Psychoneuroendocrinology (2015) 56, 190—199

Available online at www.sciencedirect.com

ScienceDirect

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

Puberty suppression and executive functioning: An fMRI-study in adolescents with gender dysphoria

Annemieke S. Staphorsius[a], Baudewijntje P.C. Kreukels[b], Peggy T. Cohen-Kettenis[b], Dick J. Veltman[c], Sarah M. Burke[a,b], Sebastian E.E. Schagen[d,e], Femke M. Wouters[d], Henriëtte A. Delemarre-van de Waal[e,1,2], Julie Bakker[a,f,*,1]

[a] *Neuroendocrinology Group, Netherlands Institute for Neuroscience, Meibergdreef 47, 1105 BA Amsterdam, The Netherlands*
[b] *Center of Expertise on Gender Dysphoria, Department of Medical Psychology, Neuroscience Campus Amsterdam, VU University Medical Center, De Boelelaan 1117, 1081 HZ Amsterdam, The Netherlands*
[c] *Department of Psychiatry, Neuroscience Campus Amsterdam, VU University Medical Center, De Boelelaan 1117, 1081 HZ Amsterdam, The Netherlands*
[d] *Department of Pediatric Endocrinology, Neuroscience Campus Amsterdam, VU University Medical Center, De Boelelaan 1117, 1081 HZ Amsterdam, The Netherlands*
[e] *Department of Pediatrics, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands*
[f] *GIGA Neurosciences, University of Liège, Avenue de l'Hôpital 1B36, 4000 Liège, Belgium*

Received 21 August 2014; received in revised form 3 March 2015; accepted 4 March 2015

KEYWORDS
fMRI;
Adolescents;
Gonadotropin releasing hormone analog;
Tower of London;
Gender dysphoria;
Sex difference

Summary Adolescents with gender dysphoria (GD) may be treated with gonadotropin releasing hormone analogs (GnRHa) to suppress puberty and, thus, the development of (unwanted) secondary sex characteristics. Since adolescence marks an important period for the development of executive functioning (EF), we determined whether the performance on the Tower of London task (ToL), a commonly used EF task, was altered in adolescents with GD when treated with GnRHa. Furthermore, since GD has been proposed to result from an atypical sexual differentiation of the brain, we determined whether untreated adolescents with GD showed sex-atypical brain activations during ToL performance. We found no significant effect of GnRHa on ToL performance scores (reaction times and accuracy) when comparing GnRHa treated

* Corresponding author at: University of Liège, GIGA Neurosciences, Avenue de l'Hôpital 1B36, 4000 Liege, Belgium. Tel.: +32 43665978; fax: +32 43665953.
E-mail address: jbakker@ulg.ac.be (J. Bakker).
[1] These authors shared senior authorship.
[2] Deceased (February 2014). Delemarre-van de Waal was the principal investigator of this study and has commented on the penultimate version of this manuscript.

http://dx.doi.org/10.1016/j.psyneuen.2015.03.007

male-to-females (suppressed MFs, $n = 8$) with untreated MFs ($n = 10$) or when comparing GnRHa treated female-to-males (suppressed FMs, $n = 12$) with untreated FMs ($n = 10$). However, the suppressed MFs had significantly lower accuracy scores than the control groups and the untreated FMs. Region-of-interest (ROI) analyses showed significantly greater activation in control boys ($n = 21$) than control girls ($n = 24$) during high task load ToL items in the bilateral precuneus and a trend ($p < 0.1$) for greater activation in the right DLPFC. In contrast, untreated adolescents with GD did not show significant sex differences in task load-related activation and had intermediate activation levels compared to the two control groups. GnRHa treated adolescents with GD showed sex differences in neural activation similar to their natal sex control groups. Furthermore, activation in the other ROIs (left DLPFC and bilateral RLPFC) was also significantly greater in GnRHa treated MFs compared to GnRHa treated FMs. These findings suggest that (1) GnRHa treatment had no effect on ToL performance in adolescents with GD, and (2) pubertal hormones may induce sex-atypical brain activations during EF in adolescents with GD.

