Psychoneuroendocrinology 68 (2016) 194–201

Contents lists available at ScienceDirect

Psychoneuroendocrinology

journal homepage: www.elsevier.com/locate/psyneuen

# Testosterone reduces functional connectivity during the 'Reading the Mind in the Eyes' Test

Peter A. Bos[a,b,*], Dennis Hofman[a], Erno J. Hermans[c,d], Estrella R. Montoya[a], Simon Baron-Cohen[e], Jack van Honk[a,b,f]

[a] Department of Experimental Psychology, Utrecht University, Utrecht, The Netherlands
[b] Department of Psychiatry, University of Cape Town, Cape Town, South Africa
[c] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands
[d] Department for Cognitive Neuroscience, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands
[e] Autism Research Centre, Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom
[f] Department of Psychiatry and Institute of Infectious Diseases and Molecular Medicine (IDM), University of Cape Town, Cape Town, South Africa

ARTICLE INFO

*Article history:*
Received 13 August 2015
Received in revised form 7 March 2016
Accepted 8 March 2016

*Keywords:*
Functional MRI
Hormones
Autism spectrum conditions
Social behavior
Mindreading

ABSTRACT

Women on average outperform men in cognitive-empathic abilities, such as the capacity to infer motives from the bodily cues of others, which is vital for effective social interaction. The steroid hormone testosterone is thought to play a role in this sexual dimorphism. Strikingly, a previous study shows that a single administration of testosterone in women impairs performance on the 'Reading the Mind in Eyes' Test (RMET), a task in which emotions have to be inferred from the eye-region of a face. This effect was mediated by the 2D:4D ratio, the ratio between the length of the index and ring finger, a proxy for fetal testosterone. Research in typical individuals, in individuals with autism spectrum conditions (ASC), and in individuals with brain lesions has established that performance on the RMET depends on the left inferior frontal gyrus (IFG). Using functional magnetic resonance imaging (fMRI), we found that a single administration of testosterone in 16 young women significantly altered connectivity of the left IFG with the anterior cingulate cortex (ACC) and the supplementary motor area (SMA) during RMET performance, independent of 2D:4D ratio. This IFG-ACC-SMA network underlies the integration and selection of sensory information, and for action preparation during cognitive empathic behavior. Our findings thus reveal a neural mechanism by which testosterone can impair emotion-recognition ability, and may link to the symptomatology of ASC, in which the same neural network is implicated.



## 1. Introduction

Human social interaction is characterized by the employment of cognitive empathy: our capacity to infer motives, intentions, thoughts and feelings from the bodily cues of others (Baron-Cohen, 1995; Frith and Frith, 1999). The ability to identify emotional expressions of others shows sexual dimorphism: on average, women outperform men, a difference for which accumulating evidence suggests an underlying role for the steroid hormone testosterone (Baron-Cohen, 2003; Baron-Cohen et al., 2005). Indeed, a single administration of testosterone has been demonstrated to reduce emotion recognition abilities in typical young women (van Honk et al., 2011a). After testosterone administration (compared to placebo) women show impaired performance on the 'Reading the Mind in the Eyes' Test (RMET: Baron-Cohen et al., 2001). On the RMET, both basic and complex mental states (emotions, motives, intentions, and thoughts) need to be inferred from pictures of the eye-region of the face alone. This effect of testosterone on the RMET varied with a proxy of fetal testosterone (fT), the 2D:4D ratio (Breedlove, 2010). A lower 2D:4D ratio, a proxy for higher levels of fT, predicted greater impairment on the RMET after testosterone administration (van Honk et al., 2011a).

Exposure to fT is also an important underlying factor in the etiology of autism spectrum conditions (ASC) (Baron-Cohen, 2003; Baron-Cohen et al., 2015a, 2005), a condition characterized by impaired ability in emotion recognition (Baron-Cohen, 1995), for which the RMET was developed as a sensitive measure (Baron-Cohen et al., 2001). In the current pharmacological functional Magnetic Resonance Imaging (fMRI) study, typical young women

* Corresponding author at: Department of Experimental Psychology, Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands.
*E-mail address:* p.a.bos@uu.nl (P.A. Bos).

http://dx.doi.org/10.1016/j.psyneuen.2016.03.006
0306-4530/

performed the RMET during neuroimaging after a single testosterone and placebo administration, in a double-blind design, to investigate the neural mechanisms by which testosterone brings about its down-regulating effect on cognitive empathic abilities.

Studies into the mechanisms supporting cognitive empathic abilities have identified several brain regions involved in the inference of mental states using the RMET. A study that included more than a hundred patients with brain lesion shows that lesions in the left inferior frontal gyrus (IFG) impair performance on the RMET (Dal Monte et al., 2014), suggesting that left IFG is crucial for the identification of subtle emotional expressions. Consistent with this, studies comparing neural responses during the RMET between cultures and throughout development have also revealed consistent activation of the IFG, together with the posterior part of the superior temporal sulcus (pSTS) (Adams et al., 2010; Baron-Cohen et al., 2006, 1999; Moor et al., 2012). Interestingly, in relation to the role of testosterone, sex differences have been observed in the function of the left IFG (Baron-Cohen et al., 2006), and in connectivity of the left IFG with the superior temporal cortices (Schmithorst and Holland, 2007). Also, the amount of exposure to fT and baseline testosterone levels predicts brain structure (Koolschijn et al., 2014; Lombardo et al., 2012; Witte et al., 2010), and structural connectivity of the IFG and pSTS (Peper et al., 2013; Rametti et al., 2012). Altered activation and connectivity of the same regions have also been implicated in the social deficits of ASC (Hadjikhani et al., 2007; Vissers et al., 2012).

Although based on these studies a relation between testosterone and activation or connectivity of the left IFG can be predicted, direct effects of testosterone administration on the IFG have not yet been reported. Studies have reported acute down-regulation of prefrontal-amygdala functional connectivity after testosterone administration (Bos et al., 2012a; van Wingen et al., 2010), both in the context of emotion processing paradigms. In these studies, altered connectivity was observed in the absence of changed activation in the prefrontal cortex, indicating that endocrine manipulations can affect neural activity and connectivity independently. Thus, testosterone can be regarded a modulator of the network involved in emotion recognition, provisionally by altering connectivity of the prefrontal cortex with other brain regions (van Honk et al., 2011b). Given these findings, we predict that testosterone will selectively affect connectivity of the prefrontal regions involved in cognitive empathic abilities (i.e. left IFG), underlying the down-regulation effect of testosterone during RMET performance (van Honk et al., 2011a).

