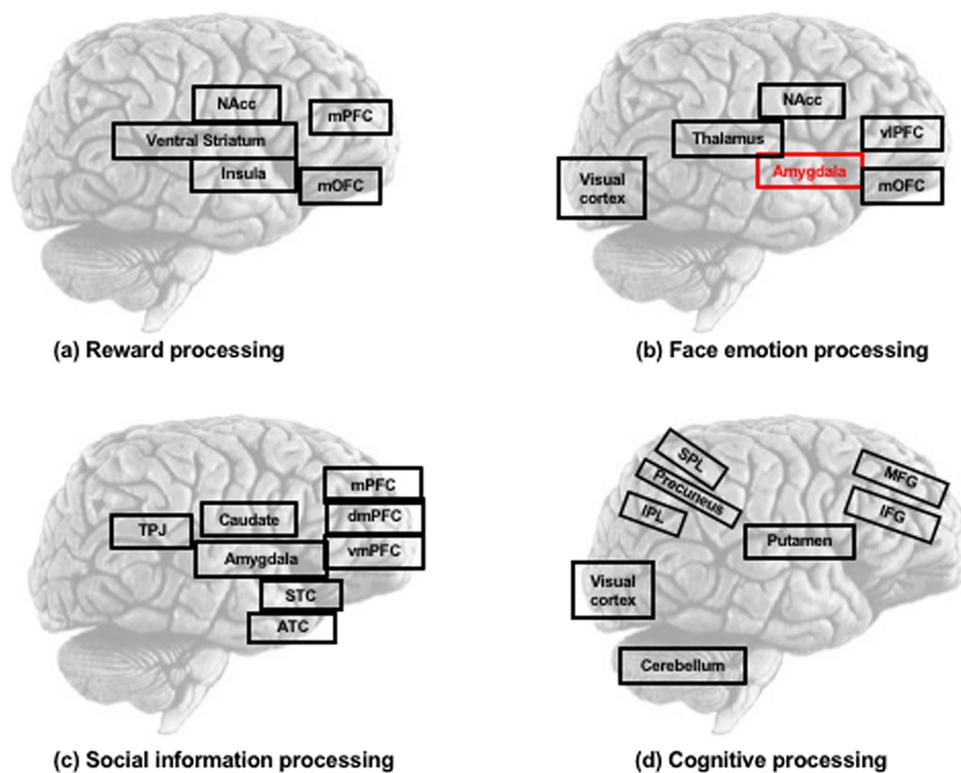
The goal of a meta-analysis is to integrate findings across different studies (paradigms, samples, protocols, etc.) and determine whether

<sup>1</sup> It is beyond the scope of this paper to provide an in-depth analysis of the strengths and weaknesses of each approach. Instead, we refer the reader to existing reviews that address these issues very comprehensively (see Dorn et al., 2006; Dorn and Biro, 2011; Shirtcliff et al., 2009).

**Table 1**  
Demographic Characteristics of Studies Investigating Pubertal and Functional Brain Development.

Study	N	Age Range	Sex	Study Design	Puberty Assessment(s)
<b>Reward Processing</b>					
Braams et al. (2015)*	299	8-26 years	52% female	Longitudinal	Self PDS, T
Braams et al. (2016)*	169	8-27 years	49% female	Longitudinal	T
Bress et al. (2012)	64	8-13 years	41% female	Cross-sectional	Self PDS, parent PDS
Forbes et al. (2010)	77	11-13 years	55% female	Cross-sectional	Physical exam, T
Ladouceur et al. (2018)	79	10-13 years	59% female	Cross-sectional	Self PDS, PBIP, T, E, DHEA
Op de Macks et al. (2011)	50	10-16 years	66% female	Cross-sectional	Self PDS, PBIP, T
Op de Macks et al. (2016)*	68	11-13 years	100% female	Cross-sectional	Self PDS, PBIP, T, E
Op de Macks et al. (2017) ^	58	11-13 years	100% female	Cross-sectional	Self PDS, T, E
van Duijvenvoorde et al. (2014)	31	12-19 years	58% female	Cross-sectional	Self PDS
<b>Facial Emotion Processing</b>					
Ferri et al. (2014)	75	8-15 years	100% female	Cross-sectional	Self and parent PDS, Self and parent PBIP
Forbes et al. (2011)	76	11-13 years	53% female	Cross-sectional	Staging by exam
Moore et al. (2012)	45	10-13 years	58% female	Longitudinal	Self PDS
Spielberg et al. (2014b)	38	11-13 years	55% female	Longitudinal	T
Spielberg et al. (2014a)	41	12-14 years	51% female	Longitudinal	T
Telzer et al. (2015)	30	9-16 years	Not reported	Cross-sectional	Parent PDS
Tyborowska et al. (2016)	47	14 years	55% female	Cross-sectional	Self PDS, T
<b>Social Information Processing</b>					
Goddings et al. (2012)	42	11-14 years	100% female	Cross-sectional	Visual inspection of Tanner stage, self-report menarche status, T, O, DHEA
Jankowski et al. (2014)	20	11-14 years	50% female	Cross-sectional	Self PDS
Klapwijk et al. (2013)	35	11-13 years	100% female	Cross-sectional	Visual inspection of Tanner stage, self-report menarche status, T, O, DHEA
Masten et al. (2013)	16	10-13 years	56% female	Longitudinal	Self PDS
Pfeifer et al. (2013)	27	10-13 years	67% female	Longitudinal	Self PDS
Silk et al. (2014)	48	11-17 years	71% female	Cross-sectional	Self PDS
Silk et al. (2017)	49	10-18 years	86% female	Cross-sectional	Self PDS
<b>Cognitive Processing</b>					
Alarcón et al. (2014)	74	10-16 years	50% female	Cross-sectional	T
Brumback et al. (2012)	114	8-13 years	47% female	Cross-sectional	Self PDS
Cservenka et al. (2015)	44	10-15 years	50% female	Cross-sectional	Self PDS, T, E
Peters et al. (2014)	268	8-25 years	51% female	Cross-sectional	Self PDS, T, E
Schweinsburg et al. (2005)	49	12-17 years	51% female	Cross-sectional	Self PDS

Notes: Studies are organized by content domain of functional task used in the scanner to elicit neural activation. Within each content domain, the studies are organized alphabetically. PBIP – Picture-Based Interview about Puberty, PDS - Pubertal Development Scale, T – testosterone, E – estradiol, O - oestradiol, and DHEA – dehydroepiandrosterone. \* and ^ represent studies that explicitly state that they used same participant sample (e.g., Braams et al., 2015 and Braams et al., 2016).



**Fig. 1.** Illustration of convergence of findings regarding locus of effects within each content domain. Each content domain is represented graphically with the combined set of regions of interest that were investigated across studies for (a) Reward, (b) Face emotion, (c) Social information, and (d) Cognitive processing. Brain regions coded in red indicate convergence regarding locus of pubertal effects within a domain, meaning that more than 50% of the studies within the domain exhibited a significant relation between pubertal development and neural activation. Note that this convergence only existed in the amygdala for face emotion processing. ATC - anterior temporal cortex, dmPFC - dorsomedial prefrontal cortex, IFG - inferior frontal gyrus, IPL - inferior parietal lobule, mPFC - medial prefrontal cortex, MFG - middle frontal gyrus, NAcc - nucleus accumbens, OFC - orbitofrontal cortex, TPJ - temporoparietal junction, vIPFC - ventrolateral prefrontal cortex, SPL - superior parietal lobule, STC - superior temporal cortex.

there are consistent patterns in the findings. This is important because individual studies are vulnerable to measurement error and biases in the estimation of effect sizes, an issue that is particularly problematic in neuroimaging studies (see Wager et al., 2007). Unfortunately, fMRI studies often do not report the data that are required for traditional meta-analyses to estimate an overall effect size, including effect sizes, means and standard deviations, or the results of *t*-tests (see Radua and Mataix-Cols, 2012). Of the studies we reviewed here, only 32% reported an effect size of the measure of association between neural activation and pubertal development. Therefore, a traditional meta-analysis was not possible with these data.

Activation likelihood estimation (ALE) is a meta-analytic tool for fMRI data that is designed to determine regions of consistent activation across tasks. However, this was also not a useful tool to address the questions we asked about this literature. First, only significant results are submitted to an ALE analysis and then evaluated for consistency in location. We wanted to evaluate convergence in findings across both null and significant results. Second, ALE requires a minimum of 10–15 studies (Laird et al., 2009). There are not enough data within any of the functional domains in this literature to do an ALE analysis. Third, different patterns of blood-oxygen-level-dependent (BOLD) responses (negative, positive) have to be analyzed in separate ALE analyses. We wanted to evaluate convergence in the directionality of the association between functional activation and metrics of pubertal development across studies.

Therefore, to conduct a quantitative meta-analysis of this literature, we used a version of a *label-based* approach to assess convergence in the patterns of findings within each functional domain. A label-based approach is especially useful when information from neuroimaging studies are insufficient for conducting ALE or traditional meta-analyses (Radua and Mataix-Cols, 2012). It is a version of a region of interest (ROI) based meta-analysis and involves counting the number of times a particular ROI is reported as having significant activation (Laird et al., 2005; Phan et al., 2002). This allowed us to determine consistency in the regional locus of findings across studies.

Using this approach, we focused on determining whether there is convergence of findings regarding the association between a measure of

pubertal development and functional brain development. Importantly, there were too few studies within each functional domain to also consider the specific index of pubertal development (e.g., hormonal assay versus self-reported staging) in our assessment of convergence. Of note, these indices are all moderately correlated (Shirtcliff et al., 2009), which indicates some convergent validity across measures and supports the notion that it is reasonable to include them in the same meta-analysis. However, we do discuss potential differences between measurement types when summarizing the findings. Finally, it is important to acknowledge that the majority of the studies are cross-sectional in nature, but that longitudinal studies have more power to uncover the causal relations between puberty and functional brain development.

We identified two kinds of convergence. First, we looked for *convergence in the location of effects*. In other words, we asked whether within each functional domain there is convergence of findings indicating that a particular region(s) is consistently identified as a place where neural activation and metrics of pubertal development are linked. Importantly, because of limitations in neural source localization for ERP studies, we excluded them when assessing convergence in location of effects. Second, we looked for *consistency in directionality of the effects*, specifically in the relation between increasing pubertal development and neural activation. In evaluating the directionality of the effects, it is important to remember that pubertal development is generally increasing at the level that is being measured in these studies (i.e., increasing Tanner stages, increasing hormone levels). However, task modulations can either increase or decrease neural activation. As a result, it is important to consider whether there is consistency in the *directionality of relations between increasing pubertal development and neural activation* as a function of task (i.e., positive correlation or negative correlation). We evaluated consistency in each of these two kinds of convergence (i.e., locus and direction of effects) across studies in each domain separately. As in previous label-based meta-analyses of fMRI data (e.g., Phan et al., 2002), we used a > 50% criterion for convergence, which represents a simple majority of studies with the pattern of results. Specifically, when the majority of studies (> 50%) satisfied either criterion, convergence was met.

