

Accepted Manuscript

Title: Sex-specific associations of testosterone with prefrontal-hippocampal development and executive function

Author: Tuong-Vi Nguyen Jimin Lew Matthew D. Albaugh
Kelly N. Botteron James J. Hudziak Vladimir S. Fonov D.
Louis Collins Simon Ducharme James T. McCracken



PII: S0306-4530(16)30999-4
DOI: <http://dx.doi.org/doi:10.1016/j.psyneuen.2016.12.005>
Reference: PNEC 3484

To appear in:

Received date: 2-7-2016
Revised date: 23-11-2016
Accepted date: 7-12-2016

Please cite this article as: {<http://dx.doi.org/>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Title: Sex-Specific Associations of Testosterone with Prefrontal-Hippocampal Development and Executive Function

Abbreviated title: Testosterone, Brain Structure, and Executive Function

Authors:

- 1) Tuong-Vi Nguyen, MD, MSc^a
- 2) Jimin Lew^b
- 3) Matthew D Albaugh, PhD^c
- 4) Kelly N Botteron, MD^{d,e}
- 5) James J Hudziak, MD^{c,e}
- 6) Vladimir S Fonov, PhD^f
- 7) D. Louis Collins, PhD^f
- 8) Simon Ducharme, MD, MSc^{f,g}
- 9) James T McCracken, MD^{e,h}

Affiliations:

^a Department of Psychiatry and Department of Obstetrics-Gynecology, McGill University Health Center (Royal Victoria Hospital at the Glen site), McGill University, Montreal, QC, Canada, H4A 3J1

^b Department of Psychology, McGill University, Montreal, QC, Canada, H4A 3J1

^c University of Vermont, College of Medicine, Burlington, VT, USA, 05405

^d Washington University School of Medicine, St. Louis, MO, USA, 63110

^e Brain Development Cooperative Group

^f McConnell Brain imaging Centre, Montreal Neurological Institute, Montreal, QC Canada H3A 2B4

^g McGill University Health Centre, Department of Psychiatry and Department of Neurology & Neurosurgery, McGill University, Montreal, QC, Canada, H3A 1A1

^h Department of Child and Adolescent Psychiatry, University of California in Los Angeles, Los Angeles, CA, USA, 90024

Communications should be directed to Dr. Tuong-Vi Nguyen (tuong.v.nguyen@mcgill.ca)
McGill University Health Center, Royal Victoria Hospital at the Glen site, 1001 Decarie C6.1188-4,
Montreal, QC, Canada, H4A 3J1
Phone number: (514) 299-3373

ACKNOWLEDGEMENTS

This work was supported by Federal funds from the National Institute of Child Health and Human Development, the National Institute on Drug Abuse, the National Institute of Mental Health, and the National Institute of Neurological Disorders and Stroke (Contract #s N01-HD02-3343, N01-MH9-0002, and N01-NS-9-2314, -2315, -2316, -2317, -2319 and -2320).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

HIGHLIGHTS

1. Testosterone is associated with prefrontal-hippocampal structural covariance in boys
2. Testosterone-related covariance is associated with lower executive function in boys
3. No effect of testosterone on prefrontal-hippocampal covariance was seen in girls
4. No effect of testosterone-related covariance on executive function was seen in girls

ABSTRACT (250, max 250 words)

Testosterone is thought to play a crucial role in mediating sexual differentiation of brain structures. Examinations of the cognitive effects of testosterone have also shown beneficial and potentially sex-specific effects on executive function and mnemonic processes. Yet these findings remain limited by an incomplete understanding of the critical timing and brain regions most affected by testosterone, the lack of documented links between testosterone-related structural brain changes and cognition, and the difficulty in distinguishing the effects of testosterone from those of related sex steroids such as of estradiol and dehydroepiandrosterone (DHEA). Here we examined associations between testosterone, cortico-hippocampal structural covariance, executive function (Behavior Rating Inventory of Executive Function) and verbal memory (California Verbal Learning Test-Children's Version), in a longitudinal sample of typically developing children and adolescents 6-22 yo, controlling for the effects of estradiol, DHEA, pubertal stage, collection time, age, handedness, and total brain volume. We found prefrontal-hippocampal covariance to vary as a function of testosterone levels, but only in boys. Boys also showed a specific association between positive prefrontal-hippocampal covariance (as seen at higher testosterone levels) and lower performance on specific components of executive function (monitoring the action process and flexibly shifting between actions). We also found the association between testosterone and a specific aspect of executive function (monitoring) to be significantly mediated by prefrontal-hippocampal structural covariance. There were no significant associations between testosterone-related cortico-hippocampal covariance and verbal memory. Taken together, these findings highlight the developmental importance of testosterone in supporting sexual differentiation of the brain and sex-specific executive function.

KEYWORDS: Gonadarche; Androgen; Cognition; Puberty; Adolescents; Human Brain

1. INTRODUCTION

The sex steroid testosterone is thought to play a crucial role in mediating sexual differentiation of brain structures. Previous investigations using experimental animal models have documented testosterone-related masculinization of limbic brain areas (Zuloaga et al., 2008), and we have previously documented limbic changes related to pubertal maturation in a large, longitudinal human sample (Hu et al., 2013). In particular, changes in hippocampal volumes were shown to be sex-specific, with a negative correlation with pubertal maturation in boys and a positive correlation in girls (Hu et al., 2013). There have also been several reports of testosterone-dependent, sex-specific changes in gray matter and cortical thickness across puberty in several regions involved in cognitive control and emotional regulation, from our group and others (Bramen et al., 2011; Koolschijn et al., 2014; Neufang et al., 2009; Nguyen et al., 2013a; Paus et al., 2010; Peper et al., 2011; Raznahan et al., 2010; Witte et al., 2010). Taken together, the current literature suggests that testosterone may alter the structural relationship, or covariance, between the cortex and hippocampus.

Further supporting the importance of testosterone in determining brain structure and function, studies of testosterone administration or deprivation have reported an overall beneficial effect of this hormone on cognitive control and mnemonic processes, at least in adults (Janowsky, 2006; Muller et al., 2005). For example, in rodents, androgen deprivation by gonadectomy impairs performance on hippocampal-dependent tasks such as verbal memory and executive function, whereas testosterone replacement normalizes performance, suggesting a testosterone-specific effect on hippocampal-related cognition (Edinger and Frye, 2004; Kritzer et al., 2001). In adult men, there is some evidence that these effects of testosterone may be related to increased brain activity in distinct areas such as the prefrontal cortex, whereas effects in women tend to be non-specific, associated with increased overall brain activity, and attributable to the vascular effects of testosterone rather than any specific effect of this hormone on brain structure and function (Fernandez et al., 2003; Redoute et al., 2005).

Potential mechanisms of action of testosterone on the CNS are complex and include both direct actions on the androgen receptor (AR), indirect actions on estrogen receptors through its conversion to estradiol, and genomic (nuclear ARs) as well as non-genomic (membrane ARs) pathways (Foradori et al., 2008; Zuloaga et al., 2008). Through these various mechanisms, testosterone may carry out both neuroprotective and neurotoxic actions. For example, relatively lower testosterone levels may act predominantly through nuclear ARs, leading to increased neuronal and glial survival, while higher, more extreme testosterone levels may act mostly through membrane ARs, causing faster, potentially detrimental changes in cellular function that have been linked to higher rates of apoptosis (Foradori et al., 2008). For example, testosterone treatment post-gonadectomy resulted in increased cholinergic cell count in the anterior cingulate cortex and increased cholinergic fiber density in the hippocampus (Nakamura et al., 2002). In contrast, administration of high doses of testosterone propionate have been associated with increased glutamatergic neurotoxicity and decreased cholinergic tone (Sotomayor-Zarate et al., 2011). Given the rapidly changing testosterone levels during development, it is unclear whether and over which age range testosterone would carry out neuroprotective vs. neurotoxic actions.

Limitations of the current literature include discrepant brain-hormone findings with regards to the critical timing and brain regions most affected by testosterone-related effects during development (Koolschijn et al., 2014), the predominant use of adult animal models and elderly human samples in studies of the cognitive effects of testosterone (Janowsky, 2006), and the difficulty in distinguishing the effects of testosterone from those of related sex steroids such as of estradiol and dehydroepiandrosterone (DHEA) (Nguyen et al., 2013a; Nguyen et al., 2016; Nguyen et al., 2013c). Therefore, a consistent link between testosterone, brain structure, and cognition in the developing brain has yet to be firmly established.

To address these limitations, we examined associations between testosterone, cortico-hippocampal structural covariance, and cognition, in a longitudinal sample of typically developing children and adolescents 6 to 22 years of age. First, we tested for associations between testosterone and cortico-

hippocampal structural covariance. Second, we examined the relationship between testosterone-related cortico-hippocampal covariance and tests of executive function and verbal memory. To our knowledge, our analyses stand out as unique from the extant literature.

2. METHODS AND MATERIALS

2.1 Sampling and Recruitment

The National Institutes of Health (NIH) MRI Study of Normal Brain Development is a multi-site project that aimed to provide a normative database to characterize healthy brain maturation. Subjects were recruited across the United States with a population-based sampling method seeking to achieve a representative sample in terms of income level, race and ethnicity (Evans, 2006). All experiments on human subjects were conducted in accordance with the Declaration of Helsinki. All procedures were carried out with the adequate understanding and written parental consent, as well as assent of the subjects (or consent, if ≥ 18 years old). Subjects underwent repeated magnetic resonance brain imaging (MRI) every 2 years, with a maximum of 3 scans over 4 years. The sample was limited to developmentally healthy children with rigorous exclusion criteria, described in detail elsewhere (Evans, 2006). In particular, any children with a current or past treatment for language disorder (simple articulation disorders not exclusionary); and a lifetime history of Axis I psychiatric disorder (except for simple phobia, social phobia, adjustment disorder, oppositional defiant disorder, enuresis, encopresis, nicotine dependency) were excluded from the study. After strict quality control of MRI data (see section 2.2) and the exclusion of scans without hormonal measurements or behavioral parameters, 218 subjects were used for hormonal-related analyses (324 scans) and 155 to 186 subjects (199 to 245 scans) for cognitive analyses, depending on data available for each cognitive test (see **Table 1** for more details).