1. Introduction

Gender dysphoria (GD) is a condition in which people suffer from an incongruence between their natal sex and their gender identity, *i.e.* their experienced gender (American Psychiatric Association, 2013). For young individuals with GD, puberty is a period that causes great distress because it is characterized by unwanted physical changes, the development of the secondary sex characteristics. Therefore puberty inhibiting hormones may be prescribed such as the gonadotropin releasing hormone analogs (GnRHa) leuprolide or triptorelin (Gooren and Delemarre-van de Waal, 1996).

Some researchers expressed concerns about the possible disadvantages of GnRHa administration during adolescence (Spriggs, 2004; Viner, 2005; Houk and Lee, 2006; Korte et al., 2008). They fear that it may lead to misdiagnosis or that adolescents cannot make complex life decisions. Moreover, some have questioned whether hormonal suppression affects psychological functioning and if it may entail medical risks. Indeed, during adolescence the brain is still developing. Furthermore, puberty has been suggested to represent a second organizational period during brain development in rodents (Juraska et al., 2013) and in humans (Romeo, 2003; Sisk and Zehr, 2005). The prefrontal cortex (PFC) in particular appears to develop much later than other brain areas (Huttenlocher, 1979). Histological studies suggest that there is a second wave of synaptic proliferation in the PFC at the onset of puberty (Huttenlocher, 1979; Bourgeois et al., 1994; Woo et al., 1997), followed by a plateau phase and synaptic pruning. Executive functioning (EF), which is believed to depend heavily on prefrontal activation, also develops relatively slowly. For instance, performance on the Tower of London task (ToL), a frequently used EF task, improves with age until early adulthood (De Luca et al., 2003; Huizinga et al., 2006; Asato et al., 2006; Albert and Steinberg, 2011).

Since puberty marks an important period in the development of EF, the question arises if pubertal suppression affects this development. Therefore, in the present study, adolescents with GD who received GnRHa to suppress their puberty were compared with a group of control adolescents regarding ToL performance and brain activation patterns (using functional magnetic resonance imaging, fMRI). To check whether potential differences between the groups were due to the suppression (and not due to GD), we also compared them with a group of age matched adolescents with GD who were not — yet — using GnRHa but were already in puberty.

Most of the previous ToL neuroimaging studies did not report — or perhaps did not look for — any sex effects (Owen et al., 1996; Baker et al., 1996; Dagher et al., 1999; Lazeron et al., 2000; Rowe et al., 2001; Van den Heuvel et al., 2003; Newman et al., 2003; Schall et al., 2003; Wagner et al., 2006; Boghi et al., 2006). However, one study reported sex differences in precuneus and dorsolateral prefrontal cortex (DLPFC) activation (Boghi et al., 2006). Therefore we examined sex differences as well.

Furthermore, it has been hypothesized that sexual differentiation of the brain might be different in individuals with GD (Cohen-Kettenis and Gooren, 1999; Van Goozen et al., 2002; Swaab, 2004). Functional neuroimaging studies comparing adults with GD (before the start of treatment) to controls demonstrated that MFs differed from their natal sex in parietal activation during a mental rotation task (Schöning et al., 2010) and showed female-like activity during processing of erotic stimuli (Gizewski et al., 2009) and after exposure to androstadienone, an odorous steroid compound (Berglund et al., 2008). In a verbal fluency study with adolescents performed by our group (Soleman et al., 2013), activation levels of untreated FMs and MFs fell in between those of the control groups. Structural neuroimaging studies have also shown intermediate values in adult FMs and MFs compared to control groups (Rametti et al., 2011a,b; Kranz et al., 2014) and several structural studies have shown differences between adults with GD and controls sharing their natal sex (Luders et al., 2009, 2012; Simon et al., 2013; Zubiaurre-Elorza et al., 2013; Hoekzema et al., 2015) although another study reported brain volumes largely in line with their natal sex (Savic and Arver, 2011).

As mentioned above, our group has examined the effect of GD on VF performance and brain activation in untreated adolescents (Soleman et al., 2013). Although the VF task may be considered an executive functioning task, the effect of GnRHa treatment on VF performance and brain activation was not investigated. In this study we examined if ToL-related brain activation of adolescents with GD, before start of GnRHa and while on GnRHa, was more in line with that of individuals of their experienced gender or of their natal sex. We believe that the present study is the first to examine the effects of puberty suppression on executive functioning.