Finally, in describing the results, it is important to note that the study

**Table 2**  
Findings of Studies Investigating Puberty and Functional Brain Development.

Study	Analysis of Age	Brain Analysis	Sig	General Findings
<b>Reward Processing</b>				
Braams et al. (2015) *	Covariate	ROI corrected	+	• Changes in T positively related to NAcc reward activation.
Braams et al. (2016) *	Covariate	ROI corrected	φ	• No relation between T and reward activation in NAcc.
Bress et al. (2012)	None	ERP corrected	φ	• No relation between pubertal stage and FN ERP reward response.
Forbes et al. (2010)	Covariate; Narrow range	ROI corrected	+	• T positively correlated with caudate nucleus reward anticipation activation in boys.
			-	• T negatively correlated with caudate nucleus reward outcome activation in both boys and girls.
			-	• Pubertal stage negatively related to caudate nucleus outcome activation and positively related to mPFC outcome activation in combined sample.
Ladouceur et al. (2018)	Covariate; Narrow range	ROI corrected	-	• E negatively correlated with NAcc reward response in girls.
			+	• T positively correlated with functional connectivity between NAcc and insula, ACC during reward in girls.
Op de Maacks et al. (2011)	None	Whole-brain un-corrected	φ	• No relation between any metrics of pubertal development and the NAcc reward response in boys or in combined sample.
Op de Maacks et al. (2016) *	Covariate; Narrow range	Whole-brain corrected	+	• T positively correlated with the NAcc reward response.
			+	• E positively related to NAcc activation during risk-taking in girls.
			+	• T positively related to mOFC activation during risk-taking in girls.
Op de Maacks et al. (2017) *	Covariate; Narrow range	Whole-brain & ROI corrected	+	• E positively related to anterior insula social-status reward activation in girls.
van Duijvenvoorde et al. (2014)	None	ROI corrected	φ	• No relation between pubertal stage and NAcc reward activation.
<b>Facial Emotion Processing</b>				
Ferri et al. (2014)	None	ROI corrected	-	• Pubertal stage related to less amygdala neutral face activation in girls.
			-	• Mid/late pubertal stage related to less vIPFC fear face activation in girls.
Forbes et al. (2011)	Covariate; Narrow range	ROI corrected	-	• Mid/late pubertal stage related to less amygdala neutral face activation.
Moore et al. (2012)	Covariate	Whole-brain corrected	+	• Pubertal stage positively related to more amygdala, thalamus, and visual cortex activation to emotional faces at age 10 and 13.
Spielberg et al. (2014b)	Covariate; Narrow range	ROI-PPI corrected	+	• T positively related to amygdala and NAcc emotional face activation.
Spielberg et al. (2014a)	None	ROI-PPI corrected	-	• T negatively related to amygdala-OFC functional connectivity during face expression.
Telzer et al. (2015)	Covariate	ROI corrected	+	• Pubertal stage positively related to amygdala response to opposite-sex emotional faces.
Tyborowska et al. (2016)	Narrow range	Whole-brain & ROI partially corrected	+	• T positively related to more prefrontal cortex activation during emotional contexts.
			-	• T negatively related to amygdala and pulvinar activation during emotional contexts.
			+	• T positively related to functional connectivity between amygdala and pulvinar during emotional contexts.
<b>Social Information Processing</b>				
Goddings et al. (2012)	Covariate; Narrow range	Whole-brain corrected	+	• O and DHEA positively related to ATC social word activation in girls.
Jankowski et al. (2014)	Narrow range	Whole-brain & ROI corrected	+	• Pubertal stage positively related to bilateral VS activation during social self-evaluation task.
Klapwijk et al. (2013)	Covariate; Narrow range	Whole-brain corrected	+	• O positively related to functional connectivity between dmPFC and TPJ when evaluating social scenarios in girls
			+	• Pubertal stage positively related to functional connectivity between dmPFC and ATC in girls.
Masten et al. (2013)	Narrow range	Whole-brain corrected	+	• Pubertal stage positively related to dmPFC, PCC, TPJ, and temporal lobe peer rejection activation at age 13.
			+	• Change in pubertal stage from age 10 to 13 positively related to change in dmPFC and temporal lobe peer rejection activation at age 13.
Pfeifer et al. (2013)	Covariate; Narrow range	ROI corrected	+	• Pubertal stage positively related to social self-evaluation activation in vmPFC.
Silk et al. (2014)	Covariate	Whole-brain corrected	+	• Pubertal stage positively related to more amygdala and caudate activation during peer rejection.
Silk et al. (2017)	Covariate	Whole-brain & ROI corrected	φ	• No correlation between PDS and network activation to social status words.
<b>Cognitive Processing</b>				
Alarcón et al. (2014)	Covariate	Whole-brain corrected	φ	• No relation between T and spatial working memory activation.
Brumback et al. (2012)	Covariate	ERP Not reported	φ	• Pubertal stage not related to P300 response.
Cservenka et al. (2015)	Covariate	Whole-brain corrected	-	• T negatively related to frontal and putamen activation in boys during emotion conflict task.
			-	• T negatively related to parietal and cerebellar activation in girls during emotion conflict task.
			-	• E negatively related to frontal and cerebellar activation in boys during emotion conflict task.
			+	• E positively related to occipital activation in girls during emotion conflict task.
Peters et al. (2014)	None	ROI corrected	φ	• Neither PDS nor hormones related to frontoparietal network activation during feedback learning.
Schweinsburg et al. (2005)	None	Whole-brain corrected	-	• PDS stage negatively related to activation in superior parietal lobe during working memory task.

Notes: Studies are organized by content domain of functional task used in the scanner to elicit neural activation. When associations between a measure of pubertal development and functional brain development were reported, we included it in the table. They are indicated with a (+) positive or (-) negative sign to reflect the direction of the association. Otherwise, null results are reported with (φ) symbols. In other words, we have not included all the null results from a study with a positive result. Within each content domain, the studies are organized alphabetically. PBIP – Picture-Based Interview about Puberty, PDS – Pubertal Development Scale, T – testosterone, E – estradiol, O – oestradiol, and DHEA – dehydroepiandrosterone; ROI – regions of interest; ATC – anterior temporal cortex, dmPFC – dorsomedial prefrontal cortex, FN – feedback negativity, mPFC – medial prefrontal cortex, NAcc – nucleus accumbens, OFC – orbitofrontal cortex, STC – superior temporal cortex, TPJ – temporoparietal junction, vIPFC – ventrolateral prefrontal cortex, vmPFC – ventromedial prefrontal cortex. \* and ^ represent studies that explicitly state that they used the same participant sample (e.g., Braams et al., 2015 and Braams et al., 2016).

designs are observational and correlational by nature. However, given the compelling evidence from animal models showing that experimental manipulations in sex hormones influence organizational changes in brain circuitry and behavior in adolescence (Sisk and Foster, 2004; Sisk and Zehr, 2005a, 2005b), many researchers describe their findings using strong, more causal language. Therefore, as we reviewed the literature in the following section, we described the findings using the language that the researchers themselves used to interpret their findings.

## 4. Results

### 4.1. Pubertal and functional brain development during reward processing

The studies investigating the relations between pubertal development and the neural basis of reward processing in adolescents are grounded in empirical findings that sensation-seeking behaviors increase in adolescence (Steinberg et al., 2008). Sensation seeking refers to the tendency to seek out novel, varied, or highly stimulating experiences, and the willingness to take risk in order to attain them (Zuckerman et al., 1978). Based on animal models (Spear, 2000), Steinberg (2008) proposed that this increase in sensation-seeking during adolescence is likely related to changes in the dopaminergic system that co-occur with pubertal development. He proposed that puberty-related changes in the dopaminergic system enhance the rewarding value of stimuli to adolescents, thereby motivating them to engage in more risk-taking behaviors to experience such stimulation (Shulman et al., 2015; Steinberg et al., 2008). Following this argument, researchers have hypothesized a positive association between pubertal development and neural activation in brain regions that are associated with reward processing and that are highly influenced by dopamine, with a primary focus on the ventral striatum (VS), and nucleus accumbens (NAcc) within the VS, and the ventromedial prefrontal cortex (vmPFC; see Fig. 1a).

These studies have largely employed adapted versions of monetary gambling tasks as their primary measure of risk taking. In these tasks, participants make an active choice (e.g., to play or pass on a card) and received feedback on each trial about whether they win or lose money. Paradigms like this reliably elicit activation in the reward circuitry in adolescents, including the ventral striatum, dorsal striatum, insula, and posterior cingulate cortex and there is some evidence that activation in these regions changes with age (for review see Silverman et al., 2015). The primary analysis strategy across studies is to use a ROI approach, with a focus on the NAcc, and correlate the magnitude of the signal in this ROI with various measures of pubertal development. To preview the findings, 50% (i.e., 7 out of 14; see Table 2) of the findings reported a *positive* association between some measure of pubertal development and neural activation within some region of the reward processing system (Braams et al., 2015; Forbes et al., 2010; Op de Macks et al., 2011, 2016, 2017). In contrast, 21% of the findings reported a *negative* association between some metric of pubertal development and neural activation in the reward system (Ladouceur et al., 2018; Forbes et al., 2010). Finally, 29% of the findings reported a non-significant association between neural activation in the reward system and measures of pubertal development (Ladouceur et al., 2018; Braams et al., 2016; Bress et al., 2012; van Duijvenvoorde et al., 2014).

Specifically, all of the studies that reported a positive association between reward-related activation and pubertal development measured sex hormones as the index of puberty. For example, Op de Macks et al. (2011) measured testosterone in 50 adolescent boys and girls. They used an fMRI event-related gambling task and acquired BOLD signal from the NAcc during reward processing. The authors reported a positive correlation between activation in the NAcc during reward versus loss trials and testosterone levels in both boys and girls. However, this same correlation did not hold with the self-report measures of pubertal development in either boys or girls. These findings are consistent with the notion that more advanced pubertal development, as measured by increased testosterone, is related to higher magnitude neural responses during reward processing in the ventral striatum. The same researchers

did a subsequent study to evaluate risk-taking and reward processing (reward versus loss) in female adolescents (Op de Macks et al., 2016). The authors reported that activation during risk-taking (play vs. pass trials) in the medial orbitofrontal cortex (mOFC) and NAcc was positively associated with both estradiol and testosterone in female adolescents. However, in contrast to their previous findings (Op de Macks et al., 2011), there was no relation between activation during reward processing (reward versus loss trials) in the NAcc and hormone levels. In their most recent study, Op de Macks and colleagues revised their Jackpot gambling task to include both monetary and social status rewards (Op de Macks et al., 2017). They scanned 11-13-year-old female adolescents in this study and measured both self-reported PDS as well as hormone assays. In contrast to previous findings using the monetary version of the Jackpot gambling task, the ROI-based analyses revealed no association between NAcc activation and pubertal development. However, they reported a positive association between estradiol levels and differential activation between social and monetary reward in the bilateral anterior insula from a whole-brain analysis.

In the fourth positive finding study, Braams et al. (2015) also measured testosterone in a large sample of adolescents and young adults between the ages of 8 to 27 years in a longitudinal study. They employed a heads-or-tails gambling fMRI event-related task in which participants won or lost money on each trial (Braams et al., 2015). The researchers reported a positive association between the magnitude of NAcc reward-related activation (win versus loss trials) and testosterone. However, there were two problematic confounds in these analyses. They combined 1) individuals from the entire age range of the sample (8-27 years), including sexually mature adults, with the highest levels of testosterone, and 2) girls and boys, who have much higher testosterone values, in the same analysis. As a result, these results may be influenced by age- and sex-related differences in testosterone that are not specifically related to the process of pubertal development. In a separate analysis of a smaller subset of these same participants, the authors did not replicate the association between NAcc reward-related activation and testosterone levels (Braams et al., 2016).

There are two studies with mixed findings, both of which measured hormones and various methods of pubertal staging as metrics of pubertal development. In one of the studies, Forbes and colleagues tested two similarly aged groups of adolescents (approximately age 12) who varied in pubertal development, as determined by Tanner staging via physical examination (Forbes et al., 2010). During each trial of the slow-event-related fMRI paradigm, participants guessed whether the value of a card was higher or lower than 5. Using an ROI analysis approach, the researchers evaluated pubertal group differences in reward-related activation (reward vs. fixation) for two parts of the trial (reward anticipation, reward feedback). They reported that the advanced puberty group exhibited decreased activation in the caudate nucleus and increased activation in mPFC to reward feedback compared to the early puberty group. There were no group differences in activation during reward anticipation. In addition, testosterone levels were negatively associated with activation in the caudate nucleus during reward feedback in both boys and girls, but positively associated with activation related to reward anticipation in boys. In the second study, Ladouceur and colleagues collected self-report pubertal information (i.e., PBIP, PDS) and hormone assays from adolescents ages 10-13 years old (Ladouceur et al., 2018). They created a novel reward paradigm that unconfounded the predictability and reward received on individual trials. Participants were required to identify the spatial location of a gopher and were cued on each trial about whether it was a potential reward or non-reward trial. Within a set of predefined regions, they contrasted the correct reward and non-reward trials and submitted this activation to multiple correlational analyses with the pubertal metrics. The researchers reported that as estradiol level increases in girls, caudate activation decreases. There were no significant associations between any metrics of pubertal development and functional activation to reward in the boys. Functional connectivity analyses revealed that as testosterone increased in girls, connectivity

between the NAcc and putamen increased. These complicated sets of findings are largely inconsistent with the predicted positive association between reward-related neural activation and pubertal development.

The remaining studies all failed to find an association between indices of pubertal development and reward-related neural activation. Specifically, van Duijvenvoorde and colleagues tested a group of adolescents using the same gambling task as described in Op de Macks et al. (2011). They conducted the same ROI-based analysis but used the self-report PDS as the measure of pubertal development, rather than sex hormones (Van Duijvenvoorde et al., 2014). They reported no significant relation between ventral striatum activation to reward and pubertal staging in adolescents (12–19 years). In an event-related potentials (ERPs) study, researchers investigated the feedback negativity (FN) response to rewards and losses in a group of 8- to 13-year-old adolescents (Bress et al., 2012). The FN potential is an ERP component that is sensitive to reward outcome (Dunning and Hajcak, 2007) and source localization techniques suggest that it originates from the striatum (Foti et al., 2011). The researchers reported no association between the magnitude of the FN and either the adolescent- or parent-reported PDS.

It is important to note that there are inconsistencies in findings in spite of the similarity in tasks used across studies. In fact, several pairs of studies employed the exact same paradigm but reported different results. For example, the two studies by Braams and colleagues (Braams et al., 2015, 2016) employed the same flip coin guessing task with the same participants and found an association between testosterone and reward-related activation in one study (Braams et al., 2015) but not the other (Braams et al., 2016). Similarly, Op de Macks et al. (2011, 2016) and van Duijvenvoorde et al. (2014) both used the same gambling task but reported inconsistent results. As a result, it does not appear that the task demands or participant sample is the primary cause of the inconsistency in findings across studies.