## 2. Methods and materials

### 2.1. Participants

16 right-handed typical young women (mean age 20.8 years, SD 2.0) participated in the study, and were recruited at the university campus of Utrecht University. Only women using single-phase oral contraceptives were included, since contraceptive use suppresses cyclic fluctuations in levels of estrogens and progesterone (Fleischman et al., 2010). Also, no scans were performed during menstruation. Also, participants were scanned at the same time of day on two separate days with an interval of at least a week. Only women were included, since our earlier study showing behavioral effects of testosterone on the RMET was also only performed in women (van Honk et al., 2011a), and because ethical approval was restricted to females. (For further information see: substance administration). Participants had no history of psychiatric, neurological, or endocrine abnormalities. Participants did not smoke and used no medication other than oral contraceptives. When compared to normative data from the general population, the participants in the current study did not differentiate on self-reported empathy (using the Empathy Quotient or EQ; see Procedure). Our sample had a mean EQ of 44.1 (SD = 10.8) compared to an average of 49.9 (SD = 12.0) in the normative data (Baron-Cohen et al., 2015b). The experimental protocol was approved by the Ethics Committee of the University Medical Centre Utrecht and in accordance with the latest declaration of Helsinki. Participants gave written informed consent prior to participation and received payment afterwards.

### 2.2. Testosterone administration

The testosterone sample consisted of 0.5 mg of testosterone, 5 mg of the carrier cyclodextrin, 5 mg of ethanol, and 5 ml of water. The placebo sample was identical to the drug sample, but without containing testosterone. Both testosterone and placebo were administered sublingually under supervision of the experimenter. Extensive prior research established that 0.5 mg of testosterone in young women, without exception, results in an approximately 10-fold increase in blood levels of testosterone 15 min after administration, and a return to baseline in 90 min (Tuiten et al., 2000). This 10-fold increase mostly reflects testosterone that is bound to sex hormone binding globulin (SHBG) and albumin, whereas the increase of the free fraction, which passes the blood brain barrier and exerts its effects in the brain, is increased to a moderate extent. A study into dose-dependent effects of sublingual testosterone administration indeed shows that the increase of the free fraction by testosterone depends on the SHBG level in women (van Rooij et al., 2011). Behavioral and physiological effects of the administration, as demonstrated by enhanced vaginal pulse amplitude, peak four hours after intake (Tuiten et al., 2000). Numerous studies using various cognitive, emotional, and neural measures have reported effects after this four hour delay (for a review, see: Bos et al., 2012b). Therefore, the same four-hour delay was used in the present study.

### 2.3. Procedure

The experimental setup of this study follows a randomized, counterbalanced, cross-over, placebo-controlled, testosterone administration paradigm. Upon arrival, the administration and fMRI procedure were explained to the participants who were given the opportunity to ask questions. Afterwards participants gave informed consent and were screened on a questionnaire for alcohol and drug use. They also filled out the Empathy Quotient (EQ) and Systemizing Quotient-Revised (SQ-R) (Baron-Cohen, 2009), as these questionnaires have shown clear on average sex differences (Wheelwright et al., 2006) and are correlated with exposure to fT (Auyeung et al., 2006). Four hours prior to fMRI data acquisition, participants received testosterone or placebo administration, due to the delay in physiological effects as explained above. After administration participants were instructed to refrain from physically and psychologically intensive tasks. Before being asked to take their place in the MRI scanner, participants filled out a shortened version of the profile-of-mood-states (POMS) questionnaire to obtain an index of their mood (Shacham, 1983). After this, participants were screened using a MRI-checklist and a metal detector, and were instructed to position themselves on the scanner bed as comfortably as possible and to try to relax. Head movement was minimized by foam pads that were placed between the RF-coil and the participant's head. Further instructions during the scan session were given by intercom. During the scan session, the participants performed two additional experimental tasks (published elsewhere: Bos et al., 2010, 2012a). The current task was always performed first. After the second session, participants were

debriefed, given payment, and asked to guess the day of administration to control for blindness regarding to the administration.

### 2.4. RMET

The task currently used is an adaptation of the original 'Reading the Mind in the Eyes' Test (Baron-Cohen et al., 2001), which was developed to be sensitive to individual differences and to impairment in adults with ASC. It consists of 36 pictures of the eye-region of faces that depict a specific emotion. In the original version of the task each picture has four possible answers among which the participant has to choose the correct one. To simplify the display, in order to avoid excessive eye movement during reading, we first showed the participants an emotion-word (e.g. shy, hostile, playful) for one second, after which the picture of the eyes was presented for three seconds (Fig. 1A). Participants were instructed to indicate whether the word matched the picture after the offset of the picture by pressing either left (match) or right (does not match) using a button box in their right hand. Reaction times were calculated for the matched and non-matched trails of the emotion and control condition, and was collapsed over correct and incorrect responses. The matched and non-matched answers were taken from the original RMET. As a control condition, the same pictures were shown while preceded by a matched or non-matched non-emotional word describing a feature of the pictures (e.g. woman, curly hair, heavy eye-brows, ear visible). This resulted in 144 trials (36 pictures, where the preceding word was emotion matched, emotion non-matched, control condition matched, control condition non-matched). Since the participants performed the task twice (both in the placebo and testosterone condition), the 144 trials were divided over two versions of the task to avoid learning effects. Every stimulus was displayed twice in the both versions of the task, and all answers categories were equally divided over the two versions. These two versions of the task were counterbalanced to the drug order and stimuli were randomly presented within both sessions.

All trials were presented in the middle of the screen, and were interleaved with a fixation cross with a random duration (averaged at 5.5 s). In addition, 15 null events (fixation cross a 9.5 s) were added in both task versions to secure enough rest throughout the experiment. Our design was developed to avoid excessive eye-movement during the task, which might have been a confound in previous studies (Bartlett et al., 2011), where the stimuli were presented simultaneously with several answer categories (Adams et al., 2010; Baron-Cohen et al., 2006; Moor et al., 2012). In addition, differences in difficulty between the conditions is better controlled for, as in these previous experiments the control condition only included sex and age judgments. Although our design was optimized for imaging purposes, the changes in the task have made it less sensitive for measuring behavioral effects. Since participants were no longer able to see both matched and non-matched condition at the same time (as in the original version), the difficult items in our version of the task are no longer discriminative for proficiency in emotion-detection. In difficult items, both emotional categories might apply, and deciding that a non-matched answer is correct, might reflect a more liberal response bias instead of less proficiency on the RMET.