2.2 Neuroimaging Measures

A three-dimensional T1-weighted (T1W) Spoiled Gradient Recalled (SPGR) echo sequence from 1.5 Tesla scanners was obtained on each participant, with 1mm isotropic data acquired sagittally from the

entire head for most scanners. In addition, T2-weighted (T2W) and proton density-weighted (PDW) images were acquired using a two-dimensional (2D) multi-slice (2mm) dual echo fast spin echo (FSE) sequence.

Fully automated analysis of whole-brain cortical thickness was done through the CIVET pipeline, developed at the Montreal Neurological Institute (MNI). First, a multistage quality control process was implemented, as described previously (Nguyen et al., 2013a; Nguyen et al., 2013b), excluding subjects with white or gray matter artifacts. All quality-controlled MR images were subsequently processed through the CIVET pipeline. These processing steps have been described at length in other publications (Nguyen et al., 2013a; Nguyen et al., 2013b).

Volumetric measures of the hippocampus were obtained from MRI data using a fully automated segmentation method validated in human subjects (Collins and Pruessner, 2010). This method utilizes a large, manually labeled MRI dataset ($n = 80$) of young healthy adults that serves as a template library (Pruessner et al., 2001). The manual segmentation was done by four different raters, and intra-class intra-rater and inter-rater reliability varied between $r=0.83$ for the right and $r=0.95$ for the left hippocampus (Pruessner et al., 2000). From this manual segmentation, a fully automated method was derived, characterized by label fusion techniques that combine segmentations from a subset of 'n' most similar templates. Specifically, each template is used to produce an independent segmentation of the subject using the ANIMAL pipeline (Collins and Evans, 1997), followed by a thresholding step to eliminate cerebrospinal fluid, which results in 'n' different segmentations. To fuse the segmentations at each voxel, a voting strategy is used; the label with the most votes from the 'n' templates is assigned to the voxel. Combining multiple segmentations minimizes errors and maximizes consistency between segmentations. When using $n = 11$ templates, the label fusion technique has been shown to yield an optimal median Dice Kappa of 0.886 and Jaccard similarity of 0.796 for the hippocampus (Collins and Pruessner, 2010).

2.3 Hormonal and Pubertal Measures

During each MRI visit, children provided two separate 1–3 cm³ samples of saliva, collected on the day of the scan, which were assayed by enzyme-linked immunosorbent assay (ELISA) methods, and the average results used as a measure of hormonal levels. The intra-assay and inter-assay coefficients of variation (COVs) were 6.1% and 13.5% for testosterone, 4.1% and 9.1% for estradiol and 6.5% and 16.2% for DHEA, respectively (Salimetrics ELISA, State College, PA; Salimetrics Salivary ELISA Kit, State College, PA). At the next MRI, a similar procedure was followed and the child again provided two separate saliva samples for hormonal measurements.

Salivary sampling measures the unbound, biologically active portions of circulating hormonal levels, which freely crosses the blood-brain barrier and is therefore more relevant to studies of brain-hormone associations than total plasma hormonal levels (Khan-Dawood et al., 1984; Worthman et al., 1990). Testosterone levels have been shown to follow diurnal and seasonal patterns in response to the pulsatile release of adrenocorticotrophic hormone and gonadotropin-releasing hormone, particularly in boys (Brambilla et al., 2009; Stanczyk, 2006). To control for this, we have included collection time, sex and season as covariates in hormonal-related analyses (see section 2.5). Finally, to measure pubertal maturation, the Pubertal Development Scale (PDS) was administered by a physician to all subjects included in this study (Petersen et al., 1988). This scale has been shown to have good reliability (coefficient alpha: 0.77) and validity ($r^2=0.61-0.67$) compared to physical examination (Petersen et al., 1988). During an interview with the child/adolescent, questions were asked about physical development. We computed a puberty variable consisting of 5 stages, representing increasing levels of physical maturity similar to Tanner staging, previously described (Nguyen et al., 2013a).

2.4 Cognitive Measures

Cognitive measures were administered each time the subject underwent a scan (during the same research visit). We selected measures of cognitive control (i.e. executive function) and mnemonic processes (i.e.

verbal memory), the Behavior Rating Inventory of Executive Function (BRIEF) and the California Verbal Learning Test –Children’s Version (CVLT). The BRIEF is a parent- and teacher-completed rating scale, developed to assess everyday manifestations of children’s executive control functions (Gioia et al., 2000a; Gioia et al., 2002b). It includes eight subscales: Inhibit (inhibiting distractions and interference); Emotional Control (emotional regulation); Shift (flexibly shifting to new actions); Working Memory (short-term memory); Initiate (initiating action at an appropriate time/context); Plan/Organization (anticipating, planning); Organization of Materials (getting the materials necessary for the planned actions); and Monitor (monitoring the action process through internal and external feedback) (Gioia et al., 2000a; Gioia et al., 2002b). These executive functions refer to a collection of related abilities that direct and control goal-oriented cognitive, behavioral and emotional function, and that can be differentiated from ‘primary’ abilities such as language, visuo-spatial and mnemonic processes (Gioia et al., 2000a; Gioia et al., 2002b). The main strengths of the BRIEF lie in its ability to differentiate between executive function and ‘primary’ cognitive functions as well as between various components of executive function (Gioia et al., 2000a; Gioia et al., 2002b). Further, its use of ecologically valid measurements allows for a ‘real-world’ snapshot of executive function that includes aspects of complex, everyday problem-solving demands (Gioia et al., 2000a; Gioia et al., 2002b). The BRIEF has demonstrated good reliability, with high test-retest reliability ($r \approx .88$ for teachers, $.82$ for parents), internal consistency (Cronbach's alphas $\approx .80 - .98$), and moderate correlations between parent and teacher ratings ($r \approx .32 - .34$) (Gioia et al., 2000b).

That being said, the debate about the adequate/optimal way to assess executive function is still very much ongoing (Gioia et al., 2000a; Gioia et al., 2002a; Gioia et al., 2002b; Gioia et al., 2001; Goldman-Rakic, 1987; Welsh et al., 1991). Some have supported the use of direct, performance-based testing of discrete components of executive functions, while others have suggested that most of these performance-based tests are inadequate in assessing executive function because they attempt to separate integrated functions into component parts (Burgess, 1997; Goldberg and Podell, 2000; Shallice and Burgess, 1991). Further,

current performance-based tests measure individual components of executive function over a short timeframe (Burgess, 1997; Goldberg and Podell, 2000; Shallice and Burgess, 1991). In this context, some authors have regarded performance-based tests as representing a somewhat skewed and limited version of ‘real-life’, ‘*in vivo*’ executive function. Therefore, we selected the BRIEF scale because it measures executive function in an integrated, multidimensional, relativistic way, outlining the complex, priority-based decision-making that is demanded in real-world situations (Gioia et al., 2000a; Gioia et al., 2002a; Gioia et al., 2002b; Gioia et al., 2001).

In contrast, the CVLT directly evaluates more ‘primary’, performance-based aspects of cognitive function, i.e. verbal learning and memory, including components of short-term and long-term delayed recall of verbal items (Baldo et al., 2002; Cattie et al., 2012; Goodman et al., 1999; Woods et al., 2006). This test measures performance with regards to semantic clustering, serial clustering, free vs. cued recall, perseveration and intrusion errors, response bias, response consistency, and learning slope, and yields several sub-scores, of which 5 are particularly relevant for verbal memory: long-delay (cued and free recall), short-delay (cued and free recall) and total number of words recognized. The CVLT is one of the most frequently used children’s measures of verbal learning and memory, and has high test-retest reliability: $r \approx 0.62-0.93$ (Mottram and Donders, 2005). The previously demonstrated construct validity and temporal stability of the CVLT also make it a measure of episodic verbal learning and memory supported by a considerable body of research (Jacobs and Donders, 2008).

The CVLT has been shown to predict ecological measures of scholastic and academic achievement, and specific CVLT indices correlated with parent reports of overall attention level (Muir-Broaddus et al., 2002). Thus, although it is a performance-based test, it is thought to have moderate to high ecological validity. In contrast, a measure of global executive functioning accounted for minimal variance in several CVLT indices (Hill et al., 2012), suggesting that using global ecological (e.g. parent- or teacher-) ratings may not adequately measure verbal learning. Further, to our knowledge, there are no validated ecological

tests to evaluate verbal learning. The use of CVLT is therefore preferred to test this specific cognitive parameter, as it builds on vast existing data demonstrating its clinical and ecological validity.

2.5 Statistical Analyses

Statistical analyses were done using SurfStat (Matlab toolbox designed by Keith J. Worsley) and SPSS 21.0 (SPSS, Inc., Chicago, Illinois). Please see **Table 2** for more details on statistical models used in this section.

2.5.1 Sample characteristics

We tested for differences in the complete sample used for brain analyses and the subsamples used for cognitive analyses using Chi-Squared goodness-of-fit and Student's T-tests, examining the distribution of scans, the number of males vs. females, age, total brain volume, pubertal stages, DHEA levels, collection times, handedness and hippocampal volumes. Differences across samples were considered significant at $p < 0.05$.