In sum, the studies investigating the relation between indices of pubertal development and neural activation related to reward processing failed to meet our criterion for convergence in the *directionality* of effects. Specifically, four studies reported greater activation during reward tasks as a function of increasing testosterone level and two studies reported greater activation as a function of increasing estradiol level (but only in female adolescents; see Tables 1 and 2). However, four studies reported null effects and two studies reported three negative associations such that greater hormone levels were associated with lower levels of neural activation (see Table 2). Also, this set of studies does not meet our criterion for convergence in the locus of effects. This is true in spite of the fact that many of the studies hypothesized that the nucleus accumbens would be a region of interest in which such an effect would be observed. This mix of findings does not support the hypothesized positive association between pubertal hormones and neural activation to reward processing in brain regions that are highly influenced by dopamine.

#### 4.2. Pubertal and functional brain development during face emotion processing

The studies investigating relations between pubertal development and face perception in adolescents are grounded in behavioral findings that emotional expression processing develops late into adolescence (Brown and Dunn, 1996; Herba and Phillips, 2004; Herba et al., 2006; Motta-Mena and Scherf, 2016; Pine et al., 2004; Thomas et al., 2007) and is disrupted during the ages of 10–13 years when adolescents are undergoing pubertal development (Lawrence et al., 2015; McGivern et al., 2002). Given the affective nature of emotional expressions, researchers have primarily focused on the subcortical structures contributing to face processing, and the amygdala in particular. Indeed, meta-analyses of fMRI studies with adults across multiple face emotion paradigms implicate the amygdala (Fusar-Poli et al., 2009). Empirical studies investigating adolescent responses to emotional faces also consistently indicate activation in the amygdala (Tahmasebi et al., 2012). Also, research on neural activation to affective information indicates

that there are age-related changes in amygdala responsivity during the transition from childhood to adolescence (see Scherf et al., 2013, for review). This has led to the prediction by many researchers that the amygdala is “more reactive to facial displays” during adolescence (Guyer et al., 2008; Hare et al., 2008).

To investigate the association between measures of pubertal development and neural activation to face emotion displays, many of the existing studies employed an experimental paradigm established by Hariri et al. (2000) using adult faces. In this paradigm, participants are presented with three faces of different identities in which a target face is at the top of the display and two non-target faces are at the bottom. Participants are required to pick the non-target face that displays the matching expression exhibited by the target face. The contrasting condition involves presenting circles and ovals that have to be matched for similar shape. Previous adult studies have shown that the contrast of face versus shape reliably elicits activation in amygdala (Hariri et al., 2000). To preview the findings in this domain, there is convergence in the *locus* of effects. All 7 of the studies reported an association between some metric of pubertal development and neural activation during face emotion processing in the amygdala (see Table 2; Fig. 1b). However, the findings across studies did not meet our criterion for convergence in the *direction* of effects (i.e., > 50%).

Five of the seven studies investigating face emotion processing (see Table 2) used the previously described expression matching paradigm that was specifically designed to elicit amygdala activation (i.e., Hariri et al., 2000). In the most recent study using this approach, the researchers also manipulated the sex of the face stimuli (Telzer et al., 2015). Telzer and colleagues presented separate blocks of male and female angry, happy, and neutral adult faces to 9- to 16-year-old adolescents. The authors were interested in evaluating participants' differential responses to opposite- versus same-sex faces, regardless of expression. Using a whole-brain regression, they reported a positive association between pubertal status, from parent-reported PDS, and amygdala activation to opposite-sex faces. Specifically, processing emotional expressions from opposite- compared to same-sex adult faces was related to higher amygdala activation in individuals with more advanced pubertal status. This finding is fairly consistent with the predictions regarding the association between pubertal development and amygdala activation during facial emotion expression processing.

Similarly, Spielberg et al. (2014a, 2014b) presented the face expression-matching task with fear and angry adult faces to adolescent participants at age 11 and again at age 13. The researchers reported a positive correlation between increasing testosterone levels over time and an increase in the magnitude of expression-related activation in both the left amygdala and left NAcc (Spielberg et al., 2014a). In another study, and colleagues explored the relation between changes in pubertal development and in functional connectivity from the amygdala and rest of the brain over a two-year period (Spielberg et al., 2014b). They used a ROI-based psychophysiological interaction (PPI) analysis strategy with the amygdala as the PPI seed, and submitted these PPI maps to a whole-brain regression analysis with change in serum testosterone levels across the two-year period as the predictor. The authors reported that increasing testosterone (from time 1 to time 2) was associated with decreased functional connectivity between the bilateral amygdala and right centromedial orbitofrontal cortex (OFC) during the emotion expression compared to shape processing task. In other words, the functional coupling between the amygdala and OFC decreased as a function of increasing testosterone across this two-year period in the entire sample of male and female adolescents. The authors argued that this functional decoupling between the amygdala and OFC, which was associated with a rise in testosterone, might facilitate increased amygdala reactivity to emotion.

Two studies using the same emotion-matching task reported conflicting results regarding the relation between pubertal development and amygdala activation to emotion expressions. In the first study, the researchers tested adolescents when they were approximately 11 years old using adult faces in the task (Forbes et al., 2011). Pubertal development was evaluated by physical examination and participants were categorized

into an early or late stage of development for group level comparisons. The researchers reported that the more advanced pubertal group exhibited less amygdala and ventrolateral prefrontal cortex (vlPFC) activation to neutral faces compared to shapes than did the early puberty group. There was no difference between the groups when viewing the angry faces versus shapes. These findings stand in contrast to those from the previously described studies, which all report a positive association between pubertal development and amygdala activation during processing of emotion expression in adolescence (e.g., Telzer et al., 2015; Spielberg et al., 2014a, 2014b). In the second study, Ferri and colleagues used the same emotion-matching experimental paradigm, but included adolescent faces, and tested adolescent girls ages 8–15 years (Ferri et al., 2014). They combined adolescent- and parent-report PDS and Picture-Based Interview responses to compute a latent factor to measure pubertal development. Ferri and colleagues reported a negative association between perceived pubertal development, as measured by the latent factor, and amygdala activation to adolescent neutral faces, compared to shapes. Importantly, this negative association was also observed between amygdala activation to neutral faces compared to shapes and increasing age.

Two other face expression studies did not use the emotion-matching paradigm. In the first study, adolescent participants passively viewed images of whole-face emotional displays (i.e., adult facial expressions including anger, fear, happy, sad; Moore et al., 2012) during scanning at age 10 and age 13. The researchers evaluated activation from each expression versus a fixation in each scan. They reported relations between activation and pubertal development, as measured by self-report PDS, separately for each age. The central finding was that at 10 years of age, there was a positive association between individual differences in pubertal development and neural activation to emotional expressions in the amygdala, thalamus, and extrastriate cortex. By 13, these same positive associations were present and extended to the fusiform gyrus, temporal pole, vlPFC, and vmPFC. In the second study of facial expressions that did not use the expression-matching paradigm, the researchers used a social approach-avoidance task in which 14-year-old adolescents had to pull (approach) or push (avoid) a lever in separate blocks as they viewed images of happy and angry faces (Tyborowska et al., 2016). The authors reported that higher pubertal development, as measured by testosterone, was associated with stronger activation during the angry-approach and happy-avoid trials than the angry-avoid and happy-approach trials in the anterior prefrontal cortex but lower activation in the amygdala and pulvinar of the thalamus. The authors interpreted these results to reflect the influence of puberty, and testosterone specifically, on the developing neural substrates supporting emotional actions.

In conclusion, across these seven studies, the findings converge on a consistent locus of the effect because they all report that activation in the amygdala during processing of facial expressions is related to pubertal development. Importantly, this *convergence in the locus* of findings is consistent across measures of pubertal development (physical examination, self-report, hormone). However, there was no convergence across studies in the direction of the association between the metrics of pubertal development and functional activation. Specifically, three studies reported a positive association between measures of pubertal development and neural activation to displays of facial emotion in the amygdala (Moore et al., 2012; Spielberg et al., 2014b, Telzer et al., 2015). Two of these studies reported the association between pubertal staging and neural activation and one reported an association between testosterone and neural activation. In contrast, three studies reported a negative association between measures of pubertal development and neural activation to displays of facial emotion in the amygdala (Forbes et al., 2011; Ferri et al., 2014; Tyborowska et al., 2016). Two of these studies reported the association between pubertal staging measures and neural activation and one study reported an association between testosterone and neural activation. Finally, two studies investigated functional connectivity between the amygdala and other regions and found opposite associations with testosterone (Tyborowska et al., 2016; Spielberg et al., 2014a).

#### 4.3. Pubertal and functional brain development during social information processing

The 7 studies in this domain share a common interest in investigating the association between pubertal development and changes in neural responses during the evaluation of social information by adolescents. The studies vary with respect to the kinds of social information being evaluated; they include scenarios depicting 'social emotions', social and academic traits in oneself and a fictional character, social status words, and simulated experiences of peer acceptance and rejection. In spite of the surface level differences, these paradigms activate a common underlying neural circuitry in the 'social brain' in both adults (Saxe, 2006) and adolescents (Richardson et al., 2018), which includes the precuneus, temporoparietal junction (TPJ), anterior temporal pole, and middle, dorso-, and ventromedial prefrontal cortex.

Across studies, researchers consistently predict that these behavioral abilities *increase* with increasing pubertal development, leading to a clear directional hypothesis about the relation between neural activation and pubertal development. However, there are less clear hypotheses about the specific regions in which the predicted positive association between activation and measures of pubertal development will be expressed. Most of the researchers generally target regions in the broadly defined 'social brain'. To preview the findings in this domain, 6 of 7 studies (> 50%; 8 of 9 findings in total) reported a positive association between indices of pubertal development and functional activation during social information processing in adolescents (see Table 2; Goddings et al., 2012; Jankowski et al., 2014; Klapwijk et al., 2013; Masten et al., 2013; Pfeifer et al., 2013; Silk et al., 2014). This indicates convergence in the direction of findings. However, across these same studies, there is no convergence in the locus of this association in any particular region.

In the first of these social information processing studies, Goddings and colleagues scanned female adolescents (ages 11–14 years) as they read and rated scenarios designed to evoke social (embarrassment and guilt) or basic emotions (Godings et al., 2012). The researchers collected multiple measures of pubertal development, including several hormonal assays and multiple kinds of staging measures (See Table 1). They reported positive associations between DHEA, oestradiol levels, and neural activation for the social versus basic scenarios in the left anterior temporal cortex, but not between any of the staging measures of pubertal development and neural activation. In their subsequent investigation, the researchers explored the relation between pubertal hormones and functional connectivity between social brain regions (Klapwijk et al., 2013). The researchers used dmPFC as the seed for a PPI analysis of the social versus basic emotion functional activation contrast throughout the brain. This analysis revealed regions that were functionally connected with the dmPFC as it differentially responded to these two kinds of emotional scenarios. In spite of a previous finding that activation in the dmPFC was related to age (not pubertal development; Goddings et al., 2012), the authors ran a correlational analysis with this PPI map and oestradiol levels. They reported a significant positive association between oestradiol and connectivity between the dmPFC and right TPJ. In other words, they reported increasing functional connectivity between the right TPJ and dmPFC specifically during the processing of more complex emotional scenarios as a function of increasing pubertal development (as measured by oestradiol level).

Several research groups design experimental paradigms that engage social information processing related to peers, given the importance of peers in the lives of adolescents. For example, Pfeifer and colleagues asked participants to evaluate social and academic traits in themselves and a familiar fictional character (Pfeifer et al., 2013). In this task, participants read the phrase, "I am popular," and indicated whether the phrase described themselves or the fictional character with a yes/no response. Participants were scanned at age 10 and again at age 13. When contrasting the trials in which participants made self- versus character-evaluations, the researchers found an increase in activation over time in the vmPFC. In a post-hoc ROI-based analysis of this region, they reported a positive association between the

magnitude of change in activation with age and the change in pubertal status, as assessed by self-report PDS, particularly for social, but not academic, self-evaluations. However, complementary whole-brain analyses evaluating a relation between self-report pubertal stage and longitudinal change in brain activation did not converge with this ROI-based finding.

Another study used a similar trait-evaluation task, but invited adolescents to rate the academic, physical, and social traits of themselves and of a close peer (Jankowski et al., 2014). The researchers first used a whole-brain analysis approach that identified the ventral striatum (VS) as a region in which adults and adolescents differed in the magnitude of responses during this trait evaluation task. Then, in a post-hoc analysis of the parameter estimates extracted from the bilateral VS, they found a positive association between self-report pubertal stage and VS activation during the social self-evaluations.