### 2.5. 2D:4D ratio/testosterone saliva measurement

The ratio of the length of the right hand's second digit divided by the length of the fourth digit (2D:4D) is a marker for prenatal androgens in humans (Breedlove, 2010), and has in our previous study shown to predict the effect of testosterone administration on RMET performance (van Honk et al., 2011a). 2D:4D ratio was measured by two independent raters (inter-rater correlation: $r = 0.91$; $p < 0.001$) from an image scan of the right-hand of the subjects, which is a valid method to measure finger lengths. Lengths of the second and fourth digits were measured from the ventral proximal crease of the digit to the fingertip using an Adobe® Photoshop tool (Breedlove, 2010).

Baseline testosterone levels were obtained since these have been shown to correlate with neural responses to faces (Bos et al., 2012a; e.g. Derntl et al., 2009). Endogenous testosterone levels were obtained using saliva sampling according to Granger et al. (2004), which has been successfully applied in several previous studies (Bos et al., 2012a; e.g. Eisenegger et al., 2010). Testosterone in saliva was measured after diethylether extraction using a competitive radio-immunoassay employing a polyclonal antitestosterone-antibody (Dr. Pratt AZG 3290). [1,2,6,7-3H]-Testosterone (TRK402, Amersham Nederland B.V.) was used as a tracer following chromatografic verification of its purity. The lower limit of detection was 10 pmol/l and inter-assay variation was 16.1; 11.5; and 5.1% at 21; 100 and 230 pmol/l respectively ($n = 4,5,5$).

### 2.6. Scanning parameters

Scanning was performed on a 3 T Philips Achieva MRI scanner (Philips Medical Systems, Best, The Netherlands). Blood oxygen level dependent (BOLD-) response was measured with functional T2*-weighted sagittal whole- brain images obtained throughout the task. A 3D PRESTO sequence (Neggers et al., 2008) was used with the following parameters: 23 ms echo time, 16 ms repetition time; $224 \times 224 \times 136$ mm field of view; flip angle of 9°. 1020 scans with a volume acquisition time of 0.813 s were obtained, each comprised of 39 sagittal slices; voxelsize 3.5 mm isotropic.Subsequently a high resolution T1-weighted anatomical scan with the following parameters was acquired for co-registration and normalization purposes: 4.7 ms echo time, 9.5 ms repetition time, $240 \times 221 \times 160$ mm field of view, 266 sagittal slices, flip angle of 8.0°, voxelsize 0.6 mm isotropic.

### 2.7. fMRI data processing and analyses

Preprocessing and subsequent analyses were performed with SPM8 (http://www.fil.ion.ucl.ac.uk/spm). Functional scans of both sessions were motion corrected to the first dynamic scan. Brain extraction was performed on the individual anatomical scans by using unified segmentation, and applying the gray and white matter maps as a mask for the anatomical scans. Extracted brains were used in further preprocessing steps leading to better fit during coregistration. Subsequently, all volumes were normalized to MNI-space using the MNI152 template and were resliced at 2 mm isotropic voxelsize. Smoothing with an 8 mm full width at half maximum gaussian kernel was applied to the normalized functional volumes.

A general linear model (GLM) was applied to both sessions to investigate the effect of the emotion versus the control condition, and its interaction with testosterone administration. Within both sessions, neural responses to the onset of the pictures were modelled using a 2 s boxcar function convolved with a hemodynamic response function (hrf) as implemented in the SPM8 software. Additional regressors of no interest which are entered into the analyses to reduce unexplained variance in the data include the realignment parameters, and a discrete cosine transform high pass filter with a cut-off of 128 s. None of the participants had to be removed from the data due to excessive head movement (more than 2 mm).

The contrast maps of emotion stimuli versus baseline and control stimuli versus baseline in both sessions were entered in a two-factorial ANOVA, with drug (testosterone versus placebo) and emotion condition (emotion versus control stimuli) as within subject factors. Matched vs. non-matched conditions were not added

to the model, since during both conditions neural activation represents a cognitive process during which a facial expression is compared to a word, and by dividing the trials in two categories, part of the variance is explained of our effect of interest. Order was entered as a between subjects factor, and was removed from the analyses if it showed not significant within the regions of interest. All calculated linear contrasts were subjected to a cluster-forming threshold of $p < 0.001$. To control for multiple comparisons this voxelwise cluster-forming threshold was combined with a cluster threshold, and only clusters that survive a $p < 0.05$ Family Wise Error (FWE) correction are reported. In addition to the whole brain analysis, small volume corrections ($p < 0.05$ FWE cluster-level) were applied for the predefined regions of the interest: the STS, and the IFG (orbital and triangular part), as based on the automated anatomical labelling (AAL) template (Tzourio-Mazoyer et al., 2002). The superior temporal gyrus (STG) was used as the anatomical template since the STS is not incorporated in the AAL-template as a separate mask.

Finally, we performed a psychophysiological interaction (PPI) analysis to investigate the possible effect of testosterone on functional brain connectivity during mindreading. As a seed region, we took the peak value with a 6 mm sphere radius of the contrast emotion versus control stimuli. This sphere was masked by an anatomical template of the IFG based on the AAL template, since a small part of the sphere was located outside of the IFG as defined in the AAL template (Tzourio-Mazoyer et al., 2002). The time course of activation in the seed region, was extracted from the data, and an interaction term for this time course was calculated for the contrast of emotion versus the control condition. Both time course and PPI term were entered in the regression model. Subsequently, PPI contrast maps of the testosterone and placebo sessions were subjected to a paired sample *T*-test. Additionally, to further specify which condition brought forth the effect, one-sampled *T*-tests were performed on the emotion versus control contrast maps in both drug sessions. The same corrections and thresholds were applied as described in the analysis above.

## 3. Results

### 3.1. Behavioral data

First, the data was checked for outliers based on a criterion of 3 times the SD above or below the mean. No outliers were detected. The distribution of the data was further checked using the Shapiro-Wilk normality test. With the exception of the unmatched control condition after testosterone ($W = 0.834$, $p = 0.008$), and the matched control condition after placebo ($W = 0.825$, $p = 0.006$), none of the number of errors made by the participants deviated from normality (all $W > 0.896$, $p > 0.05$). The reaction times all violated the assumption of normality (all $W < 0.849$, $p < 0.05$). Since normality was violated, we adopted a more conservative significance threshold of $p = 0.01$ in our further analyses (log-transformations with a conventional threshold of $p = 0.05$ yielded similar results). Mean and SD of the behavioral data in the different conditions are provided in Table 1. Please note that subjects were instructed to wait for the offset of the stimulus, which lasted 3 s, before responding. Reaction times are given from the onset of the stimulus, since some participants did not follow up on instruction (reaction times below 3 s).