2.5.2 Testosterone-Related Cortico-Hippocampal Networks

Mixed effects designs were used to model the relationship between testosterone and covariance of the hippocampus with whole-brain, native-space cortical thickness (CTh), taking into account the within- and between- individual variances in this longitudinal sample, and controlling for the effects of age, sex, total brain volume, scanner, handedness, and time of salivary sampling. All continuous variables were centered using their respective means. A correction for multiple comparisons across the whole brain, using random field theory (RFT, $p < 0.05$), was applied to all analyses (Worsley et al., 1992). To examine associations between testosterone and structural covariance of the hippocampus, we examined the significance of the term 'Testosterone*Hippocampus', while controlling for all the aforementioned control variables. Because testosterone-related effects on brain structure have been consistently found to be sexually dimorphic, we also tested the term 'Testosterone*Hippocampus*Sex' for significance. We

examined testosterone-related structural covariance between whole-brain CTh and mean hippocampal volume, defined as the average volume of the left and right hippocampi. To examine any distinct effects of testosterone above and beyond those related to estradiol, DHEA, pubertal stage, or season of collection, these variables were also included as control variables in additional testosterone-related models. Finally, to test for age effects on the relationship between testosterone and cortico-hippocampal networks, we tested for interactions with age, i.e. ‘Testosterone*Hippocampus*Age’ and ‘Testosterone*Hippocampus*Age*Sex’ on whole-brain CTh.

2.5.3 Cortico-Hippocampal Networks, Executive Function and Verbal Memory

To examine associations between testosterone-related cortico-hippocampal networks and cognitive measures, we averaged CTh of brain regions found to be significant in section 2.5.2 (see **Table 2** for more details) and examined the impact of cortico-hippocampal covariance on cognitive measures, while controlling for all the aforementioned control variables, including estradiol, collection time, and season of collection. More specifically, we tested for sex-specific associations between prefrontal-hippocampal covariance and executive function (as measured on the BRIEF) and verbal memory (as measured on the CVLT-C), i.e. testing for the significance of the interaction term ‘CTh*Hippocampus*Sex’ on cognitive measures (‘CTh’ referring here to average cortical thickness of the prefrontal brain area identified in section 2.5.2, see **Table 2** for more details).

2.5.4 Mediation Effects of Cortico-Hippocampal Networks

We formally tested whether cortico-hippocampal covariance (covariance between the average CTh of the brain area identified in section 2.5.2 and mean hippocampal volume) could mediate the relationship between testosterone and the cognitive components found to be significant in section 2.5.3. To examine the relationship between testosterone and cortico-hippocampal networks, we extracted the coefficients and standard errors of the significant interaction term ‘Testosterone*Hippocampus*Sex’ for peak vertices (i.e. vertices with the highest coefficient; see Methods, 2.5.2 for more details). Similarly, to examine the

relationship between testosterone-related cortico-hippocampal networks and tests of executive function and verbal memory, we extracted the coefficients and standard errors of the significant interaction term ‘CTh*Hippocampus*Sex’ (see Methods, 2.5.3 for more details). It is important to emphasize that no additional analyses were run for these two sections of the mediation analysis other than those already described above (see Methods, 2.5.2 and 2.5.3). Additional analyses were only run to test the relationship between testosterone and cognition, using mixed effect models in SPSS, and controlling for age, sex and collection time. Testosterone can be converted to estradiol by the enzyme aromatase. Therefore, given the potential importance of estradiol in influencing the association between testosterone, cortico-limbic covariance, and executive function through indirect conversion effects, we also tested the association between estradiol, structural brain covariance, and cognition.

Extracted coefficients and standard errors were then entered in the Sobel-Goodman test calculator to formally test mediation effects (<http://quantpsy.org/sobel/sobel.htm>). This more traditional approach to test mediation and moderation effects, using Baron-Kenny’s criteria and augmented by a formal Sobel’s test, was preferred by our group to more recent methods that include bootstrapping. This is because of the complexity of our longitudinal data (multiple scans per subjects and different number of scans per subject). The traditional method treats each relationship (between predictor and moderator, and then between moderator and outcome) separately, allowing us to model the longitudinal component of the data. Finally, please note that the same set of control variables (as listed in Methods, 2.5.2) was used for the mediation analyses.

3. RESULTS

3.1 Sample Characteristics

Table 1 details sample characteristics, including number of longitudinal scans and covariates of interest. The sample used for testosterone-related analyses included 218 participants (F=125) and 324 scans (F=190), with an age range of 6 to 22 years (mean: 13.5 years +/- 3.6). The sample used for BRIEF

analyses included 186 subjects (F=109) and 245 scans (F=146), with an age range of 6 to 19 years (mean: 12.7 years \pm 3.2). Finally, the sample used for CVLT-C analyses included 155 subjects (F=91), 199 scans (F=118), with an age range of 6 to 16 years (mean: 11.7 years \pm 2.5). There were sex differences in testosterone levels, as previously reported (Nguyen et al., 2013a), but no sex difference in the levels of estradiol or DHEA, or in the cognitive test scores.

3.2 Testosterone-Related Cortico-Hippocampal Networks

As shown in **Figure 1** (top brain figures and middle graphs), whole-brain analyses, controlling for the effects of age, sex, total brain volume, scanner, handedness and collection time of salivary samples, revealed a sex-specific effect of testosterone, such that testosterone levels were significantly associated with the structural covariance between mean hippocampus (average of the right and left hippocampi) and CTh of the right dorsal anterior cingulate cortex (linear mixed models, Brodmann areas 32 and 33, $r=2.3 \times 10^{-1}$, $SE=5.6 \times 10^{-2}$, cluster-level $p=2.0 \times 10^{-7}$, peak vertex id 56429 [$x=1.9$, $y=30.7$, $z=10.8$]) in boys, but not girls. More specifically, in boys, lower testosterone levels were associated with a negative prefrontal-hippocampal covariance, and higher testosterone levels, with positive prefrontal-hippocampal covariance, while no significant effects of testosterone were seen in girls.

No other brain region met the threshold for significance (RFT, $p<0.05$). Adding estradiol, DHEA, pubertal stage, and season of sampling as control variables (one at a time, to limit decreases in power) resulted in identical findings (exactly the same prefrontal brain region was found to be significant). Finally, there were no significant interactions between testosterone, sex and age on cortico-hippocampal networks (no significant effects of ‘Testosterone*Hippocampus*Age’ and ‘Testosterone*Hippocampus*Age*Sex’ on CTh).

3.3 Cortico-Hippocampal Networks, Executive Function and Verbal Memory

As shown in **Figure 1** (bottom graphs), analyses controlling for the effects of age, sex, total brain volume, scanner and handedness, revealed a sex-specific effect of prefrontal-hippocampal covariance on the BRIEF Monitor (Executive Function) subscale ($r=4.2*10^{-2}$, $SE=1.6*10^{-2}$, $p=1.1*10^{-2}$) and the BRIEF Shift subscale ($r=3.0*10^{-2}$, $SE=1.4*10^{-2}$, $p=3.8*10^{-2}$). In boys, but not girls, lower BRIEF scores were associated with a positive covariance between prefrontal-hippocampal covariance (similar to the covariance seen with higher testosterone levels), and higher BRIEF scores were associated with negative prefrontal-hippocampal covariance (similar to the covariance seen with lower testosterone levels). Of note, no significant relationship emerged between prefrontal-hippocampal structural covariance and CVLT-C (Verbal Memory) scores ($p \geq 0.05$).

3.4 Mediation Effects of Cortico-Hippocampal Networks

Coefficients and standard errors were extracted for: (1) the relationship between testosterone and prefrontal-hippocampal covariance (as moderated by sex); (2) the relationship between prefrontal-hippocampal covariance and BRIEF scores (as moderated by sex); and (3) the relationship between testosterone and BRIEF scores (as moderated by sex), and used to formally test mediating effects of prefrontal-hippocampal covariance on the relationship between testosterone and the components of executive function found to be significant in section 3.3, using different versions of the Sobel-Goodman mediation test. Prefrontal-hippocampal covariance was found to significantly mediate the relationship between testosterone and BRIEF Monitor scores ($p < 0.05$; see **Figure 2**) but not BRIEF Shift scores ($p \geq 0.05$). Of note, estradiol was not associated with ratings of executive function, verbal memory or HPC/PFC covariance, making it unlikely that indirect effects of testosterone through estradiol could have accounted for these results. Taken together, these findings suggest that prefrontal-hippocampal structural covariance may mediate the relationship between testosterone and a specific component of executive function (BRIEF Monitor).

4. DISCUSSION

Results from this study show important and complex sex-specific associations between testosterone, prefrontal-hippocampal structural covariance and specific aspects of executive function, namely monitoring and shifting abilities. Prefrontal-hippocampal covariance varied as a function of testosterone levels, but only in boys. Boys also showed a specific association between positive prefrontal-hippocampal covariance (as seen at higher testosterone levels) and lower performance on specific components of executive function, i.e. monitoring of the action process through internal and external feedback and flexible shifting between actions. Finally, we found the association between testosterone and a specific aspect of executive function (monitoring) to be significantly mediated by prefrontal-hippocampal structural covariance.

These specific alterations in executive function related to testosterone-dependent prefrontal-hippocampal covariance are consistent with observations that prefrontal-hippocampal interactions regulate working and reference memory (which relate to monitoring processes) and response perseveration (which relate to shifting abilities) (Yoon et al., 2008). In addition, our findings are consistent with previously documented correlations between hippocampal-neocortical ratios and inter-species differences in executive function (Shultz and Dunbar, 2010).

In contrast, it was relatively surprising to us, though not entirely unexpected, to find detrimental effects of testosterone-related prefrontal-hippocampal covariance on monitoring and shifting abilities (in boys only), as well as no effects of testosterone-related structural covariance on verbal memory in either sex. Indeed, the standing assumptions about the beneficial effects of testosterone on cognition are based on adult animal samples and human studies involving older men (Janowsky, 2006), which may not be applicable to this typically developing sample of children and adolescents ranging from 6-22 years of age. Although studies in children and adolescents are scarce, at least one study has not found any effects of testosterone replacement on psychomotor speed and switching between cognitive sets in the men at the

younger end of the age spectrum in an older male cohort (Haren et al., 2005). Similarly, another study found no beneficial cognitive effects of testosterone supplementation in young adult men (Bhasin et al., 2001). In addition, another study found treatment with testosterone to result in improved verbal memory for men if, and only if, conversion to estradiol occurred (Cherrier et al., 2005). In terms of developmental samples, our group previously found levels of aggressive behavior to vary according to testosterone-related prefrontal-amygdala covariance (possibly related to disinhibition and release from top-down inhibition) from childhood to young adulthood, which may compound the effects of prefrontal-hippocampal covariance on monitoring and shifting abilities, at least in boys (Nguyen et al., 2016). Taken together, these findings suggest that testosterone may have a specific detrimental effect on these specific aspects of executive function (by way of altering prefrontal-limbic covariance) during middle childhood and adolescence, possibly due to its effects on behavioral activation.