In another peer evaluation task, Silk and colleagues showed participants images of similarly aged peers with fictitious biographical profiles and asked them to pick who they would like to interact with on a subsequent visit to the lab (Silk et al., 2014). Two weeks later during the fMRI scan, the adolescents believed that they were engaging in a live online chat task with the adolescents they picked. During the task, the adolescents were accepted or rejected for social interactions by the selected peers. The researchers reported that increased activation in both the left amygdala and caudate nucleus during rejection trials was positively related to more advanced adrenarche (i.e., pubic hair and body odor) as indicated by self-report PDS.

Masten and colleagues were interested in longitudinal changes in empathic processing toward peers in adolescence and the role that pubertal development might play in these changes (Masten et al., 2013). They measured empathic skills and pubertal development via self-report PDS in adolescents at age 10 and again at age 13. They used fMRI to scan the adolescents only at age 13. The participants passively viewed fictitious peers playing a collaborative game of Cyberball. As the game progressed, one peer was systematically rejected from the game. The researchers measured neural activation to observed peer rejection compared to observed peer collaboration. Using a whole-brain correlational analysis, the authors reported a positive association between neural activation to peer rejection and self-report PDS scores in the bilateral DMPFC, PCC/precuneus, TPJ, and temporal pole. They also reported a similar positive association between the magnitude of this activation to observed peer rejection at age 13 and the change in self-reported PDS scores from age 10 to age 13 in the dmPFC and temporal pole. However, ROI-based analyses did not reveal converging results.

Finally, Silk and colleagues asked 10- to 18-year-old adolescents to identify the emotional valence of social status words that were previously identified as words used by adolescents to name “popular” and “unpopular” peers (Silk et al., 2017). The authors reported an age-related increase in neural activation in the mPFC in response to social status compared to neutral words; however, they reported no association between self-reported measures of pubertal development (as assessed using the PDS) and such activation.

To summarize, across these studies that evaluated an association between neural activation elicited during social information processing and various measures of pubertal development, there is *convergence in the direction* of findings. Specifically, 6 of these 7 studies (8 out of 9 findings) reported a positive relation between measures of neural activation during social information processing tasks and pubertal development. Importantly, this convergence in directionality of findings was reported across multiple tasks of social information processing and multiple measures of pubertal development (i.e., hormones, PDS; see Table 1). In contrast, there is *no convergence* in the location of the findings across regions showing an association between neural activation and pubertal development. Across these 7 studies, many regions were implicated (ATC, dmPFC, vmPFC, TPJ, amygdala, caudate nucleus, VS, and insula), which may be related to the reliance on whole-brain analysis as the primary analytic approach.

#### 4.4. Pubertal and functional brain development during cognitive processing

Finally, there are 5 studies investigating the relation between metrics of pubertal development and functional activation during cognitive tasks. Across these studies, researchers employed a variety of cognitive tasks, including spatial working memory, information processing, feedback learning, and inhibitory control tasks. The diversity in components of cognitive processing and, therefore underlying neural systems across these studies (see Fig. 1d) prevented us from evaluating convergence in the locus of findings within this domain. To preview the findings, only two of five studies (i.e., < 50%) reported an association between some measure of pubertal development and functional neural activation during cognitive processing (Cservenka et al., 2015; Schweinsburg et al., 2015). As a result, there is no convergence in the directionality of findings regarding a potential association between pubertal and functional brain development in the cognitive processing domain.

Two studies investigated the association between metrics of pubertal development and neural activation elicited during spatial working memory tasks (SWM). Interestingly, despite the similarity in the paradigms, the studies reported very different findings. Alarcón and colleagues scanned 10- to 16-year-old adolescents as they completed a 2-back SWM task with letters in multiple locations (Alarcón et al., 2014). Activation in the SWM condition was compared to that elicited during a vigilance condition in which participants attended to and identified dots in the same locations. Although boys and girls did not differ in behavioral responses, there were sex differences in the overall patterns of activation elicited during the SWM task. However, differences in testosterone did not mediate these sex-specific brain responses. Also, testosterone was not related to activation in the left inferior parietal lobule, where sex differences in activation were observed. In contrast, Schweinsburg and colleagues used a similar SWM task with adolescents, but the participants had to remember the spatial location of abstract line drawings instead of letters (Schweinsburg et al., 2005). The researchers reported that adolescents’ self-report PDS scores were negatively related to memory activation in the right superior parietal lobe, which remained significant after controlling for the effects of age.

Two other studies in the cognitive domain were interested in evaluating the association between metrics of pubertal development and the neural activation elicited during very basic cognitive processes, including categorization and rule-learning. Brumback and colleagues instructed 8- to 13-year-olds to categorize stimuli in an event-related potentials odd-ball paradigm (Brumback et al., 2012). The authors reported a sex difference in the latency of the P300, an ERP component that corresponds with novelty detection. However, pubertal status, as measured by the self-report PDS, was not related to either the latency or the amplitude of the P300. Peters and colleagues investigated rule-learning among children, adolescents, and adults (Peters et al., 2014). Participants were presented with a series of objects and had to learn to map each object onto one of three locations on the basis of feedback from the researchers. Neither adolescents’ self-report PDS scores, testosterone, nor estradiol levels were related to activation in the frontoparietal network during the rule-learning task.

Finally, one group employed an inhibitory control task using face emotion expression stimuli to investigate associations between pubertal development and brain activation (Cservenka et al., 2015). In this study, adolescents completed an emotion Stroop task, which required them to identify the expression on a face in spite of emotion-congruent or incongruent words printed across the faces (e.g., sad face with the word “happy” or “sad” above it). The authors computed separate whole-brain maps contrasting emotion-incongruent versus emotion-congruent trials in boys and girls. They ran separate regressions on each map with testosterone and estradiol levels, while controlling for age. The authors reported that in boys testosterone levels were negatively related to emotion-incongruent activation in the putamen and middle frontal gyrus (MFG), while estradiol levels were negatively related to activation in the cingulate gyrus and a cerebellar region. In contrast, they reported that in girls testosterone was negatively related to cerebellar and precuneus

**Table 3**  
Recommendations for Investigating the Relation between Pubertal and Functional Brain Development in Adolescence.

Specify Theory-driven Hypotheses
<p><b>Four essential elements to specify in hypothesis:</b></p> <ol style="list-style-type: none"> <li>1. Which aspect of pubertal development (andrenarche, gonadarche)?</li> <li>2. Where in the brain?</li> <li>3. What kind of behavior/task condition?</li> <li>4. What is the directionality of association between pubertal development and neural activation?</li> </ol> <p><b>If start with behavior:</b></p> <ul style="list-style-type: none"> <li>• What behaviors change as a function of pubertal development</li> <li>• What brain regions/network support this behavior</li> <li>• Then investigate the relation between identified brain regions/network and pubertal development</li> </ul> <p><b>If start with brain:</b></p> <ul style="list-style-type: none"> <li>• Which brain regions contain hormone receptors?</li> <li>• What behaviors does this brain regions/network support?</li> <li>• Then investigate the relation between identified brain regions/network and pubertal development</li> </ul>
Pubertal Development Measures – Optimize to Research Question
<p><b>Hormones:</b></p> <ul style="list-style-type: none"> <li>• Provide clear rationale about which selected hormones are optimal measure of pubertal development for specific research question</li> <li>• Consider cyclic nature of hormones when sampling (hourly, daily, monthly)</li> <li>• Consider the difference in hormone level and functionality across males and females</li> <li>• Implement measurement strategy for assessing individual differences in hormone levels to determine whether developmental changes attributed to puberty are bigger than individual differences, especially in cross-sectional studies</li> <li>• Report data collection details (assay sensitivity, menstrual cycle, within- and between- subject reliability)</li> </ul> <p><b>Pubertal staging:</b></p> <ul style="list-style-type: none"> <li>• Provide rationale about why staging and which measure is used for specific research question</li> <li>• Determine physical exam vs. perceived self-/parent-report</li> <li>• Report information regarding training of examiners and inter-rater reliability when conducting physical exams</li> <li>• Report explicit staging criteria for examiners when conducting physical exams</li> <li>• Report explicit scoring strategies for self-/parent-report measures</li> <li>• Consider the frequency of measures to capture individual differences in tempo</li> </ul>
Experimental Controls
<p><b>Strategies for dissociating pubertal and age effects:</b></p> <ul style="list-style-type: none"> <li>• Match adolescents who different in puberty on age (e.g., 12-year-olds in early vs. late pubertal development)</li> <li>• Measure and report potential age differences between groups when using group comparisons</li> <li>• Add age as covariant in statistical models when comparing groups who differ in pubertal stage</li> <li>• Include within-subject conditions that dissociate effects of age and pubertal development</li> </ul> <p><b>Strategies for controlling for sex differences:</b></p> <ul style="list-style-type: none"> <li>• Balance the number of male and female participants in the sample (as a function of group)</li> <li>• If studying a single sex, rationalize the specific questions and hypothesis, narrow the interpretation of findings</li> </ul>
Neuroimaging Data Analysis
<p><b>Head motion:</b></p> <ul style="list-style-type: none"> <li>• Prospective strategies: mock scanning and PACE acquisition</li> <li>• Retrospective strategies: include motion parameters (e.g., framewise displacement) as covariates in analyses</li> <li>• Report head motion statistics and group comparisons</li> </ul> <p><b>False positive activation:</b></p> <ul style="list-style-type: none"> <li>• Employ best practices for false positive correction, including voxel-wise inference or permutation strategies</li> <li>• Ensure that sample sizes are sufficiently powered for whole-brain correlational analyses</li> </ul> <p><b>Conduct independent tests of effect sizes:</b></p>

**Table 3 (continued)**

<ul style="list-style-type: none"> <li>• Ensure that criteria for voxel selection and estimation of effect size are independent</li> <li>• Avoid selecting hypothesized areas after results are known (SHARKing)</li> <li>• Use a priori selected anatomically defined ROIs</li> <li>• Use separate localizer task to functionally define a priori selected ROIs</li> </ul>
---

activation, but estradiol levels were positively related to occipital activation.

In sum, only 2 of the 5 studies in the cognitive domain reported an association between metrics of pubertal development and functional neural activation elicited during cognitive processing (Cservenka et al., 2015; Schweinsburg et al., 2005). This lack of convergence in findings across studies may be due, in part, to the different underlying cognitive processes and supporting neural systems that were studied. However, even when highly similar tasks were used, researchers reported very different findings regarding the potential relation between metrics of pubertal development and functional brain development (see Alarcón et al., 2014; Schweinsburg et al., 2005).

#### 4.5. Conclusions from the review

Using a label-based meta-analytic approach, we analyzed 28 studies in total from the past 15 years in which developmental neuroscientists investigated the relation between various metrics of pubertal development and functional brain development. We critically evaluated the extent to which findings converge to reflect consistency in the location and/or the direction of the association in each of four functional domains, using > 50% (simple majority) as a criterion for convergence. We summarize this evidence as follows:

- 1 **Reward processing:** there is no convergence in the locus or directionality of the relation between pubertal development and functional activation during reward processing, in spite of the strong theoretical focus on the nucleus accumbens (i.e., NAcc) in this work.
- 2 **Facial emotion processing:** studies converge on the amygdala as a neural region where metrics of pubertal development are associated with neural activation during emotion processing. However, there are an equal number of findings reporting that this association is positive and negative. As such, the directionality of this relation is unclear.
- 3 **Social information processing:** there is a positive relation between metrics of pubertal development and functional activation during social information processing. However, the locus of this effect is highly inconsistent.
- 4 **Cognitive processing:** there is no convergence in the locus or directionality of the relation between pubertal development and functional activation during cognitive processing.

This set of conclusions is based on our quantification of the patterns of findings in the existing literature. As a result, our conclusions are somewhat different from those of other recent articles that summarily reviewed this literature (e.g., Goddings et al., 2019; Vijayakumar et al., 2018). Specifically, our label-based approach of quantifying convergence in findings in both the locus and directionality of associations between metrics of pubertal and functional brain development allow us to provide both general and specific claims about the status of the literature. As a result, we conclude that there are hints of convergence in findings within two functional domains (i.e., face emotion processing, social information processing). Importantly, there is no domain in which the majority of studies reach the highest criterion of convergence (i.e., in both locus and directionality of effects). This conclusion leads to important questions about this lack of convergence in findings. Note that the lack of convergence is not simply due to a series of null effects. What we found is *inconsistency in the specificity of the effects*.

We do acknowledge limitations of this label-based meta-analysis

approach, which involves binarizing each result (significant association or not). In so doing, each finding carries equal weight in the analysis. As a result, we did not differentially weight the findings on the basis of sample size or the quality of the data as is often done in traditional meta-analyses. However, because we were unable to do a meta-analysis of effect sizes here, we provided much of the relevant information that readers can use to evaluate findings themselves. Similarly, we did not weight the findings by the magnitude of the effects because these data are not available in the majority of the studies. Also, in the analysis of the locus of effects, we choose to collapse across directionality of effects as a first step. There were too few studies in the current literature to properly assess interactions between the location and direction of effects or to assess the potential influence of moderating variables like sex, pubertal measurement, or study design. This is a limitation of the literature, as opposed to of our approach. We hope that as this literature grows, these questions can be addressed with a similar analysis approach.