A 2(drug) × 2(emotion/control) × 2(matched/non-matched) ANOVA on the number of errors showed a main effect of emotion condition ($F(1,15) = 59.9$, $p < 0.001$), but no effect of drug ($F(1,15) = 0.18$, $p > 0.01$), or interactions of drug with emotion condition ($F(1,15) = 0.002$, $p > 0.01$) or matched/non-matched condition ($F(1,15) = 1.12$, $p > 0.01$). Post-hoc pairwise comparisons showed that participants were more accurate in the control condition compared to the emotion condition (mean number of errors = 2.27 versus 4.53, $p < 0.001$).

A 2(drug) × 2(emotion/control) × 2(matched/non-matched) ANOVA on the reaction times showed a similar main effect of emotion condition ($F(1,15) = 60.37$, $p < 0.001$), but no effect of drug ($F(1,15) = 0.10$, $p > 0.01$), or interactions of drug with emotion condition ($F(1,15) = 0.21$, $p > 0.01$) or matched/non-matched condition ($F(1,15) = 1.28$, $p > 0.01$). Post-hoc pairwise comparisons showed that participants were faster in the control condition compared to the emotion condition (reaction time = 2.04 versus 2.20, $p < 0.001$). Further covariation analyses showed that there were no main effects, or interactions with drug administration or emotion category of the 2D:4D ratio, baseline testosterone, or the Systemizing Quotient/Empathy Quotient (SQ-R/EQ) (Baron-Cohen, 2009; Wheelwright et al., 2006) on the data (all $p$'s > 0.01).

Based on our previous behavioral study (van Honk et al., 2011a) we ran exploratory correlational analyses to investigate a possible effect of the 2D:4D ratio on the behavioral data. None of the correlations survived the lowered statistical threshold of ($p < 0.01$). Only a marginally significant negative correlation between the 2D:4D ratio and the SQ-R was observed ($\rho = -0.59$; $p = 0.016$, uncorrected for multiple comparisons), consistent with the idea that high fT results in stronger systemizing (Auyeung et al., 2006; Baron-Cohen, 2009). Furthermore, testosterone did not affect participants' mood, as there was no effect on any of the profile of mood states (POMS) scales (all $p$'s > 0.05). Also, participants' guesses on which day they had testosterone did not deviate from chance regarding the true day of testosterone administration (binominal = 0.18 NS, two-tailed).

### 3.2. Imaging data

To investigate neural activation during the RMET, we first contrasted the emotion condition to the control condition, yielding a large cluster of activation in the left IFG (Fig. 1B, Table 2 for inferential statistics), showing that our main area of interest was indeed activated. Our other region of interest, the pSTS, was not selectively activated during the cognitive empathy condition. However, when we looked at the activation of both conditions versus rest, it was significantly activated (Table 2), indicating that the control condition in our adapted RMET also elicited activation in this region. For the opposite task-contrast (more activation of the control compared to the emotion task) several significantly activated clusters were observed, most prominent in the middle frontal and parietal regions, as well as in the inferior temporal cortex. These regions comprise the neural network of visual attention and eye movement (Desimone and Duncan, 1995; Lynch and Tian, 2006),

**Table 1**
Behavioral data of the different condition (18 trials); mean and SD.

| condition | number of errors | | reaction time | |
|---|---|---|---|---|
| | placebo | testosterone | placebo | testosterone |
| emotion matched | 3.75 (1.73) | 4.75 (2.46) | 2.09 (0.84) | 2.21 (1.00) |
| emotion non-matched | 5.13 (2.83) | 4.50 (1.93) | 2.25 (0.86) | 2.26 (0.96) |
| control matched | 2.44 (2.34) | 2.56 (1.82) | 1.97 (0.90) | 2.06 (1.03) |
| control non-matched | 1.88 (1.15) | 2.19 (1.64) | 2.02 (0.87) | 2.10 (1.00) |