2. Methods

2.1. Subjects

Adolescents who were diagnosed with Gender Identity Disorder according to the DSM-IV-TR (American Psychiatric Association, 2000) at VU University Medical Center in Amsterdam were recruited (Kreukels and Cohen-Kettenis, 2011). During preparation of this manuscript the DSM-5 was published (American Psychiatric Association, 2013), therefore DSM-5 terminology is used throughout this manuscript.

Forty-one adolescents with GD were included in this study; 22 female-to-males, 12 of which were using GnRHa (suppressed FM) and 10 who were not (untreated FM) and 18 male-to-females, of which 8 were using GnRHa (suppressed MF) and 10 were not (untreated MF). The suppressed adolescents with GD had been receiving 3.75 mg Triptorelin (Decapeptyl-CR®) every 4 weeks, subcutaneously or intramuscularly (mean duration ± standard deviation: 1.6 ± 1.0 years). To receive GnRHa, participants had to be at least 12 years old. Furthermore, girls needed to have breast development as described in Tanner stage B2 (Marshall and Tanner, 1969) and the genital development of the boys had to be at Tanner stage G2-G3 (testicular volume of 6—8 ml) (Marshall and Tanner, 1970) with measurable estradiol and testosterone levels, respectively. Relatives and friends of the participants were asked to participate, serving as age-matched controls. Only 3 siblings participated as controls: both a brother and sister of one GnRHa treated FM and one sister of an untreated FM. Thus, the majority of the controls were friends. The control group consisted of 24 girls (F) and 21 boys (M). Subject characteristics are presented in Table 1. According to the Declaration of Helsinki, all participants and their legal guardians gave their informed consent, and the study was approved by the Ethics Committee of the VU University Medical Center Amsterdam.

For all groups exclusion criteria were: (1) insufficient command of the Dutch language, (2) unadjusted endocrine disorders, (3) neurological or psychiatric disorders that could lead to deviant test results, (4) use of psychotropic medication, and (5) contra indications for an MRI scan. Adolescents receiving puberty delaying medication or any form of hormones besides oral contraceptives were excluded as controls. In consultation with the treating clinicians, only adolescents with GD functioning within the normal range were asked to participate in the study. One instrument (amongst other things) to measure psychological functioning was the Dutch translation of the Child Behavior Check List (Achenbach and Edelbrock, 1983; Verhulst et al., 1996), a well-known parent report questionnaire measuring psychological and behavioral problems. Control subjects received the CBCL as part of this study (average CBCL scores of the groups are depicted in Table 1). The gender identity of all controls was in line with their natal sex and was checked by asking them if they felt they belonged to the other gender or wished to be the other gender. All controls had a heterosexual orientation. The adolescents with GD were all sexually attracted to partners of their natal sex.

From the initial selection ten subjects had to be removed from further data analysis due to excessive movement during scanning, two because of scan artifacts (MR signal dropout) due to braces, fifteen due to insufficient mask coverage, two because of performance at chance level, and one due to scanner failure.

2.2. Experimental setup and procedure

In this study an event-related parametric version of the ToL was used (for a detailed description see Van den Heuvel et al., 2003). On each trial, a start configuration (top) and a target configuration (bottom) were displayed simultaneously (see Fig. 1). In the planning condition subjects were asked to work out the minimum number of steps (ranging from 1 to 5) required to reach the target configuration. As a baseline condition, participants had to count the total amount of blue and yellow beads. The task lasted about 12 min and timing of the stimuli was self-paced, with a maximum response duration of 60 s per trial.

Participants practiced the task outside the scanner and performed some practice trials inside the scanner immediately before starting the task. Three other cognitive tasks were performed as well, a verbal fluency task (Soleman et al., 2013), a mental rotation task and a face recognition task (data to be published elsewhere). The four tasks were presented randomly during the scanning session and the entire session lasted 1 h. Prior to the MRI session a physical examination was performed by a clinician and intelligence was estimated with four subscales (arithmetic, vocabulary, picture arrangement, and block design) of the Wechsler Intelligence Scale for Children, third edition (WISC-III®) (Wechsler, 1991) or the Wechsler Adult Intelligence Scale, third edition (WAIS-III®) (Wechsler, 1997), depending on the participant's age. Handedness was measured by means of a Dutch questionnaire (Van Strien, 1992).