To conclude, this meta-analysis reveals that the limited convergence in findings is likely due to experimental and analytical factors. Going forward, we discuss several factors that we think are specifically problematic in this literature and how the field could address them in subsequent work. Our hope is that new work will uncover patterns of findings that do converge with respect to the locus and directionality of effects in ways that lead researchers to discover mechanisms linking pubertal and functional brain development.

## 5. Going forward

In this final section, we identify four experimental/analytic issues that are especially problematic for the extant literature and likely contribute to the relative lack of reproducible findings (see Table 3). Briefly these issues include, a critical need for theory-driven hypotheses to guide experimental design and analysis approaches in the study of pubertal and functional brain development; strategies to optimize the metric of pubertal development (e.g., which aspects measured, frequency of measurement) for the specific research question; methods for managing confounding effects, such as age and sex differences between groups; and, more rigor in the collection and analysis of neuroimaging data, particularly in terms of implementing protections against confounding effects of motion, false positive activations, and spurious correlations. We also provide strategies for how these issues can be addressed in future work.

### 5.1. Identify theory-driven hypotheses

Much of the language in the existing literature is causal in nature, describing the *influence* of pubertal development (e.g., pubertal hormones) on neural activation. However, this language is used in the absence of a clear hypothesis about a mechanism of action (i.e., which components of pubertal development influence neural activation in which parts of the brain). In the vast majority of the existing work, hypotheses about the relation between pubertal and functional brain development are stated at a general level (e.g., as pubertal development increases, so does brain activation) or are acknowledged to be preliminary and exploratory. For example, there is little explicit justification for focusing on gonadarche versus adrenarche as the mechanism of pubertal development that is potentially related to variations or changes in neural activation. A potentially serious consequence of these very general, non-specific and exploratory hypotheses is that the subsequent analysis strategy is vulnerable to *p*-hacking (Simmons et al., 2011) and *p*-HARKing (Kerr, 1998), which is susceptible to false positive findings.

While it is true that the work investigating the relation between pubertal and functional brain development in humans is in its infancy; there is a wealth of findings in the animal literature to draw upon to inform and constrain hypotheses (e.g., Schulz and Sisk, 2016, 2006). The articulation of clear theory-driven hypotheses about a causal relation between pubertal and functional brain development is essential for

guiding and organizing the experimental design and research/analysis approach of future studies. Going forward, researchers need to articulate *a priori* hypotheses about how specific aspects of pubertal development (e.g., adrenarche, gonadarche) influence the directionality of changes in neural activation (e.g., increase versus decrease) in specific regions/networks of the brain (e.g., amygdala versus vmPFC), and under what behavioral/task conditions (e.g., during face but not object processing).

In thinking about strategies for forming these more specific hypotheses, a useful place to start is with behavior. There is a relatively large body of literature investigating the relation between pubertal and behavioral developmental in adolescence (see review; Forbes and Dahl, 2010). For example, in the domain of face processing, we have found that biases in face recognition behavior (Picci and Scherf, 2016) and in perceptual sensitivity to socially complex but not basic facial expressions (Motta-Mena and Scherf, 2016) emerge as a function of pubertal development. Given these findings, a concrete hypothesis can be formed about developmental changes in the underlying neural circuitry (i.e., functional activation) that support these behavioral changes in face processing, which are also associated with changes in particular metrics of pubertal development. In other words, one strategy is to target specific behaviors that reliably change as a function of pubertal development and then identify the neural circuitry related to these specific behavioral changes. This approach will provide researchers with a clearer way to make concrete hypotheses about how specific aspects of pubertal development are associated with task-related functional activation in specific neural circuits.

Alternatively, researchers can form more concrete hypotheses about the relation between pubertal and functional brain development by targeting specific brain regions that are known to have sex hormone receptors. For example, animal work indicates that the hippocampus is dense with estrogen receptors and amygdala is dense with both estrogen and androgen receptors. The functional activation within these regions may be more directly impacted by changes in pubertal development given their sensitivity to the very hormones that drive gonadarche (see review; Scherf et al., 2012, 2013). The task for researchers is then to design tasks that elicit functional activation within these regions that can then be associated with variations or changes in pubertal development.

### 5.2. Optimizing the metric of pubertal development for the research question

There are two broad categories of metrics to assess pubertal development that are differentially optimized to address questions about how puberty might influence functional brain development. Staging metrics are designed to capture global consequences of the pubertal process at a given time. Therefore, measuring pubertal development via a staging metric is particularly good for addressing questions about how the general process of pubertal development is related to changes in brain function. Hormonal assays provide information about the mechanistic drivers of the pubertal process at a given moment in time. Therefore, measuring pubertal development via hormonal assay might be particularly useful for testing hypotheses about the ontogenesis of pubertal effects on brain function since *in vivo* concentrations of hormones likely increase long before measurable physical changes can be reported in pubertal staging procedures (Dorn et al., 2006).

Across the majority of studies that we reviewed in this literature, researchers did not provide a clear scientific rationale for the metrics of pubertal development that were employed in their studies. In this literature, the majority of the studies (75%) used an adolescent self-report or parent-report version of the Peterson Development Scale (PDS) as the primary metric of pubertal development. The reliance on this measure likely reflects that it is easy to administer, provides privacy for the adolescent/parent, and does not require special training or space to administer for the research staff.

However, there are significant limitations to this survey as a measure of the biological process of pubertal development, particularly because it was not designed to be a staging measure, although many studies use it in this way. The questions on the PDS capture the adolescent's or parent's

*perception of pubertal development.* As a result, it is recommended that studies using the PDS report the self-perception nature of this methodology, as in “perceived pubertal stage” (see Dorn et al., 2006). However, none of the studies in the existing literature that used the PDS describe the findings in this way. The self-perception nature of this measure is not trivial. This is especially evident when researchers attempt to use the PDS to assess Tanner staging. The original PDS scoring system was not designed to align with Tanner stages (see Petersen et al., 1988). Even when using the recent scoring algorithm that does attempt to translate PDS scores to Tanner stages (e.g., Shirtcliff et al., 2009), the correspondence between PDS Tanner stages and expert physical exam assessments of Tanner stages is only modest (Shirtcliff et al., 2009). Finally, many of the events described in the PDS occur late in the pubertal process (e.g., menarche, growth of facial hair), which limits its ability to capture early pubertal events (Dorn et al., 2006). Together, these findings indicate that the PDS should not be used if researchers are interested in having an assessment of specific Tanner stage for the metric of pubertal development.

Given the critical limitations of the PDS, we implore researchers to acquire Tanner staging information via expert physical exams, which is the gold standard for assessing Tanner staging of pubertal development. Some ways to make the physical exams more tolerable to adolescents include, providing a clear, concrete description of the exam during the consenting process, allowing adolescents the choice of having privacy with the examiner or having a parent accompany them during the exam, acquiring consent again through each step of the exam, explaining the process of pubertal development and normal variations of development during the exam, and showing adolescents where they are in pubertal growth charts. If physical examination by trained research personnel is not an option for pubertal staging, we recommend using the Sexual Maturation Scale (SMS; Morris and Udry, 1980) or the Picture-Based Interview about Puberty (PBIP; Dorn and Susman, 2002). These are also self-report staging measures, but they do map directly onto Tanner stages and also assess andrenarche and gonadarche staging separately. Adolescents and/or parents examine line drawings of models at each Tanner stage and indicate which one the adolescent most closely resembles. Although these are self-perception methods, they resolve all of the other limitations that plague the PDS. They assess specific Tanner stages separately for andrenarche and gonadarche beginning with the pre-pubertal stage and ending with the sexually mature stage. Also, the kappa correspondence between expert physical exam and the self-reported pubertal stage is in the moderate range (Coleman and Coleman, 2002; Shirtcliff et al., 2009).

Studies that choose to collect hormonal assays as the metric of pubertal development need to address the fact that multiple hormones are responsible for the process of pubertal development. This is important because a single hormone sampled at one time of day does not represent a global view about the status of pubertal development. This acknowledgement will greatly impact how researchers interpret findings of an association between hormone concentrations and neural activation. Also, most of the reviewed studies collected hormones via saliva, whereas some studies collected hormones via blood. The reliability in hormone measurement between saliva and blood is still largely unknown.

Perhaps most importantly, the majority of studies in this literature review that collected hormone samples collected a single sample and compared hormone concentrations between individuals. This approach can be problematic because hormone concentrations fluctuate within an individual on multiple temporal schedules (e.g., daily, monthly), which complicates the interpretation of a single hormone measurement. Perhaps more importantly, hormone concentrations can vary widely across individuals even within a pubertal stage (Dorn et al., 2006). These individual differences in hormone levels might be as large or larger than differences in hormone levels between pubertal stages (Shirtcliff et al., 2009). Therefore, a significant correlation between hormone concentrations and brain activation could merely reflect individual differences in this association within pubertal stages rather than an effect of increasing pubertal development on brain activation. For example, a similar effect might exist in sexually mature individuals. To address this potential

confound, researchers need to implement an experimental strategy for assessing individual differences in hormone levels. Perhaps assessing individual differences in hormone levels in a comparable sample of sexually mature adolescents could provide a benchmark from which to assess differences related to pubertal development.

Just over 50% of the studies in this literature collected hormonal assays as the primary or secondary metric of pubertal development. Of the studies that acquired hormones, 100% collected testosterone, 33% collected a form of estrogen, and 20% collected DHEA. Importantly, most of the studies did not provide a clear scientific rationale for their choice of hormone measurement. The high number of studies that collected testosterone may reflect the relative ease in collecting this hormone from both sexes, particularly in comparison to the relative difficulty in collecting estrogen from girls and boys. Estradiol sensitivity is particularly poor in pre- and peri-pubertal girls who often exhibit concentrations below the detection limit of the assay (Dorn and Biro, 2011). Estradiol levels change over the course of the menstrual cycle, which must be taken into consideration during data collection. This is especially difficult to accommodate in pubertal girls because many of them do not exhibit regularity in their cycles for almost 2 years following menarche (Grumbach and Styne, 2003; Dorn and Biro, 2011) and many girls are anovulatory (i.e., do not release an egg) during menstruation at this stage (Bulun and Adashi, 2011).

Researchers must also consider the frequency with which they will acquire the measurements of pubertal development. As puberty is a developmental process that unfolds over the course of approximately 8–10 years, longitudinal studies will be essential for understanding the dynamic and potentially causal interactions between pubertal development and functional brain development. A typical approach in longitudinal studies of adolescent development, as is the case in the largest federally-funded longitudinal study of adolescent brain development to date (Adolescent Brain Cognitive Development - see <https://addictionresearch.nih.gov/abcd-study>) is to assess brain function and/or pubertal development on a yearly or biennial basis. This approach is useful since it systematizes the timing of measurement across participants. However, annual or biennial measurement of pubertal development for studies primarily interested in questions about the influence of puberty on other processes (e.g., functional brain development) is likely problematic. This is because the tempo of pubertal development is non-linear and exhibits vast individual differences in boys and girls (Marceau et al., 2011). The tempo of pubertal development describes how quickly or slowly individuals progress through the pubertal process. For example, one study reported that individual differences in the tempo of boys' genital development ranged from 0.29 to 1.25 Tanner stages per year (Marceau et al., 2011). Similarly, in the same study, the tempo of girls' breast development ranged from 0 to 1.2 Tanner stages per year. Therefore, annual or biennial assessments of pubertal development are not likely to capture these individual differences.

Finally, because puberty is a process that is difficult to capture in a single measurement at a single point in time, researchers need to consider how they can assess *change* in this process and relate it to *change* in functional activation. An ideal strategy is to employ a longitudinal design with frequent, repeated assessments of pubertal development and functional activation so that growth curve analyses can be conducted at the individual level. However, this is often logistically difficult and too expensive for researchers. In the current literature, only 25% of the studies to date are longitudinal in nature and all of them only include two assessments of pubertal development, preventing the use of growth curve analyses. In cross-sectional designs, the notion of heterogeneity in variance across the sample or between groups becomes critical. For example, when including pre-pubescent children and/or sexually mature adults as comparison groups/individuals to adolescents there may be floor and/or ceiling effects in metrics of pubertal development (e.g., pubertal stages). As a result, analyses of the full spectrum of scores might need to include non-linear models. Importantly, researchers still need clear, theory-driven hypotheses about specific patterns of change in the association between brain activation and metrics of pubertal development, particularly if they are non-linear.