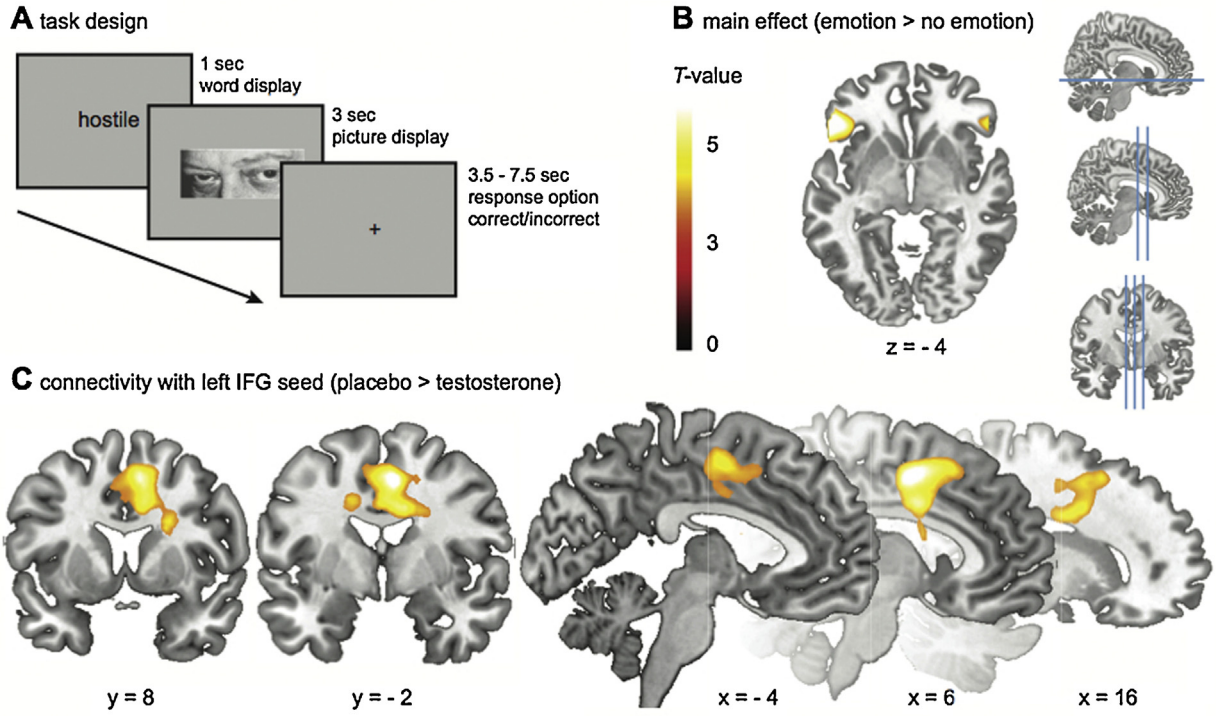


**Fig. 1.** (A) Display of one trial setup. (B) An axial slice of the T-maps for the contrast of emotion versus the control condition in the placebo and testosterone condition overlaid onto a T1-weighted canonical image. (C) Coronal and sagittal slices of the T-maps for the contrast placebo > testosterone, for the connectivity analysis based on the contrast between the emotion condition and the control condition After testosterone administration, there is a significant reduction in connectivity between the seed region (left IFG) and the bilateral SMA and ACC compared to placebo. Accompanying MNI-coordinates are presented below the slices, and for display purposes all T-maps are thresholded at $p < 0.001$ (see Table 2 for inferential statistics). For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.

**Table 2**
Peak T-values, *p*-values, cluster sizes, and MNI coordinates for significantly activated voxels.

| experimental effect | | peak voxel location | | | T-value | cluster size | p-values |
|---|---|---|---|---|---|---|---|
| region | | *x* | *y* | *z* | | voxels | |
| full factorial: main effect: stimuli > rest | | | | | | | |
| fusiform gyrus | L | −32 | −82 | −17 | 19.21 | 66.895 | <0.001* |
| lingual gyrus | R | 24 | −88 | −12 | 15.54 | s.c. | <0.001* |
| inferior frontal gyrus | R | 34 | 28 | −6 | 10.34 | s.c. | <0.001** |
| inferior frontal gyrus | L | −56 | 14 | 30 | 10.15 | s.c. | <0.001** |
| superior temporal gyrus | R | 58 | −44 | 12 | 5.04 | s.c. | 0.001** |
| main effect: rest > stimuli | | | | | | | |
| precuneus | L | −10 | −56 | 14 | 5.57 | 793 | 0.005* |
| angular gyrus | L | −50 | −70 | 30 | 6.38 | 462 | 0.037* |
| superior temporal gyrus | L | −54 | −32 | 10 | 4.41 | 127 | 0.032** |
| main effect: emotion > control | | | | | | | |
| inferior frontal gyrus | L | −54 | 32 | −4 | 7.17 | 714 | 0.008* |
| main effect: control > emotion | | | | | | | |
| middle frontal gyrus | R | 32 | 18 | 56 | 11.34 | 3746 | <0.001* |
| | L | −30 | 12 | 56 | 7.69 | 1981 | <0.001* |
| superior parietal cortex | L | −26 | −68 | 52 | 9.93 | 13.984 | <0.001* |
| inferior temporal cortex | L | −56 | −54 | −10 | 9.26 | 904 | 0.003* |
| | R | 60 | −44 | −18 | 7.51 | 708 | 0.008* |

R, right; L, left; s.c., same cluster as above.
* Whole brain FWE corrected at cluster-level.
** Small volume FWE corrected at cluster-level.

which corresponds to our control condition, in which participates had to indicate whether the faces in the pictures contained specific physical features (e.g. hair style, looking direction, detection of wrinkles).

Testing for effects of testosterone on these regions, a two-factorial ANOVA revealed no significant effects of drug, nor an interaction of drug with the emotion-condition. Thus, based on our whole brain analysis, testosterone does not seem to affect brain activity during cognitive empathy.

However, to test our hypothesis of whether testosterone would act on functional connectivity during mindreading, we ran a connectivity analysis for the contrast between the emotion condition and the control condition. Since the left IFG was specifically activated during the recognition of emotions, this region was chosen

**Table 3**
Peak T-values, *p*-values, cluster sizes, and MNI coordinates for significantly activated voxels.

| experimental effect | | peak voxel location | | | Cluster size | T-values | p-values |
|---|---|---|---|---|---|---|---|
| region | | *x* | *y* | *z* | voxels | | |
| PPI analysis emotion > control condition | | | | | | | |
| Paired sample T-test placebo > testosterone | | | | | | | |
| Supplementary motor area | R | 6 | −2 | 48 | 985 | 6.73 | <0.001* |
| Placebo: one sample T-test positive contrast | | | | | | | |
| Inferior parietal lobule | L | −28 | −50 | 46 | 439 | 5.44 | 0.016* |
| Supplementary motor area | R | 6 | −2 | 48 | 408 | 5.00 | 0.021* |
| Lingual gyrus | L | −16 | −92 | −18 | 474 | 4.77 | 0.012* |
| Testosterone: one sample T-test positive contrast | | | | | | | |
| Postcentral gyrus | L | −60 | −16 | 38 | 585 | 7.29 | 0.003* |
| Inferior occipital lobe | L | −38 | −84 | −6 | 317 | 5.47 | 0.041* |
| Testosterone: one sample T-test negative contrast | | | | | | | |
| White matter | L | −30 | −18 | 28 | 1369 | 6.63 | <0.001* |

R, right; L, left.
* Whole brain corrected (FWE) at cluster level.

as a seed region. Crucially, this analysis shows a significant cluster encompassing the bilateral supplementary motor cortex (SMA) and anterior cingulate cortex (ACC) for the contrast of testosterone compared to placebo (Fig. 1C). We subsequently ran one-sample T-tests for the emotion versus control contrast within the placebo and testosterone session (Table 3). These tests showed that the SMA/ACC cluster was activated in the placebo condition, but not in the testosterone condition. This shows that after testosterone administration, there is a significant reduction in connectivity between the left IFG and the bilateral SMA/ACC, for the contrast of the emotion condition versus the control condition.

In none of the reported analysis did order of the drug administration, digit ratio, or baseline testosterone levels have a significant effect on the data (all *p*'s > 0.05).