2.3. MRI acquisition

Imaging data were acquired on a 3.0 T Philips Intera (Best, The Netherlands) MRI scanner at the Academic Medical Center, Amsterdam, The Netherlands. Axial T2*-weighted whole-brain volumes sensitive to blood oxygen level dependent (BOLD) contrast (Ogawa et al., 1990) were acquired

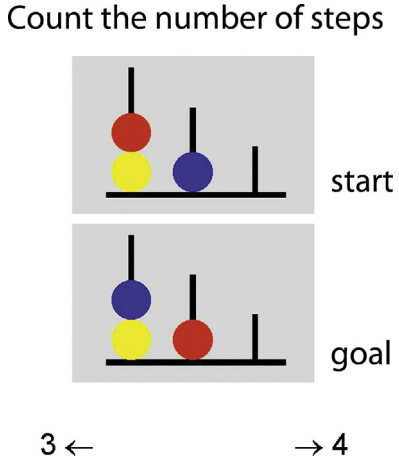

Figure 1 Example of the Tower of London task showing a trial in the planning condition.

Table 1 Sample characteristics and performance data.

Group	*n*	Age (years) (mean ± SD)	IQ* (mean ± SD)	Tanner stage (mean ± SD)	CBCL scores* (mean ± SD)	Accuracy (%)* (mean ± SD)	RT (s) (mean ± SD)
M	21	14.9 ± 1.5	110.7 ± 15.1	4.2 ± 1.2	48.4 ± 10.5■	88.5 ± 6.8	9.6 ± 2.5
F	24	14.4 ± 1.8	103.0 ± 17.3	4.3 ± 0.9	48.4 ± 10.3◦	87.2 ± 11.9	9.0 ± 1.8
MF (total)	18	15.1 ± 2.4	102.6 ± 18.5	3.9 ± 1.1	57.8 ± 9.2	79.1 ± 10.3	10.4 ± 3.5
suppressed	8	15.4 ± 0.7	94.0 ± 10.3	4.1 ± 1.0	57.4 ± 9.8	73.9 ± 9.1	10.9 ± 4.1
untreated	10	14.6 ± 3.2	109.4 ± 21.2	3.8 ± 1.1	58.2 ± 9.3	83.4 ± 9.5	9.9 ± 3.1
FM (total)	22	15.8 ± 1.9	97.1 ± 15.4	4.5 ± 0.9	60.4 ± 10.2	87.1 ± 10.0	10.0 ± 2.6
suppressed	12	16.1 ± 1.7	95.8 ± 15.6	4.1 ± 1.1	57.5 ± 9.4	85.7 ± 10.5	9.9 ± 3.1
untreated	10	15.4 ± 2.3	98.5 ± 15.9	4.9 ± 0.3	63.9 ± 10.5	88.8 ± 9.7	10.0 ± 2.0

n: number of subjects, SD: standard deviation, accuracy: percentage of correct trials, corrected for task load, RT: reaction times in seconds, corrected for task load, CBCL: Child Behavior Checklist.

* Significantly different between groups ($p < .05$, two-sided).

■ $n = 14$.

◦ $n = 16$.

M: control boys, F: control girls, MF: male adolescents with GD, FM: female adolescents with GD, suppressed: treated with GnRHa, untreated: without GnRHa treatment.

over ±12 min using an echo-planar imaging sequence (repetition time [TR] 2.3 s; echo time [TE] 30 ms; field of view: 22 cm × 22 cm × 10.5 cm; flip angle: 80 degrees; 96 × 96 matrix). A sagittal T1-weighted scan was also performed (repetition time [TR] 9 ms; echo time [TE] 3.5 ms; field of view: 25.6 cm × 23.2 cm × 17.0 cm; flip angle: 8 degrees; 256 × 256 matrix, 170 slices).