In conclusion, it is essential that researchers carefully select a metric

of pubertal development on the basis of the proposed mechanism that mediates the association between pubertal and functional brain development. This includes a justification about whether the particular focus is on measuring adrenarche or gonadarche. Much of the existing work appears to have selected metrics on the basis of ease of data collection (e.g., self-reported PDS), not on the basis of a proposed mechanism that links puberty and brain activation. It is also critical that researchers consider the frequency with which they measure pubertal development within individuals and how group comparisons will be evaluated. Consistency in measurement of the mechanisms of puberty as a process in more objective ways (versus perceived status from self-report) is likely to generate more consistency in findings across studies.

### 5.3. Excluding confounding effects of age and biological sex

Two of the most challenging aspects of studying the influence of puberty are related to the fact that its timing is fundamentally linked to age, making pubertal- and age-related effects difficult to disentangle, and that there are sex-specific trajectories and effects of the pubertal process. Given that pubertal- and age-related processes can influence behavior and functional brain activation in independent and/or complementary ways, it is critical to employ both experimental and analytical strategies to disentangle these two effects. Similarly, sex differences in either the trajectories and/or mechanisms of pubertal development and/or patterns of brain activation (in the absence of pubertal effects) can impact the assessment of the influence of pubertal development on functional brain development. Therefore, it is critical that researchers experimentally and analytically control for age- and sex-related effects when evaluating the association between metrics of pubertal development and functional brain activation.

Researchers in this literature were generally sensitive to the potential confounds of age-related effects; however, they primarily approached dealing with this confound in a post-hoc analytic way. More than 50% of the studies that we reviewed just included age as a covariate in their analyses. However, this approach only works to disassociate the influence of age- from puberty-related effects on the outcome variable (e.g., neural activation) if there is a systematic (and linear) relation between age and neural activation that can be statistically modeled. For example, if age has a quadratic influence on neural activation (e.g., high for pre-pubertal individuals and high for sexually mature individuals, low for peri-pubertal individuals), a linear covariate in a general linear model, will not disassociate the influence of age from pubertal effects on neural activation. A second, less common approach to address age effects, involved recruiting adolescents from a narrow age range (e.g., 11–14 years) in which there is extensive variability in the timing and tempo of pubertal development. Researchers taking this approach often group adolescents into broad phases of pubertal development (e.g., early versus later pubertal development) for group level comparisons. The goal is to match the groups on age, and observe differences in pubertal development. This minimizes the influence of age on group level differences and more cleanly tests the influence of pubertal development on group differences. However, researchers using this approach still need to evaluate and report potential age differences in the different pubertal groups. These results are often not reported. We suggest that researchers who use this experimental design approach also employ post-hoc analytical strategies to minimize the potential confounding effect of age from puberty. A third approach that we recommend but that has not been adopted in this literature, is to incorporate multiple within-subject conditions into the experimental design that are predicted to be differentially influenced by age and pubertal development. For example, researchers can evaluate neural activation during conditions that are not expected to be sensitive to age- or puberty-related differences and contrast them with neural activation during experimental conditions that are expected to be differentially sensitive to puberty- or age-related effects. This approach would reveal the relative sensitivity of the age- and puberty-related effects to the task. The difficulty in this approach is minimizing condition differences in difficulty level and stimulus characteristics that might

otherwise confound comparisons across conditions.

It is also important to consider that many aspects of pubertal development are sex-specific and, therefore, the impacts of pubertal development on functional brain development may be also sex-specific. For example, levels of testosterone in males are many times higher than in females. In studies including male and female participants, this requires careful sampling to acquire an even distribution of male and female participants across levels of pubertal development. Some of the studies we reviewed had more than 60% female participants in their sample (see Table 1). This lack of sex parity is problematic for several reasons. First, findings of a positive relation between pubertal and functional brain development may be driven largely by one sex in studies with a very skewed sample. For example, a study reporting an association between neural activation and estrogen in a sample of all female adolescents will surely not generalize to explain how pubertal development is related to functional brain development in boys. Second, without sex parity in the sample, one can unintentionally introduce potential sex differences into the design that interact with pubertal effects. For example, girls start pubertal development about one year earlier than boys on average (Marceau et al., 2011). One can imagine that when trying to match participants on age who vary in pubertal development, it is easier to recruit 11-year-old boys in the earlier stages of puberty than 11-year-old girls. As a result, a researcher could end up with one group of adolescents selected to be in early pubertal development (10 girls, 20 boys) with an average age of 11 and another group of adolescents (also age 11) selected to be in later pubertal development (20 girls, 10 boys). In this case, differences between the two groups could reflect a pubertal stage by sex interaction, rather than a main effect of pubertal development. Third, when researchers are using hormonal assays as the primary metric of pubertal development, it is essential to articulate sex-specific hypotheses about how and where the hormones will influence neural activation given that hormone receptors likely vary as a function of sex and these hormones have sex-specific actions. Finally, when researchers collect data from single-sex samples, it is essential that they provide a clear scientific rationale for studying a single-sex population, particularly in terms of understanding the effects of pubertal and functional brain activation given the limitations in the ability to generalize findings to both sexes.

### 5.4. Rigor in neuroimaging data collection and analyses

This body of literature investigating the relation between pubertal development and functional brain development was conducted over the last 15 years when the best practices for analyzing neuroimaging data have changed dramatically. We suggest that one of the biggest reasons for the lack of convergence in findings within domains has to do with these changing standards. Here, we highlight several critical methodological and analytic limitations that we observed in this literature (see Table 4). We provide recommendations about how to improve rigor in this work and hope that consistency in the use of these strategies in future work will lead to findings that build upon this early research to sort out where and how the processes of pubertal

**Table 4**  
Rigor of Neuroimaging Methods in Studies Investigating Puberty and Functional Brain Development.

Rigorous Methods Used	(%)
<b>fMRI Head Motion Control</b>	
Mock scanning	7.7%
Motion correction	84.6%
Include as covariate in analyses	34.6%
Evaluation of group differences	15.4%
<b>Data Analysis Strategies</b>	
Sufficient False Positive Threshold	19.1%
Independent Analysis	84.6%
Effect Size Reported	32.1%

development influence functional brain development.

#### 5.4.1. Head motion

Participant motion is a persistent methodological issue in human neuroimaging, particularly for developmental studies. Studies report that age is systematically related to the amount of motion in the scanner (Kaufmann et al., 2017). Head motion alters the uniformity of the magnetic field, which directly affects the initial magnetization and changes locations of distortions and signal drop out boundaries (Murphy et al., 2013). It also induces changes in steady state magnetization by changing the time between excitations in parts of tissue that have moved from one slice to the next (i.e., spin history effects), which can induce changes in signal intensity that are up to *two times the expected BOLD signal changes* (Muresan et al., 2005). As a result, motion can induce artifacts in the BOLD signal, and contribute to difficulties in the process of co-registration between functional and structural images. There are *prospective* and *retrospective* strategies for dealing with motion artefacts. The most common strategy employed in the studies reviewed here was to adjust for them *retrospectively* using motion correction algorithms (84.6% of studies) and/or to account for residual inaccuracies in these “corrected” data by including the detected motion traces into the fMRI analyses (34.6% of studies). However, it is important to note that 15% of the studies did not report using any strategy for dealing with motion (see Table 4). Also, modeling motion as a covariate in the analysis does not *correct for intra-volume motion or spin-history effects*, or for motion that is even weakly correlated with the task, all of which can produce false activations (Power et al., 2012).

We recommend that researchers employ both prospective and retrospective approaches to dealing with motion artefacts. There was very limited information from the studies we reviewed about the use of any prospective strategies for minimizing motion artefacts. For example, only about 8% of studies reported using some kind of procedure to train participants in a mock scanner prior to collecting data in the real scanner. Mock scanning reduces anxiety and motion in both pediatric and adult populations. We strongly encourage researchers to employ child-oriented scanning preparation procedures to minimize participant anxiety and motion prior to scanning (e.g., Barnea-Goraly et al., 2004; de Bie et al., 2010; Epstein et al., 2007; Raschle et al., 2012; Rosenberg et al., 1997; Scherf et al., 2015). In particular, we recommend that researchers prepare participants by simulating the scanning protocol using mock scanners prior to scanning participants in the experimental protocol (Scherf et al., in preparation).

Second, we recommend that researchers consider using prospective motion correction algorithms during fMRI data collection (e.g., PACE; Thesen et al., 2000). PACE is the only technique for fMRI applications that allows for adequate correction of spin-history effects (Yancey et al., 2011) and intra-volume distortions (Speck et al., 2006). The use of prospective motion correction via PACE together with retrospective motion correction is superior in fMRI analyses to either alone (Zaitsev et al., 2016). None of the studies reported using online motion correction algorithms.

In terms of *retrospective* motion correction strategies, we urge researchers to continue including motion estimates in their analyses of the fMRI data. In particular, we suggest that researchers consider using the normalized relative mean framewise displacement (FD) as a covariate in subsequent analyses because it is comparable to “scrubbing data” but, critically, it does not lead to the loss of data (Gotts et al., 2012). Also, it is essential that researchers report analyses of potential group differences in motion estimates (e.g., FD) in all neuroimaging studies since head motion is reportedly related to age (e.g., van Duijvenvoorde et al., 2014; Cservenka et al., 2015). Only 15% of the studies in this literature report analyses of motion between participant groups. Group differences in motion could potentially explain group differences in activation patterns (younger = early puberty = more motion = different activation). This is a huge potential confound of the existing work.

In sum, we recommend that researchers train participants in child-friendly protocols prior to scanning, use online motion-correction algorithms while collecting fMRI data, retrospectively correct for motion,

statistically evaluate and report group differences in motion and/or associations between motion and age or metrics of pubertal development, and then consider whether to include motion vectors in the statistical analysis of the BOLD responses. This multifaceted approach of prospective and retrospective strategies will dramatically reduce the likelihood that motion explains differences in activation profiles between those in early and late stages of pubertal development.

#### 5.4.2. Controlling for false positive activations

In fMRI data analyses, researchers typically apply general linear models within individual voxels across the whole brain or at the region of interest level, such that they conduct between several hundred to 100,000 statistical tests. Therefore, it is essential to control for the strong likelihood of false positive outcomes (activations) in these analyses. The standards for doing so have changed over the course of the last 15 years. Recently, an empirical study revealed that the previous standards for employing *cluster-wise* inference strategies for false positive correction have a very high false positive likelihood (Eklund et al., 2016). Although the vast majority of the studies reviewed in this paper employed some form of correction for false positive activation when identifying task-related neural activation (see Tables 2 and 4), we determined that only 19% had a sufficiently strong threshold as recommended by these best practices. That indicates that the vast majority of these studies are likely to have false positive activations. Here is an explicit example of why this literature is so vulnerable to false positive activations.

The most common false positive correction strategy implemented in this literature was the cluster-wise inference strategy. In this approach, initial clusters are identified, and the most common threshold was an uncorrected cluster threshold of  $p = 0.01$ . Then surviving clusters are compared at familywise error-corrected extent threshold, usually at  $p = 0.05$ . Eklund and colleagues estimated that the false positive rate for this approach (with 20 participants and a standard smoothing kernel) is likely to range from 5 to 25% (see Supp Fig. 3; Eklund et al., 2016). This means that up to 25% of the resulting voxels in the outcome maps can be false positive activations. If researchers using this approach start by identifying initial clusters using a threshold of  $p = 0.001$ , the false positive error rate generally stays down below 10% (see Supp Fig. 4; Eklund et al., 2016). However, a much safer approach is to use *voxel-wise* inference or nonparametric permutation tests (Eklund et al., 2016). Therefore, we recommend that researchers use these approaches for false positive correction.

Relatedly, similarly conservative false positive correction procedures need to be applied to the whole-brain, voxel-wise correlational analyses that researchers use to evaluate associations between measures of pubertal development and functional neural activation. First, researchers must verify that they have sufficient power to conduct these analyses. The sample size recommendation for neuroimaging studies whose primary objective is to detect these kind of whole brain correlations is a minimum of 40 participants (Yarkoni, 2009). This sample size provides 80% power to detect a moderate size brain-behavior correlation (i.e.,  $r = .40$ ) at an alpha of  $p < 0.0125$ ; Yarkoni, 2009). Several of the studies in this literature were underpowered to report whole-brain correlations according to this recommendation.

#### 5.4.3. Conducting independent tests of effect size

A common strategy for investigating the association between pubertal development and functional brain development in the existing literature is to correlate the metrics of pubertal development with a measure of neural activation within a region of interest (ROI). In neuroimaging data, it is essential that the criteria for selecting voxels for a ROI in which to estimate an experimental effect and the estimation of the actual effect (i.e., determining the magnitude of the effect size) be *independent*. Otherwise, the effect size is likely to be overestimated. For example, a researcher conducts an analysis of a whole brain interaction to identify clusters of voxels (i.e., ROIs) in which a late puberty group responds more to reward versus loss trials than does an early puberty group and identifies an ROI in medial frontal cortex. If the researcher

then extracts the beta weights from this region and submits them to a correlation/regression analysis with pubertal status again, the correlation will likely be inflated. This is because the estimation of the magnitude of the effect size is only within the limited set of voxels that already show the pubertal effect. By one estimation, the inflation is an average of  $r = .29$  higher than an independent analysis (Poldrack et al., 2017; Poldrack and Mumford, 2009). Across the studies we reviewed, 84% were successful in keeping the voxel selection process independent from the estimation of effect size. However, that means that 16% of the existing studies likely have inflated effect size estimates. Also, only 32% of studies reported effect size estimates (see Table 4).