## 4. Discussion

We investigated the neural mechanism by which testosterone impairs the cognitive-empathic abilities during performance on the RMET (van Honk et al., 2011a). The data show that the left IFG is activated specifically during the emotion-recognition condition, that is, when participants had to infer the emotional state in the stimulus. Critically, compared to the placebo-condition, testosterone reduced connectivity of the left IFG with the bilateral ACC and SMA during emotion-recognition compared to the control condition. Although few studies have investigated the effect of testosterone on functional brain connectivity, two previous studies have reported reduced frontal cortex connectivity in the context of affective and trust behaviors (Bos et al., 2012a; van Wingen et al., 2010). However in these studies connectivity with the amygdala was reduced, whereas in the present study during the RMET, testosterone acted on functional connectivity of the IFG with the ACC and SMA. By altering this pattern of connectivity testosterone might have caused the earlier observed impairment in performance on the RMET (van Honk et al., 2011a). This should be confirmed in future studies, as in the current study no effects of testosterone on behavioral measures were found. This is most likely caused by our adapted task-design, which was optimized for neuroimaging, but had reduced sensitivity in finding behavioral effects (see Methods and Materials). Our adaptation of the RMET demonstrates altered functional connectivity during identification of emotions in others after testosterone, in the absence of an effect on behavioral performance on the task. Arguably, paradigms tailored to detecting effects on behavior are needed to reveal the effects of testosterone administration on a behavioral level. Indeed, a recent testosterone administration study employing the original version of the RMET replicated the effect reported in van Honk et al. (2011a), by showing impaired performance after testosterone administration (Olsson et al., 2016).

Based on neuroimaging data of both typical individuals and patients with lesions, the left IFG can be considered crucial for mindreading ability during the RMET (Adams et al., 2010; Baron-Cohen et al., 1999; Dal Monte et al., 2014; Moor et al., 2012), although its exact role in this behavior is as yet unclear. IFG importantly integrates information from different modalities (Hagoort, 2005; Willems et al., 2009), and selects between competing options (Zhang et al., 2004), thereby facilitating perception-action coupling (Liakakis et al., 2011). The orbital part of the left IFG, corresponding to the activation observed in the current study, is particularly implicated in empathic behavior (Liakakis et al., 2011) and social decision making (Bos et al., 2012a). The ACC, which together with the SMA showed decoupling with the IFG after testosterone, is tightly linked to subcortical regions (Etkin et al., 2011), and incorporates motivational and affective states in cognitive control processes (Pessoa, 2008; Shenhav et al., 2013), whereas the SMA is involved in the initiation of voluntary motor action (Desmurget and Sirigu, 2009). Thus, together these regions are involved in integrating of sensory information from different modalities with emotional states, and selecting and initiating appropriate motor responses based on this information.

Recent meta-analyses have shown the relevance of this network in both empathic behavior (Fan et al., 2011) and emotion regulation (Kohn et al., 2014). Specifically, the strength of the anatomical connections between these regions has been related to trait empathy (Parkinson and Wheatley, 2014), and functional connectivity between the left IFG and the ACC is important during response selection and inhibition (Kemmotsu et al., 2005). Our current data shows that testosterone alters the connectivity in this specific network, possibly explaining our previous finding of reduced mindreading abilities during the RMET by testosterone (van Honk et al., 2011a). Since performance during the RMET depends on language as well as on perceptual detection abilities (Olderbak et al., 2015), future studies should aim to disentangle which of these aspects of the RMET testosterone targets.

An important question is whether our data relates to the atypical neural processing in ASC, in which testosterone may be an important etiologic factor (Baron-Cohen et al., 2015a, 2005). There is evidence for reduced long-range neural connectivity in people with ASC (Abrams et al., 2013; Vissers et al., 2012). In relation to the current data, a study by Kana et al. (2007) showed not only atypical activation in the ACC and SMA in people with ASC, but also reduced synchronization between the ACC, SMA, and IFG. Thus, the

same network that is affected in ASC is also affected by testosterone administration during mindreading.

In this light, our findings are also of interest in relation to the literature on the neuropeptide oxytocin and cognitive empathy (Bos et al., 2012b). Oxytocin, in some respects the opposite to testosterone, increases cognitive mindreading abilities (e.g. Domes et al., 2007; Theodoridou et al., 2013), and also shows opposite effects on functional connectivity (Bethlehem et al., 2013; Bos et al., 2012b). This fits with the view that oxytocin and testosterone have antagonistic properties in social-emotional behavior, and particular during cognitive empathy (Bos et al., 2012b; Domes et al., 2007; van Honk et al., 2011a). The first clinical studies in which oxytocin has been administered to individuals with ASC seem promising (Andari et al., 2010; Guastella et al., 2010), and these effects might have been caused by oxytocin acting on the same network presently targeted by testosterone, but in an opposite manner. However, to what extent our findings translate to ASC needs to be tested, considering that the current sample consisted of typical young women. Also, the current study consisted of a relatively small sample size, and therefore warrants replication in larger samples including males. In addition, it is important to stress the limitations of applying connectivity analyses in the context of fMRI using drug administration. Since fMRI inherently relies on relative changes, this methodology does not allow one to control for changes in baseline metabolic activity, or changes in overall resting state connectivity. Our effects should therefore be interpreted with care.

In sum, here we show that during emotion-recognition from the eye-region of the face, testosterone changes connectivity in a neural network that comprises the IFG, ACC, and SMA. Our findings have clinical relevance, since this network is critical for the integration of sensory information, selection, and action preparation (Desimone and Duncan, 1995; Hagoort, 2005; Liakakis et al., 2011; Lynch and Tian, 2006; Willems et al., 2009; Zhang et al., 2004), and is implicated in the symptomatology of ASC (Kana et al., 2007).

## Conflict of interest

The authors report no conflict of interest. The authors alone are responsible for the content and writing of the paper.

## Contributors

Peter A. Bos, Dennis Hofman, Erno J. Hermans, Estrella R. Montoya, Simon Baron-Cohen & Jack van Honk.

Authors PAB and EJH and JvH designed the study and wrote the protocol. Authors PAB, DH, EJH and ERM managed the literature searches and facilitated data analyses. Authors PAB, EJH, and ERM undertook the statistical analysis, and authors PAB, DH, SBC, and JvH contributed to the writing of the manuscript. All authors contributed to and have approved the final manuscript.

## Role of the funding source

The work in this paper was supported by a grant from the Netherlands Society of Scientific Research to PAB (451-14-015) and to JvH (056-24-010). SBC was supported by grants from the Medical Research Council (MRC), Wellcome Trust, and the Autism Research Trust. The funding sources had no further role in study design; in the collection, analysis and interpretation of the data; in the writing of the report; or the decision to submit the paper for publication.

## References

Abrams, D.A., Lynch, C.J., Cheng, K.M., Phillips, J., Supekar, K., Ryali, S., Uddin, L.Q., Menon, V., 2013. Underconnectivity between voice-selective cortex and reward circuitry in children with autism. Proc. Natl. Acad. Sci. U. S. A. 110, 12060–12065.