2.4. Statistical analysis

Subject characteristics and ToL performance data were analyzed with the Statistical Package for the Social Sciences (SPSS), version 21. Accuracy scores (percentage of correct trials) and reaction times (RT) were corrected for task load by multiplying the scores of category 1—5 with increasing weights (1.0, 2.0, 2.5, 3.0 and 3.5, respectively — analogous to the contrast weights for measuring task load activation during first level SPM analysis) and then dividing the sum by 12. Group differences in age, IQ, and accuracy were tested with a one-way analysis of variance (ANOVA) and *post hoc* comparisons were performed using Games—Howell correction. One-way analysis of covariance (ANCOVA) was performed to examine the effect of IQ on group differences in accuracy. Group differences in RT were examined using Kruskal—Wallis test because the assumption of normality was not met. Tanner stage was examined using Kruskal—Wallis tests and a Chi-square test was used to check for group differences in handedness. A Pearson product-moment correlation was computed to determine the relationship between IQ scores and accuracy scores. The relation between IQ and RT was assessed using Spearman's Rank Order correlation.

The fMRI analysis was carried out with Statistical Parametric Mapping 8 (SPM8, Wellcome Department of Imaging Neuroscience, Institute of Neurology at the University College London, UK) implemented in MATLAB R2011b (MathWorks Inc., Natick, MA, USA). Functional images were slice-timed, realigned to the mean image, and co-registered with the individual anatomical image. Because the participants in this study were adolescents and thus did not have adult-sized brains yet, a DARTEL template of their structural scans was created for optimal spatial normalization into Montreal Neurological Institute (MNI) space (DARTEL: Diffeomorphic Anatomical Registration Through Exponentiated, see Ashburner, 2007). Functional images were smoothed with a 8 mm FWHM Gaussian filter.

First level contrast images for *planning* were calculated by subtracting the baseline condition from the 5 planning categories. Moreover, to identify brain regions that showed signal intensity variation correlated with increasing planning complexity a *task load* contrast was calculated by giving increasing contrast weights to category 1 to 5 (Van den Heuvel et al., 2003). Individual head jerks of more than 1 mm were included in every first-level design matrix (Lemieux et al., 2007) together with the six motion parameters to account for the effects of excessive head motion. Furthermore, error trials were added as regressors of no interest.

Due to the large amount of data, only the results for the *task load* contrast are displayed in Section 3, although analyses were performed for both contrasts. *Task load* was favored because it places a greater emphasis on the more difficult trials and especially those trials (involving 3 or more steps) require EF. The results for *planning* are displayed in the Supplementary Results.

The contrast images for *planning* and *task load* were entered into a second level analysis of variance. IQ scores were added as a covariate because the groups differed in IQ scores. *T*-tests were performed to investigate the activation seen in every group (main effect) and to investigate group differences. The results were examined on whole-brain level first and subsequently region of interest (ROI) analyses were performed. Based on previous ToL neuroimaging studies (see Section 1) the dorsolateral- and rostrolateral prefrontal cortex (DLPFC and RLPFC), and precuneus were chosen as ROIs. These areas were selected from the IBASPM116 atlas and separate left and right masks were created using WFU Pickatlas version 3.0.4. (Maldjian et al., 2003). Because the

frontal ROI did not distinguish between DLPFC and RLPFC, a functional RLPFC ROI from the BrainMap database (Nielsen and Hansen, 2002) was subtracted from the frontal ROI using MarsBar (version 0.43, MRC Cognition and Brain Sciences Unit, Cambridge, UK). Each set of ROIs was masked with the FWE-corrected ($p = 0.05$) main effects for *planning* and *task load* (separately). For examination of the main effect a p-value of 0.05 corrected for multiple comparisons was used (pFWE-corrected = 0.05). For the between group ROI analyses a pFWE-corrected = 0.05 was used as well, corrected for the spatial extent of the ROI.

3. Results

3.1. Sample data

No significant age differences were found between the six groups ($F(5, 79) = 1.52$, NS), but a difference was observed in IQ ($F(5, 79) = 2.32$, $p < .05$). Control boys (M) had significantly higher IQ scores than suppressed MFs ($p = .03$). Tanner stage and handedness did not differ between the groups ($p = 0.207$ and $p = 0.647$, respectively). The means and standard deviations of age, IQ and Tanner stage are presented in Table 1. There was no significant difference in duration of suppression between MFs (mean duration ± standard deviation: 1.8 ± 0.8 years) and FMs (mean duration ± standard deviation: 1.4 ± 1.1 years); $T(18) = 1.03$, *NS*.