There are multiple ways for researchers to avoid this circular analysis problem. A straightforward solution is to use anatomically defined regions of interest to select voxels. Another strategy is to include an independent functional scan (e.g., localizer scan) or separate runs of the same functional task that are specifically designed for the purpose of voxel selection. This approach is good for identifying specific ROIs for each individual participant, but also requires that researchers have a priori predictions about where to look for their effect. Importantly, the strength of these approaches is that the ROIs are delineated prior to the initiation of any analyses, which constrains exploratory analyses that can result in SHARKing (Poldrack et al., 2017).

#### 5.4.4. Estimating neural connectivity

The majority of existing studies focused on understanding where (i.e., which regions) in the brain associations between puberty and neural activation can be observed. Only four studies investigated how puberty might be associated with the functional integration of information across neural regions (Klapwijk et al., 2013; Spielberg et al., 2014a; Tyborowska et al., 2016; Ladouceur et al., 2018). We suggest that investigating the ways in which pubertal development is related to the changing functional organization of developing neural circuitry could be very fruitful. Specifically, researchers can now measure and quantify global and local patterns of functional (temporal synchrony) and/or effective (causal, directional) connectivity in neural networks. This approach could be extremely useful for testing theoretical hypothesis about the function of pubertal development on brain development. For example, one of the most prominent theoretical frameworks about the neural basis of risk-taking in adolescents, proposes that pubertal development is related to the relatively accelerated development of the limbic reward systems compared to the frontal control systems (Steinberg, 2008; Shulman et al., 2016). This hypothesis is ripe for investigation using connectivity analyses. In addition, there are now many studies describing age-related changes in functional and effective connectivity in adolescent neural networks that overlap with the timing of pubertal development. It will be important for the field to determine which of these effects might be related to the process of pubertal development.

## 6. Conclusion

Although there is a long history of studying the influence of pubertal hormones on brain function and structure in animal models (see Sisk and Zehr, 2005a, 2005b), similar research in human adolescents is still early in its own ontogeny. We reviewed the existing 28 studies in this field that have been primarily conducted in the last decade. To quantify the findings, we measured convergence in results within content domains (reward, facial emotion, social information, cognitive processing) in terms of the locus and directionality of effects. We report that facial emotion processing is the only content domain with convergence in the locus of effects, such that studies consistently find a relation between metrics of pubertal development and neural activation in the amygdala. Social information processing is the only content domain in which there is consistency across studies in the directionality of effects. Specifically, functional brain activation during a variety of social information tasks is consistently positively associated with measures of pubertal development in adolescents; however, these effects do not converge in any

particular locus of the brain. In contrast, there is no convergence in the locus or directionality of effects in either the reward or cognitive processing domains. These findings highlight important directions for scientists to pursue in future research.

Importantly, we reveal that this limited convergence in findings relating functional brain and pubertal development is not because of null findings or even the variety of experimental paradigms researchers employ. For example, in the social information processing domain, there is immense variety in paradigm, but convergence in the directionality of effects. In contrast, in the reward processing domain, there is high consistency in the experimental paradigm, and participant sample across studies, but no convergence in locus or directionality of effects. As a result, we argue that there are critical theoretical, methodological, and analytic issues that must be addressed in order to move the field forward. To tackle these issues, we suggest that this interdisciplinary work needs to be conducted by teams of scientists with complementary expertise in adolescent development, pubertal development, endocrinology, and pediatric neuroimaging.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This work was supported by a grant (RO1 MH112573-01) from the National Institute of Mental Health (NIMH) to KSS.

## References

- Alarcón, G., Cservenka, A., Fair, D., Nagel, B., 2014. Sex differences in the neural substrates of spatial working memory during adolescence are not mediated by endogenous testosterone. *Brain Res.* 1593, 40–54.
- Barnea-Goraly, Kwon, Menon, Eliez, Lotspeich, Reiss, 2004. White matter structure in autism: preliminary evidence from diffusion tensor imaging. *Biol. Psychiatry* 55 (3), 323–326.
- Braams, B., Peper, J., van der Heide, D., Peters, S., Crone, E., 2016. Nucleus accumbens response to rewards and testosterone levels are related to alcohol use in adolescents and young adults. *Dev. Cogn. Neurosci.* 17, 83–93.
- Braams, B., van Duijvenvoorde, A., Peper, J., Crone, E., 2015. Longitudinal changes in adolescent risk-taking: a comprehensive study of neural responses to rewards, pubertal development, and risk-taking behavior. *J. Neurosci.* 35 (18), 7226–7238.
- Bress, J., Smith, E., Foti, D., Klein, D., Hajcak, G., 2012. Neural response to reward and depressive symptoms in late childhood to early adolescence. *Biol. Psychol.* 89 (1), 156–162.
- Brown, J., Dunn, J., 1996. Continuities in emotion understanding from three to six years. *Child Dev.* 67 (3), 789–802.
- Brumback, Arbel, Donchin, Goldman, 2012. Efficiency of responding to unexpected information varies with sex, age, and pubertal development in early adolescence. *Psychophysiology* 49 (10), 1330–1339.
- Bulun, S., Adashi, 2011. The Physiology and Pathology of the Female Reproductive Axis. *Williams Textbook of Endocrinology*. Saunders, Philadelphia, PA, pp. 587–664 2003.
- Coleman, Coleman, 2002. The measurement of puberty: a review. *J. Adolesc.* 25 (5), 535–550.
- Cservenka, Stroup, Etkin, Nagel, 2015. The effects of age, sex, and hormones on emotional conflict-related brain response during adolescence. *Brain Cogn.* 99, 135–150.
- Dahl, R., 2004. Adolescent brain development: a period of vulnerabilities and opportunities. Keynote address. *Ann. N. Y. Acad. Sci.* 1021, 1–22.
- de Bie, H., Boersma, M., Wattjes, M., Adriaanse, S., Vermeulen, Oostrom, K., et al., 2010. Preparing children with a mock scanner training protocol results in high quality structural and functional MRI scans. *Eur. J. Pediatr.* 169 (9), 1079–1085.
- Dorn, Biro, 2011. Puberty and its measurement: a decade in review. *J. Adolesc.* 21 (1), 180–195.
- Dorn, L., Dahl, R., Woodward, H., Biro, F., 2006. Defining the boundaries of early adolescence: a user's guide to assessing pubertal status and pubertal timing in research with adolescents. *Appl. Dev. Sci.* 10 (1), 30–56.
- Dorn, L., Susman, E., 2002. Puberty Script: Assessment of Physical Development in Boys and Girls. Cincinnati Children's Hospital Medical Center., Cincinnati, OH.
- Dunning, Hajcak, 2007. Error-related negativities elicited by monetary loss and cues that predict loss. *NeuroReport* 18 (17), 1875.
- Eklund, A., Nichols, T., Knutsson, H., 2016. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci* 113 (28), 7900–7905.
- Epstein, J., Casey, Toney, S., Davidson, M., Reiss, A., Garrett, A., et al., 2007. Assessment