Testosterone may act through nuclear or membrane androgen receptors (AR) or non-AR pathways (including, but not limited to, its conversion to estradiol) (Foradori et al., 2008). The existing literature suggests that testosterone has beneficial effects on cholinergic function (and presumably, cognition) through non-AR pathways. Indeed, there is no significant co-localization of ARs and cholinergic neurons, even though testosterone deprivation does lead to a decrease in cholinergic cell density (Nakamura et al., 2002). Thus it is likely that conversion to estradiol is a necessary step for a testosterone-related improvement in cholinergic, and perhaps, cognitive function. In contrast, AR-dependent effects of testosterone during development may have detrimental effects on cognitive function, as seen in this study.

Here we found testosterone to be associated with positive prefrontal-hippocampal covariance, even after controlling for estradiol. Taken together with previous evidence that ER-related effects of testosterone tends to *improve* cognition, the current findings are consistent with the interpretation that testosterone may act predominantly through ARs to modify prefrontal-hippocampal structural covariance and *impair*

monitoring and shifting abilities, at least during middle childhood and adolescence. However, it is important to note that our study cannot directly address the mechanisms underlying the effects of testosterone, and therefore any mechanistic hypotheses remain speculative.

Differences in the distribution of aromatase in the brain may underlie, in part, this study's findings of testosterone-related changes in cortical thickness in a very localized region of the prefrontal cortex as opposed to more global/lobar effects (Biegon et al., 2010). However, it is unclear whether this mechanism would be responsible for the reported sex differences in this developmental sample. Indeed, in contrast to previous reports from animal studies (Abdelgadir et al., 1994; Roselli and Resko, 2001), up to now the vast majority of human studies have found no sex differences in CNS levels of brain aromatase activity and gene expression (Biegon et al., 2010; Ishunina et al., 2005; Sasano et al., 1998; Steckelbroeck et al., 1999; Stoffel-Wagner et al., 1999), with the exception of a recent study demonstrating small, albeit significantly higher CNS aromatase levels in adult men compared to adult women (Biegon et al., 2015).

Detrimental effects on executive function may be related to a testosterone-related decrease in synaptic pruning, or conversely, an increase in intra-cortical myelination in the anterior cingulate cortex (Kolb et al., 2012; Koolschijn et al., 2014; Nguyen et al., 2013a). In addition, it is possible that the steep rise in testosterone levels in adolescent boys related to testicular maturation (compared to the relatively lower levels of testosterone produced by the ovaries and the adrenals in girls) creates a specific male sensitivity to its behavioral activation effects, perhaps through an excessive activation of membrane ARs, or a rapid shift in the density of steroid receptors. This increased sensitivity may in turn put adolescent boys at higher risk for a temporary decline in cognitive performance during the pubertal transition. Temporary, sex-specific dips in cognitive function during puberty have been previously documented, though they were found to be limited to the 11-17 year-old range (Blakemore and Choudhury, 2006; McGivern et al., 2002). Our results support and extend these findings by documenting a testosterone-related dip in

specific aspects of executive function in boys that last from middle childhood to young adulthood (6-22 years old).

The absence of testosterone-related effects in girls is not entirely unexpected, given that testosterone administration has been previously linked to a nonspecific increase in overall brain activity in women (likely related to the vascular effects of testosterone, rather than specific effects on brain structure or function) (Fernandez et al., 2003; Redoute et al., 2005). Further, our group also found effects of testosterone on cortical maturation to be vastly different in boys compared to girls (i.e. testosterone-related decreases in cortical thickness in several regions of the left hemisphere in post-pubertal boys, vs. increases in thickness in one restricted region of the right hemisphere in pre-pubertal girls) (Nguyen et al., 2013a). Finally, at least one randomized controlled trial has documented a lack of cognitive effects (i.e. spatial ability, verbal fluency and memory, executive function) of testosterone administration in women (Huang et al., 2015). Thus, even though there have been relatively few studies of the effects of endogenous or exogenous testosterone in women, there is some evidence to suggest that effects of testosterone are more potent in boys than girls, at least during middle childhood and adolescence. This may be because a certain testosterone threshold has to be reached before significant effects on brain structure and cognition can be seen. There is in fact some evidence that entirely different mechanisms of action may predominate depending on the levels of testosterone, with opposite effects on cell survival and apoptosis (Foradori et al., 2007). Further, sexual dimorphism in the hippocampus and dorsal anterior cingulate cortex become particularly significant during adolescence (Blakemore et al., 2010), increasing the divergent responses of the male vs. female CNS to testosterone.

4.1 Limitations

The scans were done on 1.5T scanners, which have lower resolution compared to newer 3T models. This could lead to decreased accuracy and misclassification of white vs. gray matter. Still, all quality-controlled structural MR images were processed using the highest standards (see section 2.2). An

additional concern in hormone-brain association studies is the presence of intra-individual hormonal variation due to known, or unknown, causes. Reassuringly, androgen levels have been shown to remain highly correlated for several days, weeks and possibly even an entire year, and to reliably correlate with stable measures of personality (Dabbs Jr, 1990; Granger et al., 2004; Sellers et al., 2007). Still, to limit any systematic bias related to intra-individual hormonal variation, we have controlled for sex, diurnal and seasonal variation, with no significant changes in the results. Another limitation is the lack of umbilical cord or amniotic measurements of testosterone in this study, preventing us from evaluating the effects of prenatal testosterone exposure on brain structure. However, recent observations have identified continuing and distinct organizational effects of testosterone throughout childhood and adolescence, highlighting the importance of postnatal structural brain changes related to testosterone, above and beyond those restricted to the prenatal period (Nguyen et al., 2013b).

The Baron-Kenny approach used to test mediation effects also has two additional limitations: low power (thus low likelihood of detecting a significant effect) and the assumption that we can infer the indirect effect (the effect of the mediator) by evaluating only the direct effects (from independent variable to mediator, and from mediator to outcome). The first limitation could be viewed as a strength in the context of this study because it actually makes this approach more conservative (and increases the confidence one may have in the actual results). In terms of the second limitation, it is difficult to avoid because of the complexity of our data (longitudinal data, multiple scans per subject, not the same number of scans for each subject), which prevents the use of structural equation modeling approaches.

Finally, boys with poorer testosterone-related executive function were shown here to persevere and have difficulty switching between tasks. In some settings, such as courtship, these aptitudes may in fact be adaptive rather than detrimental. However, because the BRIEF assesses executive function in an ecologically valid, ‘real-world’ setting (as opposed to more classical performance-based tests delivered in a psychology lab), the likelihood that the current results actually reflect improved executive function

related to testosterone-related structural covariance is small. On the other hand, the use of the BRIEF brings an additional limitation to this study because it measures executive function abilities as perceived by the child's parents and teachers, rather than being a direct, lab-based assessment of executive function *per se*. As such, results of our study may not be easily generalizable to those of other studies using performance-based tests of executive function.

4.2 Conclusions

Testosterone may play a sex-specific role in the regulation of executive function in boys, through structural disruption in prefrontal-hippocampal covariance. Detrimental effects of testosterone on executive function may be a hallmark of developmental changes in prefrontal-hippocampal covariance, and represent an effect specific to middle childhood and adolescence. Specific components of executive function (monitoring and shifting of responses) appear to be disproportionately affected by testosterone-related structural covariance, consistent with the importance of prefrontal-hippocampal connections in working memory and perseveration. Taken together, these findings highlight the developmental importance of testosterone in supporting sexual differentiation of the brain and sex-specific executive function.

ROLE OF FUNDING

The funding primarily served to support the Brain Development Cooperative Group and was dedicated to protocol development, data collection and image processing.

***Appendix A

The Brain Development Cooperative Group is a multi-site research group formed from personnel from several pediatric study centers.

Data was collected from 6 sites across the United States:

(1) Children's Hospital Medical Center of Cincinnati, Cincinnati, OH, USA, 45229:

Principal Investigator William S. Ball, M.D., Investigators Anna Weber Byars, Ph.D., Mark Schapiro, M.D., Wendy Bommer, R.N., April Carr, B.S., April German, B. A., and Scott Dunn, R.T.;

(2) Children's Hospital Boston, Boston, MA, USA, 02115:

Principal Investigator Michael J. Rivkin, M.D., Investigators Deborah Waber, Ph.D., Robert Mulkern, Ph.D., Sridhar Vajapeyam, Ph.D., Abigail Chiverton, B.A., Peter Davis, B.S., Julie Koo, B.S., Jacki Marmor, M.A., Christine Mrakotsky, Ph.D., M.A., Richard Robertson, M.D., and Gloria McAnulty, Ph.D.;

(3) University of Texas Health Science Center at Houston, Houston, TX, USA, 77030: :

Principal Investigators Michael E. Brandt, Ph.D., Jack M. Fletcher, Ph.D., and Larry A. Kramer, M.D., Investigators Grace Yang, M.Ed., Cara McCormack, B.S., Kathleen M. Hebert, M.A., and Hilda Volero, M.D.;

(4) Washington University in St. Louis, St. Louis, MO, USA, 63110:

Principal Investigators Kelly Botteron, M.D. and Robert C. McKinsty, M.D., Ph.D., Investigators William Warren, Tomoyuki Nishino, M.S., C. Robert Almli, Ph.D., Richard Todd, Ph.D., M.D., and John Constantino, M.D.;

(5) University of California Los Angeles, Los Angeles, CA, USA, 90024:

Principal Investigator James T. McCracken, M.D., Investigators Jennifer Levitt, M.D., Jeffrey Alger, Ph.D., Joseph O'Neil, Ph.D., Arthur Toga, Ph.D., Robert Asarnow, Ph.D., David Fadale, B.A., Laura Heinichen, B.A., and Cedric Ireland B.A.;

(6) Children's Hospital of Philadelphia, Philadelphia, PA, USA, 19104:

Principal Investigators Dah-Jyuu Wang, Ph.D. and Edward Moss, Ph.D., Investigator Robert A. Zimmerman, M.D., and Research Staff Brooke Bintliff, B.S., Ruth Bradford, and Janice Newman, M.B.A.