3.2. ToL performance data

Accuracy significantly differed between the groups ($F(5, 79) = 3.07$, $p < .05$). *Post hoc* analyses showed that the suppressed MFs had significantly lower accuracy scores than the control groups ($p = .02$ compared to control boys and $p = .04$ compared to control girls) and the untreated FMs ($p = .04$). IQ and accuracy were significantly correlated ($r = 0.31$, $n = 85$, $p < .005$), but even after correcting for IQ, a significant effect of group on accuracy remained ($F(5, 78) = 2.70$, $p < .05$). Additionally, there was a significant negative correlation between IQ and RT ($r_s(85) = -0.31$, $p < .005$). However, RT did not significantly differ between the six groups ($H(5) = 3.92$, NS). No significant correlations between age and the performance scores were found. Means and standard deviations of accuracy and RT are presented in Table 1. For the baseline condition (counting the blue and yellow balls) no significant group differences were found for accuracy ($F(5, 79) = 0.28$, NS) or RT ($F(5, 79) = 1.16$, NS).

3.3. Main effect functional MRI data

The results for the *planning* contrast can be found in the Supplementary Results. *Task load* (Table 2 and Fig. 2) showed a robust activation pattern in the bilateral DLPFC and the left supplementary motor area. The left precentral area was significantly activated as well and activation was seen in the bilateral insular cortices and right pars opercularis. The parietal activation in the left hemisphere was found in the superior gyrus, extending from the precuneus to the more lateral part of the superior parietal cortex. In the right hemisphere significant parietal activation was seen in the supramarginal and inferior parietal gyrus. Task load was also associated with significant activation of bilateral basal ganglia (caudate nucleus, putamen and globus pallidus).

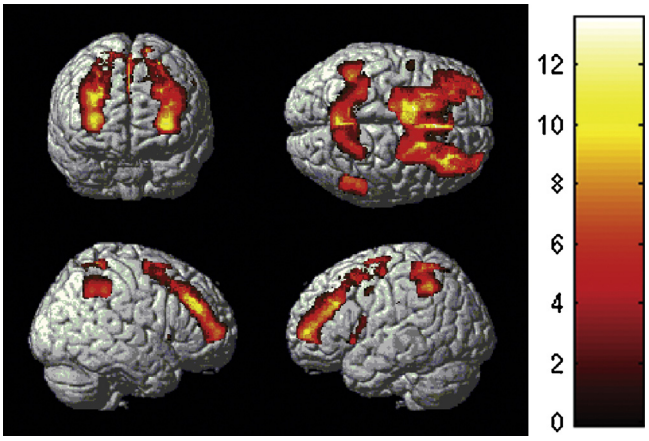

Figure 2 Brain activation pattern in all participants (main effect) for task load, with IQ added as a covariate. *P*-value (FWE-corrected) = 0.05.

3.4. Group effects functional MRI data

3.4.1. Task load activation: sex differences within groups

Whole-brain analyses for *task load* revealed no significant group differences. ROI analyses showed greater activation of the bilateral precuneus and a trend ($p < .10$) for greater activation in the right DLPFC in control boys compared to control girls. In these ROIs similar sex differences were found between the suppressed MFs and FMs. Furthermore, activation in the other ROIs (bilateral RLPFC and left DLPFC) was also greater in suppressed MFs than suppressed FMs. In contrast, no ROIs showed greater activation in the untreated MFs compared to the untreated FMs. The only sex difference between the untreated adolescents with GD was in the opposite direction; the untreated FMs showed a slightly more pronounced ($p < .10$) right DLPFC activation compared to the untreated MFs (An overview of the sex differences reported in this paragraph is given in Table 3.) To explore whether the lack of typical sexual differentiation in untreated adolescents with GD was due to activation levels being in between those of the control groups, we depicted the activation levels of the untreated GD adolescents in those voxels showing the greatest sex difference in controls (Fig. 3). Indeed, plotting effect sizes indicated that activations in untreated GD adolescents were intermediate between the two control groups.