Adams, R.B., Rule, N.O., Franklin, R.G., Wang, E., Stevenson, M.T., Yoshikawa, S., Nomura, M., Sato, W., Kveraga, K., Ambady, N., 2010. Cross-cultural reading the mind in the eyes: an fMRI investigation. J. Cogn. Neurosci. 22, 97–108.

Andari, E., Duhamel, J.R., Zalla, T., Herbrecht, E., Leboyer, M., Sirigu, A., 2010. Promoting social behavior with oxytocin in high-functioning autism spectrum disorders. Proc. Natl. Acad. Sci. U. S. A. 107, 4389–4394.

Auyeung, B., Baron-Cohen, S., Chapman, E., Knickmeyer, R., Taylor, K., Hackett, G., 2006. Foetal testosterone and the child systemizing quotient. Eur. J. Endocrinol. 155, S123–S130.

Baron-Cohen, S., Ring, H.A., Wheelwright, S., Bullmore, E.T., Brammer, M.J., Simmons, A., Williams, S.C.R., 1999. Social intelligence in the normal and autistic brain: an fMRI study. Eur. J. Neurosci. 11, 1891–1898.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., Plumb, I., 2001. The Reading the Mind in the Eyes test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. J. Child. Psychol. Psychiatry 42, 241–251.

Baron-Cohen, S., Knickmeyer, R.C., Belmonte, M.K., 2005. Sex differences in the brain: implications for explaining autism. Science 310, 819–823.

Baron-Cohen, S., Ring, H., Chitnis, X., Wheelwright, S., Gregory, L., Williams, S., Brammer, M., Bullmore, E., 2006. fMRI of parents of children with Asperger Syndrome: a pilot study. Brain Cogn. 61, 122–130.

Baron-Cohen, S., Auyeung, B., Nørgaard-Pedersen, B., Hougaard, D., Abdallah, M., Melgaard, L., Cohen, A., Chakrabarti, B., Ruta, L., Lombardo, M., 2015a. Elevated fetal steroidogenic activity in autism. Mol. Psychiatry 20, 369–376.

Baron-Cohen, S., Bowen, D.C., Holt, R.J., Allison, C., Auyeung, B., Lombardo, M.V., Smith, P., Lai, M.-C., 2015b. The reading the mind in the eyes test: complete absence of typical sex difference in ~400 men and women with autism. PLoS One 10, e0136521.

Baron-Cohen, S., 1995. Mindblindness: An Essay on Autism and Theory of Mind. MIT Press.

Baron-Cohen, S., 2003. The Essential Difference. Basic Books, New York.

Baron-Cohen, S., 2009. Autism: the empathizing-systemizing (E-S) theory. Ann. N. Y. Acad. Sci. 1156, 68–80.

Bartlett, A.M., Ovaysikia, S., Logothetis, N.K., Hoffman, K.L., 2011. Saccades during object viewing modulate oscillatory phase in the superior temporal sulcus. J. Neurosci. 31, 18423–18432.

Bethlehem, R.A.I., van Honk, J., Auyeung, B., Baron-Cohen, S., 2013. Oxytocin, brain physiology, and functional connectivity: a review of intranasal oxytocin fMRI studies. Psychoneuroendocrinology 38, 962–974.

Bos, P.A., Hermans, E.J., Montoya, E.R., Ramsey, N.F., van Honk, J., 2010. Testosterone administration modulates neural responses to crying infants in young females. Psychoneuroendocrinology 35, 114–121.

Bos, P.A., Hermans, E.J., Ramsey, N.F., van Honk, J., 2012a. The neural mechanisms by which testosterone acts on interpersonal trust. NeuroImage 61, 730–737.

Bos, P.A., Panksepp, J., Bluthe, R.M., Honk, J.V., 2012b. Acute effects of steroid hormones and neuropeptides on human social-emotional behavior: a review of single administration studies. Front. Neuroendocrinol. 33, 17–35.

Breedlove, S.M., 2010. Minireview: organizational hypothesis: instances of the fingerpost. Endocrinology 151, 4116–4122.

Dal Monte, O., Schintu, S., Pardini, M., Berti, A., Wassermann, E.M., Grafman, J., Krueger, F., 2014. The left inferior frontal gyrus is crucial for reading the mind in the eyes: brain lesion evidence. Cortex 58, 9–17.

Derntl, B., Windischberger, C., Robinson, S., Kryspin-Exner, I., Gur, R.C., Moser, E., Habel, U., 2009. Amygdala activity to fear and anger in healthy young males is associated with testosterone. Psychoneuroendocrinology 34, 687–693.

Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. Ann. Rev. Neurosci. 18, 193–222.

Desmurget, M., Sirigu, A., 2009. A parietal-premotor network for movement intention and motor awareness. Trends Cogn. Sci. 13, 411–419.

Domes, G., Heinrichs, M., Michel, A., Berger, C., Herpertz, S.C., 2007. Oxytocin improves mind-reading in humans. Biol. Psychiatry 61, 731–733.

Eisenegger, C., Naef, M., Snozzi, R., Heinrichs, M., Fehr, E., 2010. Prejudice and truth about the effect of testosterone on human bargaining behaviour. Nature 463, 356–359.

Etkin, A., Egner, T., Kalisch, R., 2011. Emotional processing in anterior cingulate and medial prefrontal cortex. Trends Cogn. Sci. 15, 85–93.

Fan, Y., Duncan, N.W., de Greck, M., Northoff, G., 2011. Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. Neurosci. Biobehav. Rev. 35, 903–911.

Fleischman, D.S., Navarrete, C.D., Fessler, D.M., 2010. Oral contraceptives suppress ovarian hormone production. Psychol. Sci. 21, 750–752.

Frith, C.D., Frith, U., 1999. Interacting minds—a biological basis. Science 286, 1692–1695.

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

Guastella, A.J., Einfeld, S.L., Gray, K.M., Rinehart, N.J., Tonge, B.J., Lambert, T.J., Hickie, I.B., 2010. Intranasal oxytocin improves emotion recognition for youth with autism spectrum disorders. Biol. Psychiatry 67, 692–694.

Hadjikhani, N., Joseph, R.M., Snyder, J., Tager-Flusberg, H., 2007. Abnormal activation of the social brain during face perception in autism. Hum. Brain Mapp. 28, 441–449.

Hagoort, P., 2005. On Broca, brain, and binding: a new framework. Trends Cogn. Sci. 9, 416–423.