In addition, the Brain Development Cooperative Group also included: a data coordinating center, a neurostatistics center, a clinical coordinating center, a diffusion tensor processing center, a scientific review center and a spectroscopy processing center:

(1) Data Coordinating Center, McGill University, Montreal, QC, Canada, H3A 1A1:

The Principal Investigator is Alan C. Evans, Ph.D., Investigators are Rozalia Arnaoutelis, B.S., G. Bruce Pike, Ph.D., D. Louis Collins, Ph.D., Gabriel Leonard, Ph.D., Tomas Paus, M.D., and Alex Zijdenbos, Ph.D., and Research Staff are Samir Das, B.S., Vladimir Fonov, Ph.D., Luke Fu, B.S., Jonathan Harlap, Ilana Leppert, B.E., Denise Milovan, M.A., and Dario Vins, B.C., and at Georgetown University, Thomas Zeffiro, M.D., Ph.D. and John Van Meter, Ph.D.

(2) Neurostatistics Laboratory, Harvard University/McLean Hospital, Belmont, MA, USA, 02478:

Investigators include Nicholas Lange, Sc.D. and Michael P. Froimowitz, M.S., who work with data coordinating center staff and all other team members on biostatistical study design and data analyses.

(3) Clinical Coordinating Center, Washington University in St. Louis, St. Louis, MO, USA, 63110:

The Principal Investigator is Kelly Botteron, M.D., Investigators C. Robert Almli Ph.D., Cheryl Rainey, B.S., Stan Henderson M.S., Tomoyuki Nishino, M.S., William Warren, Jennifer L. Edwards M.S.W., Diane Dubois R.N., Karla Smith, Tish Singer and Aaron A. Wilber, M.S.

(4) Diffusion Tensor Processing Center, National Institutes of Health, Bethesda, MD, USA, 20892: The Principal Investigator is Carlo Pierpaoli, M.D., Ph.D., Investigators Peter J. Basser, Ph.D., Lin-Ching Chang, Sc.D., Chen Guan Koay, Ph.D. and Lindsay Walker, M.S.

(5) Scientific Review, National Institutes of Health, Bethesda, MD, USA, 20892: The Principal Collaborators are Lisa Freund, Ph.D. (NICHD), Judith Rumsey, Ph.D. (NIMH), Lauren Baskir, Ph.D. (NIMH), Laurence Stanford, Ph.D. (NIDA), and Karen Sirocco, Ph.D. (NIDA) and from NINDS, Katrina Gwinn-Hardy, M.D. and Giovanna Spinella, M.D.

(12) Spectroscopy Processing Center, University of California Los Angeles, Los Angeles, CA, USA, 90024: The Principal Investigator is James T. McCracken, M.D.; Investigators are Jeffry R. Alger, Ph.D., Jennifer Levitt, M.D., and Joseph O'Neill, Ph.D.

REFERENCES

- Abdelgadir, S., Resko, J., Ojeda, S., Lephart, E., McPhaul, M., Roselli, C., 1994. Androgens regulate aromatase cytochrome P450 messenger ribonucleic acid in rat brain. *Endocrinology* 135, 395-401.
- Baldo, J.V., Delis, D., Kramer, J., Shimamura, A.P., 2002. Memory performance on the California Verbal Learning Test-II: findings from patients with focal frontal lesions. *J Int Neuropsychol Soc* 8, 539-546.
- Bhasin, S., Woodhouse, L., Casaburi, R., Singh, A.B., Bhasin, D., Berman, N., Chen, X., Yarasheski, K.E., Magliano, L., Dzekov, C., Dzekov, J., Bross, R., Phillips, J., Sinha-Hikim, I., Shen, R., Storer, T.W., 2001. Testosterone dose-response relationships in healthy young men. *Am J Physiol Endocrinol Metab* 281, E1172-1181.
- Biegon, A., Alexoff, D.L., Kim, S.W., Logan, J., Pareto, D., Schlyer, D., Wang, G.J., Fowler, J.S., 2015. Aromatase Imaging with [N-Methyl-C-11]Vorzole PET in Healthy Men and Women. *Journal of Nuclear Medicine* 56, 580-585.
- Biegon, A., Kim, S., Alexoff, D., Jayne, M., Carter, P., Hubbard, B., King, P., Logan, J., Muench, L., Pareto, D., Schlyer, D., Shea, C., Telang, F., Wang, G., Xu, Y., Fowler, J., 2010. Unique distribution of aromatase in the human brain: in vivo studies with PET and {N-methyl-11C}vorzole. *Synapse* 64, 801-807.
- Blakemore, S.J., Burnett, S., Dahl, R.E., 2010. The role of puberty in the developing adolescent brain. *Hum Brain Mapp* 31, 926-933.
- Blakemore, S.J., Choudhury, S., 2006. Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry* 47, 296-312.
- Brambilla, D.J., Matsumoto, A.M., Araujo, A.B., McKinlay, J.B., 2009. The Effect of Diurnal Variation on Clinical Measurement of Serum Testosterone and Other Sex Hormone Levels in Men. *Journal of Clinical Endocrinology & Metabolism* 94, 907-913.
- Bramen, J.E., Hranilovich, J.A., Dahl, R.E., Forbes, E.E., Chen, J., Toga, A.W., Dinov, I.D., Worthman, C.M., Sowell, E.R., 2011. Puberty Influences Medial Temporal Lobe and Cortical Gray Matter Maturation Differently in Boys Than Girls Matched for Sexual Maturity. *Cereb Cortex* 21, 636-646.
- Burgess, P., 1997. Theory and methodology in executive function research. Psychology Press, East Sussex, UK.

Cattie, J.E., Woods, S.P., Arce, M., Weber, E., Delis, D.C., Grant, I., 2012. Construct validity of the item-specific deficit approach to the California verbal learning test (2nd Ed) in HIV infection. *Clin Neuropsychol* 26, 288-304.

Cherrier, M.M., Matsumoto, A.M., Amory, J.K., Ahmed, S., Bremner, W., Peskind, E.R., Raskind, M.A., Johnson, M., Craft, S., 2005. The role of aromatization in testosterone supplementation - Effects on cognition in older men. *Neurology* 64, 290-296.

Collins, D.L., Evans, A.C., 1997. Animal: Validation and Applications of Nonlinear Registration-Based Segmentation. *International Journal of Pattern Recognition and Artificial Intelligence* 11, 1271-1294.

Collins, D.L., Pruessner, J.C., 2010. Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion. *Neuroimage* 52, 1355-1366.

Dabbs Jr, J.M., 1990. Salivary testosterone measurements: Reliability across hours, days, and weeks. *Physiology & Behavior* 48, 83-86.

Edinger, K.L., Frye, C.A., 2004. Testosterone's analgesic, anxiolytic, and cognitive-enhancing effects may be due in part to actions of its 5alpha-reduced metabolites in the hippocampus. *Behav Neurosci* 118, 1352-1364.

Evans, A.C., 2006. The NIH MRI study of normal brain development. *NeuroImage* 30, 184-202.

Fernandez, G., Weis, S., Stoffel-Wagner, B., Tendolkar, I., Reuber, M., Beyenburg, S., Klaver, P., Fell, J., de Greiff, A., Ruhlmann, J., Reul, J., Elger, C.E., 2003. Menstrual cycle-dependent neural plasticity in the adult human brain is hormone, task, and region specific. *J Neurosci* 23, 3790-3795.

Foradori, C.D., Weiser, M.J., Handa, R.J., 2008. Non-genomic actions of androgens. *Front Neuroendocrinol* 29, 169-181.

Foradori, C.D., Werner, S.B., Sandau, U.S., Clapp, T.R., Handa, R.J., 2007. Activation of the androgen receptor alters the intracellular calcium response to glutamate in primary hippocampal neurons and modulates sarco/endoplasmic reticulum calcium ATPase 2 transcription. *Neuroscience* 149, 155-164.

Gioia, G.A., Isquith, P.K., Guy, S.C., Kenworthy, L., 2000a. Behavior rating inventory of executive function. *Child Neuropsychol* 6, 235-238.

Gioia, G.A., Isquith, P.K., Guy, S.C., Kenworthy, L., Baron, I.S., 2000b. Test review: Behavior rating inventory of executive function. *Child Neuropsychology* 6, 235-238.

Gioia, G.A., Isquith, P.K., Kenworthy, L., Barton, R.M., 2002a. Profiles of everyday executive function in acquired and developmental disorders. *Child Neuropsychology* 8, 121-137.

Gioia, G.A., Isquith, P.K., Retzlaff, P.D., Espy, K.A., 2002b. Confirmatory factor analysis of the Behavior Rating Inventory of Executive Function (BRIEF) in a clinical sample. *Child Neuropsychol* 8, 249-257.

Gioia, G.A., Isquith, P.K., Retzlaff, P.D., Pratt, B.M., 2001. Modeling executive functions with everyday behaviors: A unitary or fractionated system? *Brain and Cognition* 47, 203-207.

Goldberg, E., Podell, K., 2000. Adaptive decision making, ecological validity, and the frontal lobes. *J Clin Exp Neuropsychol* 22, 56-68.

Goldman-Rakic, P., 1987. Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. Oxford University Press, NY.

Goodman, A.M., Delis, D.C., Mattson, S.N., 1999. Normative data for 4-year-old children on the California Verbal Learning Test-Children's Version. *Clin Neuropsychol* 13, 274-282.