3.4.2. Task load activation in MFs

The suppressed MFs showed greater activation compared to their experienced gender (F) in bilateral DLPFC, left RLPFC, left precuneus and right precuneus (trend), whereas untreated MFs only displayed a trend for greater activation in the right precuneus. The suppressed MFs not only showed a greater left RLPFC activation than Fs, but also relative to their natal sex (M) and untreated MFs. The suppressed MFs

Table 2 Brain regions in main effect of task load.

Brain regions — task load	# Voxels	MNI coordinates			*T*-value	*p* (FWE-corrected)
		x	*y*	*z*		
Frontal	**27,359**					
Supp_Motor_Area_L		−14	5	66	13.57	<0.001
Frontal_Mid_R		37	38	31	12.25	<0.001
Frontal_Mid_L		−33	32	31	11.89	<0.001
Parietal_L	**7465**					
Parietal_Sup_L		−29	−43	66	9.85	<0.001
Precuneus_L		−6	−60	46	9.85	<0.001
Precuneus_L		−11	−57	63	9.64	<0.001
Parietal_R	**1213**					
Supramarginal_R		54	−42	45	7.81	<0.001
Parietal_Inf_R		52	−54	46	7.14	<0.001
Basal Ganglia_L	**421**					
Putamen_L		−18	0	12	7.26	<0.001
Pallidum_L		−18	−6	0	5.83	<0.01
Basal Ganglia_R						
Caudate_R	103	15	2	16	6.62	<0.001
Frontal						
Precentral_L	56	−50	−1	45	6.04	<0.01
Frontal_Inferior_L	**477**					
Insula_L		−33	15	1	6.04	<0.01
Frontal_Inf_Oper_L		−53	6	12	5.88	<0.01
Insula_L		−29	23	−3	5.44	<0.05
Insula_R						
Insula_R	23	33	18	3	5.32	<0.05

Results are FWE corrected $p < .05$ and cluster size is > 20 voxels. Bold numbers in column 2 represent the number of voxels in the entire cluster.

showed greater left DLPFC activation than the untreated MFs as well.

3.4.3. Task load activation in FMs

Suppressed FMs differed from their experienced gender (M) by showing less bilateral precuneus activation, corresponding with the activation differences between the control groups. The untreated FMs did not show this resemblance to their natal sex (F). Besides lower right precuneus activation than boys (M), suppressed FMs also showed lower activation of this area than girls (F).

The untreated, but not the suppressed, FMs showed a trend for greater right DLPFC activation than their natal sex (F), thus showing a similarity to their experienced gender (M), who also demonstrated this trend compared to Fs. Untreated FMs displayed greater bilateral precuneus activation than suppressed FMs.

4. Discussion

In this study, we aimed to determine whether puberty suppression affected ToL performance. We found no significant effect of GnRHa on ToL performance scores (reaction times and accuracy) in either MFs or FMs when compared to untreated adolescents with GD. However, suppressed MFs had the lowest accuracy scores, which, as the analysis of covariance pointed out, did not just reflect their IQ scores, which were the lowest as well. It is possible that this is just a chance finding due to the small size of this subgroup ($n = 8$). No sex differences in performance were found in the control groups.

ROI analysis did reveal sex differences in brain activations associated with ToL performance. Control boys showed significantly greater activation in the bilateral precuneus and right DLPFC (trend) during high task load compared to control girls. In a previous study (Boghi et al., 2006) adults showed similar sex differences in the precuneus, whereas the sex difference in DLPFC activation was reversed; women exhibited greater DLPFC activation than men. A possible explanation for this discrepancy is that the DLPFC is not yet fully developed in our participants. In a Go-No-Go study children displayed greater activation of the DLPFC than adults, this was explained as resulting from greater network efficiency in adults (Casey et al., 1997). Since frontal gray matter starts developing earlier in girls than in boys (Giedd, 2008) network fine-tuning may start earlier as well. During adolescence the DLPFC may still be under the influence of pubertal hormonal effects, either activational or organizational (Romeo, 2003; Sisk and Zehr, 2005) whereas this is no longer the case for the precuneus, since a strong bilateral sex difference is present both in adolescents (present study) and adults (Boghi et al., 2006).