- and prevention of head motion during imaging of patients with attention deficit hyperactivity disorder. *Psychiatry Res.* 155 (1), 75–82.
- Ferri, J., Bress, J.N., Eaton, N.R., Proudfit, G., 2014. The impact of puberty and social anxiety on amygdala activation to faces in adolescence. *Dev. Neurosci.* 36 (3–4).
- Forbes, E., Dahl, R., 2010. Pubertal development and behavior: hormonal activation of social and motivational tendencies. *Brain Cogn.* 72 (1), 60–72.
- Forbes, E., Phillips, M., Silk, J., Ryan, N., Dahl, R., 2011. Neural systems of threat processing in adolescents: role of pubertal maturation and relation to measures of negative affect. *Dev. Neuropsychol.* 36 (4), 429–452.
- Forbes, Ryan, Phillips, Manuck, Worthman, Moyles, et al., 2010. Healthy adolescents' neural response to reward: associations with puberty, positive affect, and depressive symptoms. *J. Am. Acad. Child Adolesc. Psychiatry* 49 (2), 162–172 e5.
- Foti, Weinberg, Dien, Hajcak, 2011. Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: temporospatial principal components analysis and source localization of the feedback negativity. *Hum. Brain Mapp.* 32 (12), 2207–2216.
- Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., Benedetti, F., Abbamonte, M., Gasparotti, R., Barale, F., Perez, J., McGuire, P., Politi, P., 2009. Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J. Psychiatry Neurosci.* 34 (6), 418–432.
- Goddings, A., Beltz, A., Peper, J.S., Crone, E.A., Braams, B.R., 2019. Understanding the role of puberty in structural and functional development of the adolescent brain. *J. Res. Adolesc.* 29 (1), 32–53.
- Goddings, A.-L., Heyes, S., Bird, G., Viner, R., Blakemore, S.-J., 2012. The relationship between puberty and social emotion processing. *Dev. Sci.* 15 (6), 801–811.
- Gotts, Simmons, Milbury, Wallace, Cox, Martin, 2012. Fractionation of social brain circuits in autism spectrum disorders. *Brain* 135 (9), 2711–2725.
- Grumbach, M.M., Styne, D.M., 2003. Puberty: ontogeny, neuroendocrinology, physiology, and disorders. In: Larsen, P.R., Kronenberg, H.M., Melmed, S., Polonsky, K.S. (Eds.), *Williams Textbook of Endocrinology*, 10th ed. Elsevier Science, Philadelphia, PA, pp. 1115–1286.
- Grumbach, Styne, 1992. Puberty ontogeny, neuroendocrinology, physiology, and disorders. In: Wilson, J.D., Foster, D.W. (Eds.), *Williams Textbook of Endocrinology*. Saunders, Philadelphia (1992).
- Guyer, A., Monk, C., McClure-Tone, E., Nelson, E., Roberson-Nay, R., Adler, A., et al., 2008. A developmental examination of amygdala response to facial expressions. *J. Cogn. Neurosci.* 20 (9), 1565–1582.
- Hare, T., Tottenham, N., Galvan, A., Voss, H., Glover, G., Casey, 2008. Biological substrates of emotional reactivity and regulation in adolescence during an emotional go-nogo task. *Biol. Psychiatry* 63 (10), 927–934.
- Hariri, Bookheimer, Mazziotta, 2000. Modulating emotional responses: effects of a neocortical network on the limbic system. *Neuroreport* 11 (1), 43.
- Herba, C., Landau, S., Russell, T., Ecker, C., Phillips, M., 2006. The development of emotion-processing in children: effects of age, emotion, and intensity. *J. Child Psychol. Psychiatry* 47 (11), 1098–1106.
- Herba, C., Phillips, M., 2004. Annotation: development of facial expression recognition from childhood to adolescence: behavioural and neurological perspectives. *J. Child Psychol. Psychiatry* 45 (7), 1185–1198.
- Herting, Sowell, 2017. Puberty and structural brain development in humans. *Front. Neuroendocrinol.* 44, 122–137.
- 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.
- Kaufmann, T., Alnaes, D., Doan, N., Brandt, C., Andreassen, O., Westlye, L., 2017. Delayed stabilization and individualization in connectome development are related to psychiatric disorders. *Nat. Neurosci.* 20 (4), 513–515.
- Kerr, 1998. HARKing: hypothesizing after the results are known. *Personal. Soc. Psychol. Rev.* 2 (3), 196–217.
- Klapwijk, E., Goddings, A.-L., Heyes, S., Bird, G., Viner, R., Blakemore, S.-J., 2013. Increased functional connectivity with puberty in the mentalising network involved in social emotion processing. *Horm. Behav.* 64 (2), 314–322.
- Ladouceur, C.D., Kerestes, R., Schlund, M.W., Shirtcliff, E.A., Lee, Y., Dahl, R.E., 2018. Neural systems underlying reward cue processing in early adolescence: the role of puberty and pubertal hormones. *Psychoneuroendocrinology*.
- Laird, A.R., Eickhoff, S.B., Kurth, F., Fox, P.M., Uecker, A.M., Turner, J.A., Robinson, J.L., et al., 2009. ALE meta-analysis workflows via the brainmap database: progress towards a probabilistic functional brain atlas. *Front. Neuroinform.* 3, 23.
- Laird, A.R., McMillan, K.M., Lancaster, J.L., Kochunov, P., Turkeltaub, P.E., Pardo, J.V., Fox, P.T., 2005. A comparison of label-based review and ALE meta-analysis in the Stroop task. *Hum. Brain Mapp.* 25 (1), 6–21.
- Lawrence, K., Campbell, R., Skuse, D., 2015. Age, gender, and puberty influence the development of facial emotion recognition. *Front. Psychol.* 6, 761.
- Marceau, K., Ram, N., Houts, R., Grimm, K., Susman, E., 2011. Individual differences in boys' and girls' timing and tempo of puberty: modeling development with nonlinear growth models. *Dev. Psychol.* 47 (5), 1389–1409.
- Masten, C., Eisenberger, N., Pfeifer, J., Dapretto, M., 2013. Associations among pubertal development, empathic ability, and neural responses while witnessing peer rejection in adolescence. *Child Dev.* 84 (4), 1338–1354.
- McGovern, R., Andersen, J., Byrd, D., Mutter, K., Reilly, J., 2002. Cognitive efficiency on a match to sample task decreases at the onset of puberty in children. *Brain and Cog* 50 (1), 73–89.
- Moore, W., Pfeifer, J., Masten, C., Mazziotta, J., Iacoboni, M., Dapretto, M., 2012. Facing puberty: associations between pubertal development and neural responses to affective facial displays. *Soc. Cogn. Affect. Neurosci.* 7 (1), 35–43.
- Morris, Udry, 1980. Validation of a self-administered instrument to assess stage of adolescent development. *J. Youth Adolesc.* 9 (3), 271–280.
- Motta-Mena, N., Scherf, S., 2016. Pubertal development shapes perception of complex facial expressions. *Dev. Sci.* 20 (4). <https://doi.org/10.1111/desc.12451>.
- Muresan, Renken, JBMT, R, Duifhuis, 2005. Automated correction of spin-history related motion artefacts in fMRI: simulated and phantom data. *IEEE Trans. Biomed. Eng.* 52 (8), 1450–1460.
- Murphy, K., Birn, R., Bandettini, P., 2013. Resting-state fMRI confounds and cleanup. *Neuroimage* 80, 349–359.
- Op de Macks, Z., Bunge, S., Bell, O., Kriegsfeld, L., Kayser, A., Dahl, R., 2017. The effect of social rank feedback on risk taking and associated reward processes in adolescent girls. *Soc. Cogn. Affect. Neurosci.* 12 (2), 240–250.
- Op de Macks, Z., Bunge, S., Bell, O., Wilbrecht, L., Kriegsfeld, L., Kayser, A., Dahl, R., 2016. Risky decision-making in adolescent girls: the role of pubertal hormones and reward circuitry. *Psychoneuroendocrinology* 74, 77–91.
- Op de Macks, Z., Moor, B., Overgaauw, S., Gürolu, B., Dahl, R., Crone, E., 2011. Testosterone levels correspond with increased ventral striatum activation in response to monetary rewards in adolescents. *Dev. Cogn. Neurosci.* 1 (4), 506–516.
- Peters, S., Braams, B.R., Raijmakers, M.E., Koolschijn, C.P., Crone, E.A., 2014. The neural coding of feedback learning across child and adolescent development. *J. Cogn. Neurosci.* 26 (8), 1705–1720.
- Petersen, Crockett, Richards, Boxer, 1988. A self-report measure of pubertal status: reliability, validity, and initial norms. *J. Youth Adolesc.* 17 (2), 117–133.
- Pfeifer, J., Kahn, L., Merchant, J., Peake, S., Veroude, K., Masten, 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.
- Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I., 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage* 16 (2), 331–348.
- Picci, Scherf, 2016. From caregivers to peers: puberty shapes human face perception. *Psychol. Sci.* 27 (11), 1461–1473.
- Pine, D., Lissek, S., Klein, R., Mannuzza, S., Moulton, J., Guardino, M., Woldehawariat, G., 2004. Face-memory and emotion: associations with major depression in children and adolescents. *J. Child Psychol. Psychiatry* 45 (7), 1199–1208.
- Plant, 2002. Neurophysiology of puberty. *J. Adolesc. Health* 31 (6), 185–191.
- Poldrack, R., Baker, C., Durnez, J., Gorgolewski, K., Matthews, P., Munafò, M., 2017. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nature Review Neuroscience* 18 (2), 115–126.
- Poldrack, R., Mumford, J., 2009. Independence in ROI analysis: where is the voodoo? *Soc. Cogn. Affect. Neurosci.* 4 (2), 208–213.
- Power, J., Barnes, K., Snyder, A., Schlaggar, B., Petersen, S., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59 (3), 2142–2154.
- Radua, J., Mataix-Cols, D., 2012. Meta-analytic methods for neuroimaging data explained. *Biol. Mood Anxiety Disord.* 2 (1), 6.
- Raschle, N., Zuk, J., Ortiz-Mantilla, S., Sliva, D., Franceschi, A., Grant, E., et al., 2012. Pediatric neuroimaging in early childhood and infancy: challenges and practical guidelines. *Ann. N. Y. Acad. Sci.* 1252, 43–50.
- Richardson, H., Lisandrelli, G., Riobueno-Naylor, A., Saxe, R., 2018. Development of the social brain from age three to twelve years. *Nat. Commun.* 9 (1), 1027.
- Rosenberg, D.R., Sweeney, J.A., Gillen, J.S., Kim, J., Varanelli, M.J., O'Hearn, K.M., et al., 1997. Magnetic resonance imaging of children without sedation: preparation with simulation. *J. Am. Acad. Child Adolesc. Psychiatry* 36 (6), 853–859.
- Saxe, R., 2006. Uniquely human social cognition. *Curr. Opin. Neurobiol.* 16 (2), 235–239.
- Scherf, S., Barth, S., Elbich, D., Motta-Mena, N., Picci, G., Whyte, E., Dai, J., Griffin, J., 2018. Preventative Strategies for Reducing Motion Artifacts in High Resolution Human Imaging. In preparation.
- Scherf, S., Behrmann, M., Dahl, R., 2012. Facing changes and changing faces in adolescence: a new model for investigating adolescent-specific interactions between pubertal, brain and behavioral development. *Dev. Cogn. Neurosci.* 2 (2), 199–219.
- Scherf, S., Elbich, D., Minschew, N., Behrmann, M., 2015. Individual differences in symptom severity and behavior predict neural activation during face processing in adolescents with autism. *Neuroimage Clin.* 7, 53–67.
- Scherf, S., Smyth, J., Delgado, M., 2013. The amygdala: an agent of change in adolescent neural systems. *Horm. Behav.* 64 (2), 298–313.
- Schulz, K.M., Sisk, C.L., 2016. The organizing actions of adolescent gonadal steroid hormones on brain and behavioral development. *Neurosci. Biobehav. Rev.* 70, 148–158.
- Schulz, K.M., Sisk, C.L., 2006. Pubertal hormones, the adolescent brain, and the maturation of social behaviors: lessons from the Syrian hamster. *Mol. Cell. Endocrinol.* 254, 120–126.
- Schweinsburg, D., Nagel, J., Tarpel, F., 2005. fMRI reveals alteration of spatial working memory networks across adolescence. *J. Int. Neuropsychol. Soc.* 11 (5), 631–644.
- Shirtcliff, E., Dahl, R., Pollak, S., 2009. Pubertal development: correspondence between hormonal and physical development. *Child Dev.* 80 (2), 327–337.
- Shulman, E., Harden, P., Chein, J., Steinberg, L., 2015. Sex differences in the developmental trajectories of impulse control and sensation-seeking from early adolescence to early adulthood. *J. Youth Adolesc.* 44 (1), 1–17.
- Shulman, E., Smith, A., Silva, K., Icenogle, G., Duell, N., Chein, J., Steinberg, L., 2016. The dual systems model: review, reappraisal, and reaffirmation. *Dev. Cogn. Neurosci.* 17, 103–117.
- Silk, J., Siegle, G., Lee, K., Nelson, E., Stroud, L., Dahl, R., 2014. Increased neural response to peer rejection associated with adolescent depression and pubertal development. *Soc. Cogn. Affect. Neurosci.* 9 (11), 1798–1807.
- Silk, Lee, Kerestes, 2017. “Loser” or “Popular”? : neural response to social status words in adolescents with major depressive disorder. *Dev. Cogn. Neurosci.* 28, 1–11.
- Sisk, C.L., Zehr, J.L., 2005a. Pubertal hormones organize the adolescent brain and beha-

- viator. *Front. Neuroendocrinology* 26, 163–174.
- Sisk, C.L., Foster, D.L., 2004. The neural basis of puberty and adolescence. *Nat. Neurosci.* 7 (10), 1040–1047.
- Sisk, C.L., Zehr, J.L., 2005b. Pubertal hormones organize the adolescent brain and behavior. *Front. Neuroendocrinol.* 26 (3–4), 163–174.
- Simmons, Nelson, Simonsohn, 2011. False-positive psychology. *Psychol. Sci.* 22 (11), 1359–1366.
- Spear, L.P., 2000. The adolescent brain and age-related behavioral manifestations. *Neurosci. Biobehav. Rev.* 24 (4), 417–463.
- Speck, O., Hennig, J., Zaitsev, M., 2006. Prospective real-time slice-by-slice motion correction for fMRI in freely moving subjects. *Magnetic Resonance Mater Phys Biology Medicine* 19, 55.
- Spielberg, J.M., Forbes, E.E., Ladouceur, C.D., Worthman, C.M., Olino, T.M., Ryan, N.D., Dahl, R.E., 2014a. Pubertal testosterone influences threat-related amygdala-orbitofrontal coupling. *Soc. Cogn. Affect. Neurosci.* 10 (3), 408–415.
- Spielberg, J.M., Olino, T.M., Forbes, E.E., Dahl, R.E., 2014b. Exciting fear in adolescence: Does pubertal development alter threat processing? *Dev. Cogn. Neurosci.* 8, 86–95.
- Steinberg, L., Albert, D., Cauffman, E., Banich, M., Graham, S., Woolard, J., 2008. Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: evidence for a dual systems model. *Dev. Psychol.* 44 (6), 1764–1778.
- Steinberg, L., 2008. A social neuroscience perspective on adolescent risk-taking. *Dev. Rev.* 28 (1), 78–106.
- Tahmasebi, A.M., Artiges, E., Banaschewski, T., Barker, G.J., Bruhl, R., Buchel, C., et al., 2012. Creating probabilistic maps of the face network in the adolescent brain: a multicentre functional MRI study. *Hum. Brain Mapp.* 33, 938–957.
- Tanner, 1962. *Growth at Adolescence*. Blackwell, Oxford.
- Telzer, E., Flannery, J., Humphreys, K., Goff, B., Gabard-Durman, L., Gee, D., Tottenham, N., 2015. The cooties effect: amygdala reactivity to opposite- versus same-sex faces declines from childhood to adolescence. *J. Cogn. Neurosci.* 27 (9), 1–12.
- Thesen, S., Heid, O., Mueller, E., 2000. Prospective acquisition correction for head motion with image-based tracking for real-time fMRI. *Magnetic Resonance.*
- Thomas, L., Bellis, M., Graham, R., LaBar, K., 2007. Development of emotional facial recognition in late childhood and adolescence. *Dev. Sci.* 10 (5), 547–558.
- Tyborowska, A., Volman, I., Smeekens, S., Toni, I., Roelofs, K., 2016. Testosterone during puberty shifts emotional control from pulvinar to anterior prefrontal cortex. *J. Neurosci.* 36 (23), 6156–6164.
- van Duijvenvoorde, A., de Macks, Z., Overgaauw, S., Moor, B., Dahl, R., Crone, E., 2014. A cross-sectional and longitudinal analysis of reward-related brain activation: effects of age, pubertal stage, and reward sensitivity. *Brain Cogn.* 89, 3–14.
- Vijayakumar, N., de Macks, Z., Shirtcliff, E.A., Pfeifer, J.H., 2018. Puberty and the human brain: insights into adolescent development. *Neurosci. Biobehav. Rev.* 92 (2018), 417–436.
- Wager, T.D., Lindquist, M., Kaplan, L., 2007. Meta-analysis of functional neuroimaging data: current and future directions. *Soc. Cogn. Affect. Neurosci.* 2 (2), 150–158.
- Yancey, S.E., Rotenberg, D.J., Tam, F., Chiew, M., Ranieri, S., Biswas, L., Anderson, K.J., Baker, N.S., Wright, G.A., Graham, S.J., 2011. Spin-history artifact during functional MRI: potential for adaptive correction. *Med. Phys.* 38, 4634–4646.
- Yarkoni, 2009. Big Correlations in Little Studies: Inflated fMRI Correlations Reflect Low Statistical Power-Commentary on Vul et al. (2009). *Perspect. Psychol. Sci.* 4 (3), 294–298.
- Zaitsev, Burak, LeVan, Knowles, B.R., 2016. Prospective motion correction in functional MRI. *Neuroimage* 154, 33–42.
- Zuckerman, M., Eysenck, S., Eysenck, H., 1978. Sensation seeking in England and America: cross-cultural, age, and sex comparisons. *Journal of Consulting and Clinical Psychology* 46 (1), 139.