Granger, D.A., Shirtcliff, E.A., Booth, A., Kivlighan, K.T., Schwartz, E.B., 2004. The “trouble” with salivary testosterone. *Psychoneuroendocrinology* 29, 1229-1240.

Haren, M.T., Wittert, G.A., Chapman, I.M., Coates, P., Morley, J.E., 2005. Effect of oral testosterone undecanoate on visuospatial cognition, mood and quality of life in elderly men with low-normal gonadal status. *Maturitas* 50, 124-133.

Hill, B.D., Alosco, M., Bauer, L., Tremont, G., 2012. The Relation of Executive Functioning to CVLT-II Learning, Memory, and Process Indexes. *Applied Neuropsychology-Adult* 19, 198-206.

Hu, S., Pruessner, J.C., Coupe, P., Collins, D.L., 2013. Volumetric analysis of medial temporal lobe structures in brain development from childhood to adolescence. *NeuroImage* 74, 276-287.

Huang, G., Wharton, W., Trivison, T.G., Ho, M.H., Gleason, C., Asthana, S., Bhasin, S., Basaria, S., 2015. Effects of testosterone administration on cognitive function in hysterectomized women with low testosterone levels: a dose-response randomized trial. *J Endocrinol Invest* 38, 455-461.

Ishunina, T., Van Beurden, D., Van Der Meulen, G., Unmehopa, U., Hol, E., Huitinga, I., Swaab, D., 2005. Diminished aromatase immunoreactivity in the hypothalamus, but not in the basal forebrain nuclei in Alzheimer's disease. *2005 Neurobiol Aging*, 173-194.

Jacobs, M.L., Donders, J., 2008. Performance discrepancies on the California Verbal Learning Test - Second Edition (CVLT-II) after traumatic brain injury. *Archives of Clinical Neuropsychology* 23, 113-118.

Janowsky, J.S., 2006. Thinking with your gonads: testosterone and cognition. *Trends Cogn Sci* 10, 77-82.

Khan-Dawood, F.S., Choe, J.K., Dawood, M.Y., 1984. Salivary and plasma bound and "free" testosterone in men and women. *Am J Obstet Gynecol* 148, 441-445.

Kolb, B., Mychasiuk, R., Muhammad, A., Li, Y., Frost, D.O., Gibb, R., 2012. Experience and the developing prefrontal cortex. *Proc Natl Acad Sci U S A* 109 Suppl 2, 17186-17193.

Koolschijn, P.C., Peper, J.S., Crone, E.A., 2014. The influence of sex steroids on structural brain maturation in adolescence. *PLoS ONE* 9, e83929.

Kritzer, M.F., McLaughlin, P.J., Smirlis, T., Robinson, J.K., 2001. Gonadectomy impairs T-maze acquisition in adult male rats. *Horm Behav* 39, 167-174.

McGivern, R.F., Andersen, J., Byrd, D., Mutter, K.L., Reilly, J., 2002. Cognitive efficiency on a match to sample task decreases at the onset of puberty in children. *Brain and Cognition* 50, 73-89.

Mottram, L., Donders, J., 2005. Construct validity of the California Verbal Learning Test - Children's Version (CVLT-C) after pediatric traumatic brain injury. *Psychological Assessment* 17, 212-217.

Muir-Broaddus, J.E., Rosenstein, L.D., Medina, D.E., Soderberg, C., 2002. Neuropsychological test performance of children with ADHD relative to test norms and parent behavioral ratings. *Archives of Clinical Neuropsychology* 17, 671-689.

Muller, M., Aleman, A., Grobbee, D.E., de Haan, E.H., van der Schouw, Y.T., 2005. Endogenous sex hormone levels and cognitive function in aging men: is there an optimal level? *Neurology* 64, 866-871.

Nakamura, N., Fujita, H., Kawata, M., 2002. Effects of gonadectomy on immunoreactivity for choline acetyltransferase in the cortex, hippocampus, and basal forebrain of adult male rats. *Neuroscience* 109, 473-485.

Neufang, S., Specht, K., Hausmann, M., Gunturkun, O., Herpertz-Dahlmann, B., Fink, G.R., Konrad, K., 2009. Sex Differences and the Impact of Steroid Hormones on the Developing Human Brain. *Cereb Cortex* 19, 464-473.

Nguyen, T.V., McCracken, J., Ducharme, S., Botteron, K.N., Mahabir, M., Johnson, W., Israel, M., Evans, A.C., Karama, S., 2013a. Testosterone-related cortical maturation across childhood and adolescence. *Cereb Cortex* 23, 1424-1432.

Nguyen, T.V., McCracken, J.T., Albaugh, M.D., Botteron, K.N., Hudziak, J.J., Ducharme, S., 2016. A testosterone-related structural brain phenotype predicts aggressive behavior from childhood to adulthood. *Psychoneuroendocrinology* 63, 109-118.

Nguyen, T.V., McCracken, J.T., Ducharme, S., Cropp, B.F., Botteron, K.N., Evans, A.C., Karama, S., 2013b. Interactive Effects of Dehydroepiandrosterone and Testosterone on Cortical Thickness during Early Brain Development. *Journal of Neuroscience* 33, 10840-10848.

Nguyen, T.V., McCracken, J.T., Ducharme, S., Cropp, B.F., Botteron, K.N., Evans, A.C., Karama, S., 2013c. Interactive effects of dehydroepiandrosterone and testosterone on cortical thickness during early brain development. *J Neurosci* 33, 10840-10848.

Paus, T., Nawaz-Khan, I., Leonard, G., Perron, M., Pike, G.B., Pitiot, A., Richer, L., Susman, E., Veillette, S., Pausova, Z., 2010. Sexual dimorphism in the adolescent brain: Role of testosterone and androgen receptor in global and local volumes of grey and white matter. *Horm Behav* 57, 63-75.

Peper, J.S., Hulshoff Pol, H.E., Crone, E.A., van Honk, J., 2011. Sex steroids and brain structure in pubertal boys and girls: a mini-review of neuroimaging studies. *Neuroscience* 191, 28-37.

Petersen, A., Crockett, L., Richards, M., Boxer, A., 1988. A Self-Report Measure of Pubertal Status: Reliability, Validity, and Initial Norms. *J Youth Adolescence* 17, 117-133.

Pruessner, J.C., Collins, D.L., Pruessner, M., Evans, A.C., 2001. Age and gender predict volume decline in the anterior and posterior hippocampus in early adulthood. *Journal of Neuroscience* 21, 194-200.

Pruessner, J.C., Li, L.M., Serles, W., Pruessner, M., Collins, D.L., Kabani, N., Lupien, S., Evans, A.C., 2000. Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories. *Cereb Cortex* 10, 433-442.

Raznahan, A., Lee, Y., Stidd, R., Long, R., Greenstein, D., Clasen, L., Addington, A., Gogtay, N., Rapoport, J.L., Giedd, J.N., 2010. Longitudinally mapping the influence of sex and

androgen signaling on the dynamics of human cortical maturation in adolescence. *Proc Natl Acad Sci U S A* 107, 16988-16993.

Redoute, J., Stoleru, S., Pugeat, M., Costes, N., Lavenne, F., Le Bars, D., Dechaud, H., Cinotti, L., Pujol, J.F., 2005. Brain processing of visual sexual stimuli in treated and untreated hypogonadal patients. *Psychoneuroendocrinology* 30, 461-482.

Roselli, C., Resko, J., 2001. Cytochrome P450 aromatase (CYP19) in the non-human primate brain: distribution, regulation, and functional significance. *J Steroid Biochem* 79, 247-253.

Sasano, H., Takahashi, K., Satoh, F., Nagura, H., Harada, N., 1998. Aromatase in the human central nervous system. *Clin Endocrinol* 48, 325-329.

Sellers, J.G., Mehl, M.R., Josephs, R.A., 2007. Hormones and personality: Testosterone as a marker of individual differences. *Journal of Research in Personality* 41, 126-138.

Shallice, T., Burgess, P., 1991. Higher-order cognitive impairments and frontal lobe lesions in man. Oxford University Press, NY.

Shultz, S., Dunbar, R.I.M., 2010. Species Differences in Executive Function Correlate With Hippocampus Volume and Neocortex Ratio Across Nonhuman Primates. *Journal of Comparative Psychology* 124, 252-260.

Sotomayor-Zarate, R., Tiszavari, M., Cruz, G., Lara, H.E., 2011. Neonatal exposure to single doses of estradiol or testosterone programs ovarian follicular development-modified hypothalamic neurotransmitters and causes polycystic ovary during adulthood in the rat. *Fertil Steril* 96, 1490-1496.

Stanczyk, F.Z., 2006. Measurement of Androgens in Women. *Semin Reprod Med* 24, 078-085.

Steckelbroeck, S., Heidrich, D., Stoffel-Wagner, B., Hans, V., Schramm, J., Bidlingmaier, F., Klingmuller, D., 1999. Characterization of Aromatase Cytochrome P450 Activity in the Human Temporal Lobe. *J Clin Endocr Metab* 84, 2795-2801.

Stoffel-Wagner, B., Watzka, M., Schramm, J., Bidlingmaier, F., Klingmuller, D., 1999. Expression of CYP19 (aromatase) mRNA in different areas of the human brain. *J Steroid Biochem* 70, 237-241.

Welsh, M.C., Pennington, B.F., Groisser, D.B., 1991. A Normative Developmental-Study of Executive Function - a Window on Prefrontal Function in Children. *Developmental Neuropsychology* 7, 131-149.

Witte, A.V., Savli, M., Holik, A., Kasper, S., Lanzenberger, R., 2010. Regional sex differences in grey matter volume are associated with sex hormones in the young adult human brain. *NeuroImage* 49, 1205-1212.

Woods, S.P., Delis, D.C., Scott, J.C., Kramer, J.H., Holdnack, J.A., 2006. The California Verbal Learning Test--second edition: test-retest reliability, practice effects, and reliable change indices for the standard and alternate forms. *Arch Clin Neuropsychol* 21, 413-420.

Worsley, K.J., Evans, A.C., Marrett, S., Neelin, P., 1992. A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab* 12, 900-918.