Kana, R.K., Keller, T.A., Minshew, N.J., Just, M.A., 2007. Inhibitory control in high-functioning autism: decreased activation and underconnectivity in inhibition networks. Biol. Psychiatry 62, 198–206.
Kemmotsu, N., Villalobos, M.E., Gaffrey, M.S., Courchesne, E., Müller, R.-A., 2005. Activity and functional connectivity of inferior frontal cortex associated with response conflict. Brain Res. Cogn. Brain Res. 24, 335–342.
Kohn, N., Eickhoff, S.B., Scheller, M., Laird, A.R., Fox, P.T., Habel, U., 2014. Neural network of cognitive emotion regulation—an ALE meta-analysis and MACM analysis. NeuroImage 87, 345–355.
Koolschijn, P.C.M.P., Peper, J.S., Crone, E.A., 2014. The influence of sex steroids on structural brain maturation in adolescence. PLoS One 9, e83929.
Liakakis, G., Nickel, J., Seitz, R.J., 2011. Diversity of the inferior frontal gyrus—a meta-analysis of neuroimaging studies. Behav. Brain Res. 225, 341–347.
Lombardo, M.V., Ashwin, E., Auyeung, B., Chakrabarti, B., Taylor, K., Hackett, G., Bullmore, E.T., Baron-Cohen, S., 2012. Fetal testosterone influences sexually dimorphic gray matter in the human brain. J. Neurosci. 32, 674–680.
Lynch, J.C., Tian, J.R., 2006. Cortico-cortical networks and cortico-subcortical loops for the higher control of eye movements. Prog. Brain Res. 151, 461–501.
Moor, B.G., d. Macks, Z.A.O., Güroglu, B., Rombouts, S.A.R.B., V.d. Molen, M.W., Crone, E.A., 2012. Neurodevelopmental changes of reading the mind in the eyes. Soc. Cogn. Affect. Neurosci. 7, 44–52.
Neggers, S.F., Hermans, E.J., Ramsey, N.F., 2008. Enhanced sensitivity with fast three-dimensional blood-oxygen-level-dependent functional MRI: comparison of SENSE-PRESTO and 2D-EPI at 3 T. NMR Biomed. 21, 663–676.
Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M.W., Roberts, R.D., 2015. A psychometric analysis of the reading the mind in the eyes test: toward a brief form for research and applied settings. Front. Psychol. 6, 1503.
Olsson, A., Kopsida, E., Sorjonen, K., Savic, I., 2016. Testosterone and estrogen impact social evaluations and vicarious emotions: a double-blind placebo-controlled study. Emotion, http://dx.doi.org/10.1037/a0039765 (epub ahead of print).
Parkinson, C., Wheatley, T., 2014. Relating anatomical and social connectivity: white matter microstructure predicts emotional empathy. Cereb. Cortex 24, 614–625.
Peper, J.S., Mandl, R.C.W., Braams, B.R., de Water, E., Heijboer, A.C., Koolschijn, P.C.M.P., Crone, E.A., 2013. Delay discounting and frontostriatal fiber tracts: a combined DTI and MTR study on impulsive choices in healthy young adults. Cereb. Cortex 23, 1695–1702.
Pessoa, L., 2008. On the relationship between emotion and cognition. Nat. Rev. Neurosci. 9, 148–158.
Rametti, G., Carrillo, B., Gómez-Gil, E., Junque, C., Zubiaurre-Elorza, L., Segovia, S., Gomez, A., Karadi, K., Guillamon, A., 2012. Effects of androgenization on the white matter microstructure of female-to-male transsexuals. A diffusion tensor imaging study. Psychoneuroendocrinology 37, 1261–1269.
Schmithorst, V.J., Holland, S.K., 2007. Sex differences in the development of neuroanatomical functional connectivity underlying intelligence found using Bayesian connectivity analysis. NeuroImage 35, 406–419.
Shacham, S., 1983. A shortened version of the Profile of Mood States. J. Personal. Assess. 47, 305–306.
Shenhav, A., Botvinick, M.M., Cohen, J.D., 2013. The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron 79, 217–240.
Theodoridou, A., Penton-Voak, I.S., Rowe, A.C., 2013. A direct examination of the effect of intranasal administration of oxytocin on approach-avoidance motor responses to emotional stimuli. PLoS One 8, e58113.
Tuiten, A., Van Honk, J., Koppeschaar, H., Bernaards, C., Thijssen, J., Verbaten, R., 2000. Time course of effects of testosterone administration on sexual arousal in women. Arch. Gen. Psychiatry 57, 146–149.
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15, 273–289.
Vissers, M.E., Cohen, M.X., Geurts, H.M., 2012. Brain connectivity and high functioning autism: a promising path of research that needs refined models methodological convergence, and stronger behavioral links. Neurosci. Biobehav. Rev. 36, 604–625.
Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., Weil, L., Wakabayashi, A., 2006. Predicting autism spectrum quotient (AQ) from the systemizing quotient-Revised (SQ-R) and empathy quotient (EQ). Brain Res. 1079, 47–56.
Willems, R.M., Ozyürek, A., Hagoort, P., 2009. Differential roles for left inferior frontal and superior temporal cortex in multimodal integration of action and language. NeuroImage 47, 1992–2004.
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.
van Honk, J., Schutter, D.J., Bos, P.A., Kruijt, A.-W., Lentjes, E.G., Baron-Cohen, S., 2011a. Testosterone administration impairs cognitive empathy in women depending on second-to-fourth digit ratio. Proc. Natl. Acad. Sci. U. S. A. 108, 3448–3452.
van Honk, J., Terburg, D., Bos, P.A., 2011b. Further notes on testosterone as a social hormone. Trends Cogn. Sci. 15, 291–292.
van Rooij, K., Bloemers, J., de Leede, L., Goldstein, I., Lentjes, E., Koppeschaar, H., Olivier, B., Tuiten, A., 2011. Pharmacokinetics of three doses of sublingual testosterone in healthy premenopausal women. Psychoneuroendocrinology 37, 773–781.
van Wingen, G.A., Mattern, C., Verkes, R.J., Buitelaar, J., Fernandez, G., 2010. Testosterone reduces amygdala-orbitofrontal cortex coupling. Psychoneuroendocrinology 35, 105–113.
Zhang, J.X., Feng, C.-M., Fox, P.T., Gao, J.-H., Tan, L.H., 2004. Is left inferior frontal gyrus a general mechanism for selection? NeuroImage 23, 596–603.