Worthman, C.M., Stallings, J.F., Hofman, L.F., 1990. Sensitive salivary estradiol assay for monitoring ovarian function. *Clin Chem* 36, 1769-1773.

Yoon, T., Okada, J., Jung, M.W., Kim, J.J., 2008. Prefrontal cortex and hippocampus subserve different components of working memory in rats. *Learning & Memory* 15, 97-105.

Zuloaga, D.G., Puts, D.A., Jordan, C.L., Breedlove, S.M., 2008. The role of androgen receptors in the masculinization of brain and behavior: what we've learned from the testicular feminization mutation. *Horm Behav* 53, 613-626.

TABLES AND FIGURES

TABLE 1: Sample Characteristics

		Testosterone-Related Analyses	BRIEF_executive	CVLT-C (Verbal Memory)
Subjects & scans	1	n = 121 scans	n = 119 scans	n = 97 scans
	2	n = 110 scans	n = 91 scans	n = 73 scans
	3	n = 93 scans Total = 324	n = 35 scans Total = 245	n = 30 scans Total = 200
Number of scans	1 scan	129	132	113
	2 scans	66	52	39
	3 scans	21	3	3
Testosterone (pg/mL)	1	n = 121, mean = 99.20, SD = 74.14	n = 119, mean = 99.12, SD = 74.76	n = 97, mean = 84.89, SD = 61.58
	2	n = 110, mean = 74.00, SD = 45.83	n = 91, mean = 71.84, SD = 43.97	n = 73, mean = 61.81, SD = 31.97
	3	n = 93, mean = 105.23, SD = 101.34	n = 35, mean = 88.33, SD = 99.01	n = 30, mean = 86.24, SD = 106.14
Estradiol	1	n = 121, mean = 10.87, SD = 5.63	n = 119, mean = 10.86, SD = 5.67	n = 97, mean = 10.6, SD = 5.507
	2	n = 110, mean = 12.11, SD = 6.19	n = 91, mean = 12.42, SD = 6.16	n = 73, mean = 12.12, SD = 6.20
	3	n = 93, mean = 16.82, SD = 10.01	n = 35, mean = 19.45, SD = 10.93	n = 30, mean = 18.89, SD = 11.39
DHEA	1	n = 121, mean = 103.07, SD = 100.03	n = 119, mean = 103.90, SD = 100.66	n = 97, mean = 90.50, SD = 101.94
	2	n = 110, mean = 188.84, SD = 171.85	n = 91, mean = 173.19, SD = 157.78	n = 73, mean = 141.07, SD = 146.96
	3	n = 93, mean = 197.28, SD = 164.48	n = 35, mean = 147.44, SD = 138.07	n = 30, mean = 132.19, SD = 136.79
Season of sampling	1	Spring = 0 Summer = 11 Fall = 110 Winter = 0 Total = 121	Spring = 0 Summer = 11 Fall = 108 Winter = 0 Total = 119	Spring = 0 Summer = 8 Fall = 89 Winter = 0 Total = 97
	2	Spring = 0 Summer = 0 Fall = 108 Winter = 2 Total = 110	Spring = 0 Summer = 0 Fall = 89 Winter = 2 Total = 91	Spring = 0 Summer = 0 Fall = 73 Winter = 0 Total = 73
	3	Spring = 0 Summer = 0 Fall = 3 Winter = 90 Total = 93	Spring = 0 Summer = 0 Fall = 0 Winter = 35 Total = 35	Spring = 0 Summer = 0 Fall = 0 Winter = 30 Total = 30
Collection time (min after midnight)	1	n = 121, mean = 682.58, SD = 139.42	n = 119, mean = 684.90, SD = 139.43	n = 97, mean = 676.86, SD = 134.61
	2	n = 110, mean = 710.63, SD = 115.73	n = 91, mean = 713.27, SD = 109.35	n = 73, mean = 702.00, SD = 115.47

	3	n = 93, mean = 702.38, SD = 123.07	n = 35, mean = 699.03, SD = 132.12	n = 30, mean = 687.63, SD = 130.85
Age (years)	1	n = 121, mean = 12.71, SD = 3.22	n = 119, mean = 12.63, SD = 3.19	n = 97, mean = 11.65, SD = 2.66
	2	n = 110, mean = 13.48, SD = 3.66	n = 91, mean = 12.80, SD = 3.31	n = 73, mean = 11.58, SD = 2.49
	3	n = 93, mean = 14.50, SD = 3.77	n = 35, mean = 12.76, SD = 2.64	n = 30, mean = 12.06, SD = 2.14
Gender F = female M = male	1	F = 69, M = 52	F = 68, M = 51	F = 54, M = 43
	2	F = 71, M = 39	F = 57, M = 34	F = 47, M = 26
	3	F = 50, M = 43	F = 21, M = 14	F = 18, M = 12
Pubertal stage	1	n = 121, mean = 2.53 SD = 1.47	n = 119, mean = 2.51 SD = 1.478	n = 97, mean = 2.14, SD = 1.315
	2	n = 110, mean = 2.52, SD = 1.48	n = 91, mean = 2.46, SD = 1.463	n = 73, mean = 2.51, SD = 1.45
	3	n = 93, mean = 3.05, SD = 1.492	n = 35, mean = 3.03, SD = 1.445	n = 30, mean = 3.03, SD = 1.402
Handedness	1	L = 8, R = 113 Total = 121	L = 8, R = 111 Total = 119	L = 8, R = 89 Total = 97
	2	L = 9, R = 101 Total = 110	L = 8, R = 83 Total = 91	L = 7, R = 66 Total = 73
	3	L = 8, R = 85 Total = 93	L = 3, R = 32 Total = 35	L = 3, R = 27 Total = 30
Total brain volume (cm ³)	1	n = 121, mean = 1270.55 SD = 123.26	n = 119, mean = 1270.89, SD = 123.94	n = 97, mean = 1283.92, SD = 121.13
	2	n = 110, mean = 1268.22, SD = 126.26	n = 91, mean = 1275.79, SD = 126.51	n = 73, mean = 1284.06, SD = 124.84
	3	n = 93, mean = 1285.88, SD = 140.12	n = 35, mean = 1286.54, SD = 139.63	n = 30, mean = 1283.56, SD = 133.95
Left hippocampus (mm ³)	1	n = 121, mean = 2954.43 SD = 322.78	n = 119, mean = 2955.73, SD = 324.86	n = 97, mean = 2955.58, SD = 322.50
	2	n = 110, mean = 2971.405, SD = 296. 595	n = 91, mean = 2964.77, SD = 304.25	n = 73, mean = 2944.92, SD = 311.20
	3	n = 93, mean = 3071.97, SD = 356. 90	n = 35, mean = 2992.22, SD = 354.68	n = 30, mean = 3012.18, SD = 364.67
Right hippocampus (mm ³)	1	n = 121, mean = 3045.01 SD = 349.12	n = 119, mean = 3043.366, SD = 351.70	n = 97, mean = 3047.26, SD = 354.06
	2	n = 110, mean = 3064.97, SD = 330. 436	n = 91, mean = 3064. 55, SD = 343.96	n = 73, mean = 3030.19, SD = 345.69

Table 2: Description of statistical models

Methods section	Statistical model
2.5.2 Testosterone & Cortico-Hippocampal Networks	<p>(1) Whole-brain CTh = 1 + <u>Testosterone*Hippocampus</u> + Testosterone + Hippocampus + Collection Time + Age + Sex + Scanner + Handedness + Total Brain Volume + random (Subj) + I</p> <p>(2) Whole-brain CTh = 1 + <u>Testosterone*Hippocampus*Sex</u> + Testosterone*Hippocampus + Hippocampus*Sex + Testosterone*Sex + Testosterone + Hippocampus + Sex + Collection Time + Age + Scanner + Handedness + Total Brain Volume + random (Subj) + I</p> <p>(3) Whole-brain CTh = 1 + <u>Testosterone*Hippocampus*Age</u> + Testosterone*Hippocampus + Hippocampus*Age + Testosterone*Age + Testosterone + Hippocampus + Age + Collection Time + Sex + Scanner + Handedness + Total Brain Volume + random (Subj) + I</p> <p>(4) Whole-brain CTh = 1 + <u>Testosterone*Hippocampus*Age*Sex</u> + Testosterone*Hippocampus*Age + Testosterone*Hippocampus*Sex + Hippocampus*Age*Sex + Testosterone*Age*Sex + Testosterone*Hippocampus + Testosterone*Sex + Hippocampus*Sex + Testosterone*Age + Hippocampus*Age + Age*Sex + Testosterone + Hippocampus + Age + Sex + Collection Time + Scanner + Handedness + Total Brain Volume + random (Subj) + I</p> <p>(5) Note that in order to limit the number of control variables per model: models (1), (2), (3) and (4) were retested while adding Estradiol, DHEA, Pubertal Stage or Season of Sampling as additional covariates (one at a time)</p>
2.5.3 Cortico-Hippocampal Networks & Cognitive Tests	<p>(1) Following significant interaction Testosterone*Hippocampus*Sex in 2.5.2: Cognitive Scores = 1 + <u>CTh*Hippocampus*Sex</u> + CTh*Hippocampus + Hippocampus*Sex + CTh*Sex + CTh + Hippocampus + Sex + Age + Scanner + Handedness + Total Brain Volume + random (Subj) + I</p>
2.5.4 Mediation	<p>(1) Coefficients and p-values extracted from section 2.5.2</p> <p>(2) Coefficients and p-values extracted from section 2.5.3</p> <p>(3) Cognitive scores = 1 + <u>Testosterone</u> + Age + Sex + Collection Time + random (Subj) + I</p> <p>(4) Coefficients and p-values from (1), (2) and (3) were extracted from existing analyses, and entered in the Sobel-Goodman test calculator to formally test mediation effects of scans per subject (see ‘Methods’)</p>

‘Subj’ refers to a specific subject; and ‘I’ to the identity matrix of the mixed effects model

‘CTh’ as referred to in section 2.5.3 refers to average cortical thickness of the dorsal anterior cingulate region found to be significant in section 2.5.2

Figure 1: Testosterone, Prefrontal-Hippocampal Structural Covariance & Executive Function

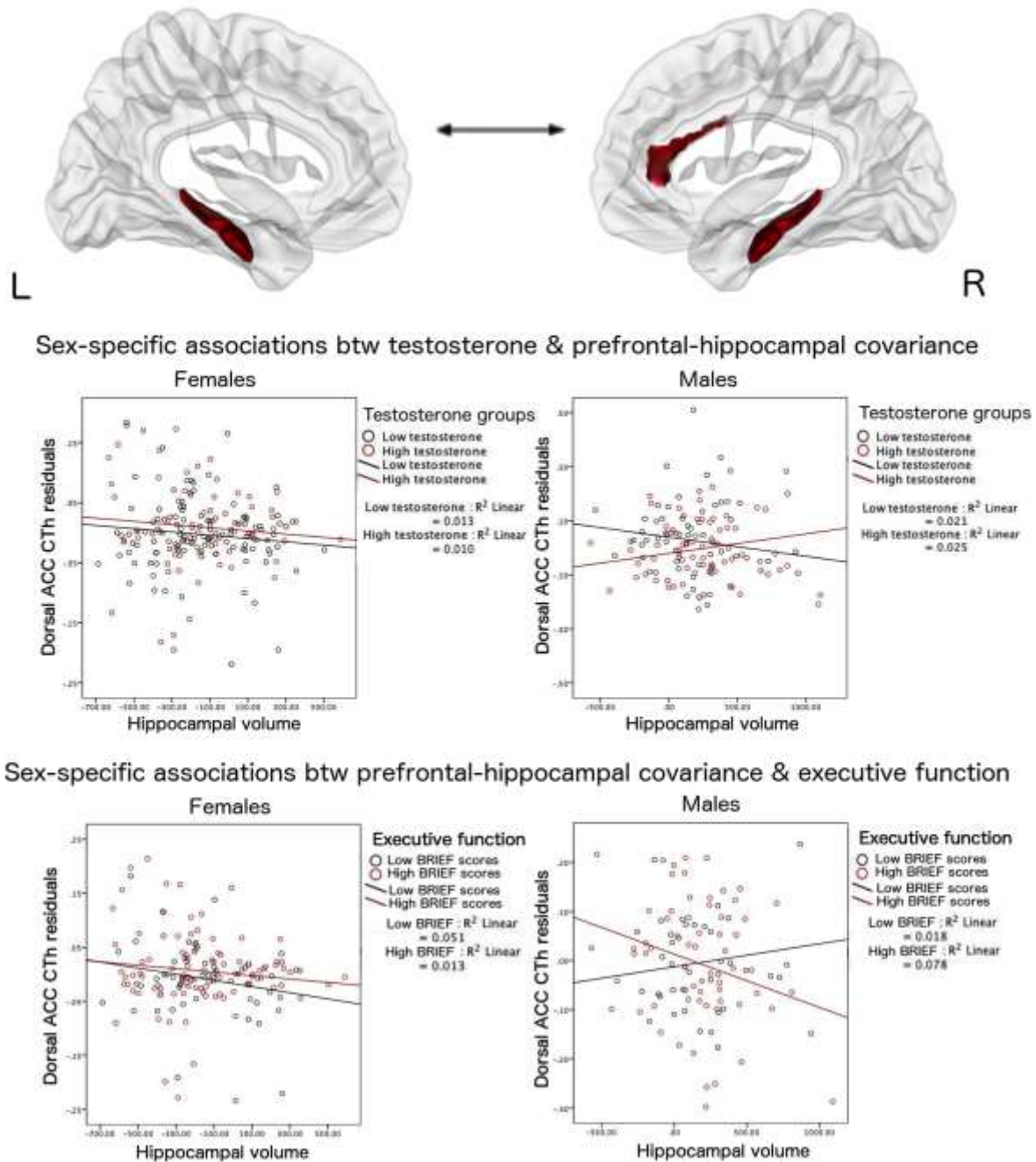


FIGURE 1 LEGEND:

This figure shows sex-specific associations between testosterone, prefrontal-hippocampal structural covariance and executive function. In boys (but not girls), prefrontal-hippocampal covariance varied according to testosterone levels. In addition, in boys (but not girls), prefrontal-hippocampal structural covariance predicted scores on different components of executive function (BRIEF Monitor and Shift subscales).

Top two brain figures: Left hippocampus (on the left) and right hippocampus (on the right) are shown, along with the significant area of the right dorsal anterior cingulate cortex found to be significant in whole-brain analyses, corrected for multiple comparisons using random field theory ($p < 0.05$).

Middle two graphs: In boys, higher testosterone levels were associated with a positive covariance between mean hippocampal volume and cortical thickness of the right dorsal anterior cingulate gyrus, while lower testosterone levels were associated with a negative covariance between these regions.

Bottom two graphs: In boys, positive prefrontal-hippocampal covariance (associated with higher testosterone levels) was associated with lower BRIEF scores (executive function), while negative prefrontal-hippocampal covariance was associated with higher BRIEF scores (BRIEF Monitor shown as an example, similar results with BRIEF Shift).

Testosterone and cognitive scores were split into high and low groups for the purposes of visualization, based on the values at which prefrontal-hippocampal covariance shifted from positive to negative, even though both variables were included as continuous variables in all analyses. Further, please note that although the individual slopes for each dichotomous group are included here for informative purposes, they do not reflect the significance of the interaction analyses, which used continuous measures of testosterone and cognitive scores.

ACC: anterior cingulate gyrus
CTh: cortical thickness

Figure 2: Prefrontal-Hippocampal Structural Covariance Mediates the Relationship between Testosterone and Executive Function

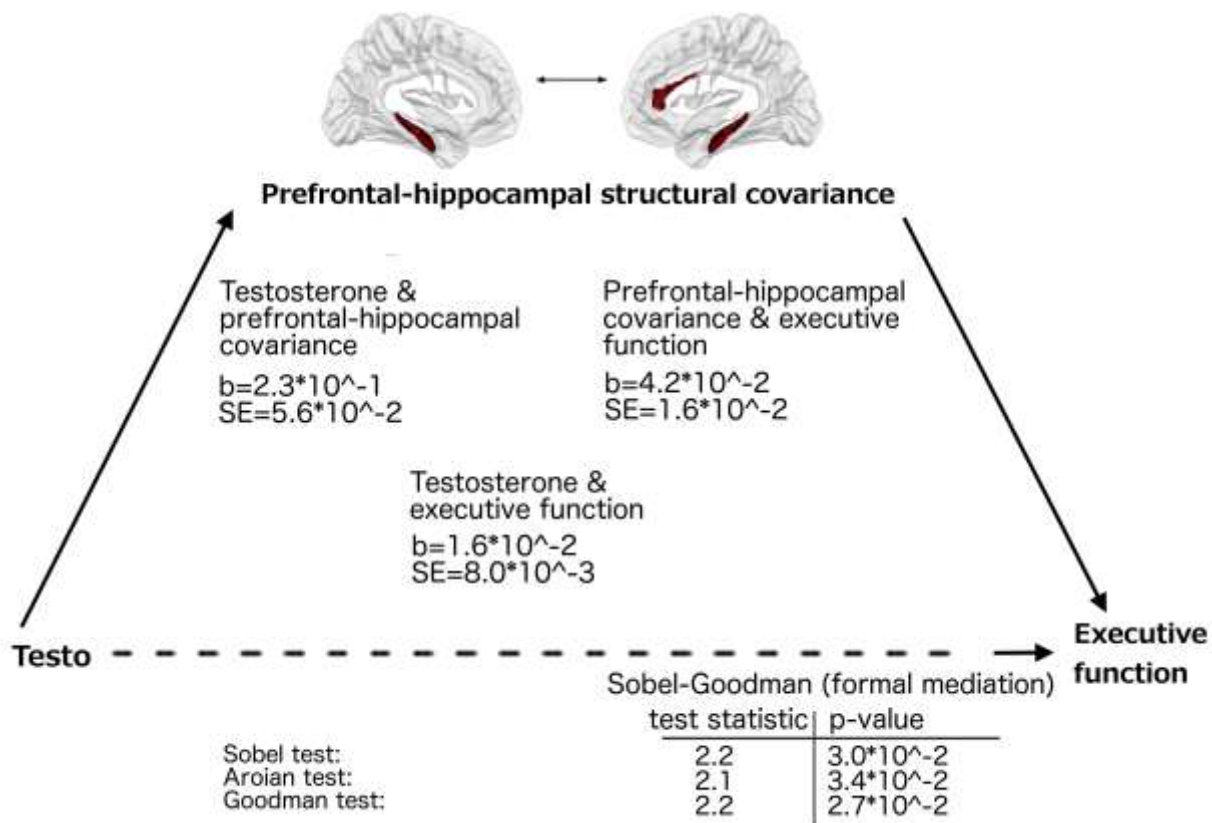


FIGURE 2 LEGEND:

This figure displays the relationship between testosterone levels, prefrontal-hippocampal covariance (for mean hippocampus), and executive function –BRIEF Monitor subscale (see Methods section 2.5, and Results section 3.4).

The beta coefficients and standard errors are displayed in the center of the figure:

- (1) on the left, for the relationships between testosterone and prefrontal-hippocampal covariance (as moderated by sex)
- (2) on the right, for the relationship between prefrontal-hippocampal covariance and executive function (BRIEF Monitor, as moderated by sex);
- (3) at the lower center, for the relationship between testosterone and executive function (BRIEF Monitor, as moderated by sex)

Test statistics and p-values for the formal mediation tests (Sobel, Aroian, Goodman) are displayed at the bottom of the figure. These tests are all different versions of the Sobel-Goodman mediation test (with or without including the third term to estimate the variance of the mediated effect), and they tend to converge closely with sample sizes greater than 50, as seen here.