It has been hypothesized that the sexual differentiation of the brain in individuals with GD may be distinct from other members of their natal sex due to organizational effects

Table 3 Voxels showing sex differences for task load activation (corrected for IQ).

Region of interest	Sex difference	MNI coordinates			*T*-value	*p* (FWE-corrected)
		x	*y*	*z*		
Precuneus L	M > F	−6	−67	51	4.76	<0.01
		−8	−57	46	3.92	<0.05
		−2	−46	46	3.38	<0.10
	MF (s) > FM (s)	−6	−52	48	3.57	<0.05
		−11	−49	46	3.55	<0.05
		−14	−61	60	3.33	<0.10
		−3	−58	48	3.24	<0.10
Precuneus R	M > F	15	−49	60	3.71	<0.05
		15	−43	58	3.63	<0.05
		2	−43	43	3.52	<0.05
		2	−46	48	3.36	<0.10
		8	−43	60	3.31	<0.10
		5	−45	51	3.21	<0.10
	MF (s) > FM (s)	12	−60	45	3.32	<0.10
		8	−54	54	3.30	<0.10
DLPFC L	MF (s) > FM (s)	−18	23	39	5.06	<0.01
		−21	23	51	4.22	<0.05
		−44	38	19	3.96	<0.05
		−17	26	52	3.95	<0.05
		−20	6	49	3.80	<0.10
		−20	18	55	3.65	<0.10
DLPFC R	M > F	27	−1	60	3.76	<0.10
	MF (s) > FM (s)	34	38	25	4.75	<0.01
	FM (u) > MF (u)	*18*	*12*	*45*	*3.59*	*<0.10*
RLPFC L	MF (s) > FM (s)	−27	59	7	4.11	<0.01
		−23	51	10	3.68	<0.05
		−23	54	0	3.46	<0.05
		−26	56	19	3.46	<0.05
		−24	53	16	3.32	<0.05
RLPFC R	MF (s) > FM (s)	30	50	−2	3.26	<0.05
		26	50	15	3.02	<0.10

MNI coordinates are given for the voxels showing a sex difference between males and females within the same group, *e.g.* male *versus* female controls. Indicated in Italic is a reversed sex difference. For these comparisons, ROIs were used (indicated in the first column) since no results were obtained at whole-brain level. All sex differences reported are FWE corrected $p < .10$. M: control boys, F: control girls, MF: male adolescents with GD, FM: female adolescents with GD, (s): suppressed by GnRHa, (u): untreated.

of sex hormones (Cohen-Kettenis and Gooren, 1999; Van Goozen et al., 2002; Swaab, 2004). This was based on findings that the development of the sexual organs and the differentiation of the brain follow separate time courses during prenatal development, implying different time windows during which these processes can be affected. Plotting effect sizes in the present study showed that brain activation levels of the untreated adolescents with GD fell in-between those of the two control groups in the areas that showed significant sex differences in the controls (Fig. 3). Hence, untreated MFs and FMs had a closer resemblance to each other than the control groups and no sex differences were found. Similar results were found in the VF study performed by our group (Soleman et al., 2013), where the controls showed a sex difference in right rolandic operculum activation but the untreated adolescents with GD, who showed intermediate activation compared to the control groups, did not. As proposed by the sexual differentiation hypothesis of GD (Cohen-Kettenis and Gooren, 1999; Van Goozen et al., 2002; Swaab, 2004), the absence of a sex difference in untreated GD might be a result of a different hormonal milieu during prenatal development. However, possible effects of pubertal hormones on establishing atypical differentiation cannot be ruled out based on the results of the untreated participants. To this end, examination of sexual differentiation in puberty suppressed adolescents with GD, as was performed in the present study, provided a useful model. Interestingly, the suppressed MFs showed greater activation than the suppressed FMs in the same ROIs that were more active in control boys than control girls, indicating sex-typical brain activations. This similarity to their natal sex was also observed when comparing the suppressed adolescents with GD to the control groups. Like control boys, suppressed MFs showed greater ROI activation than control

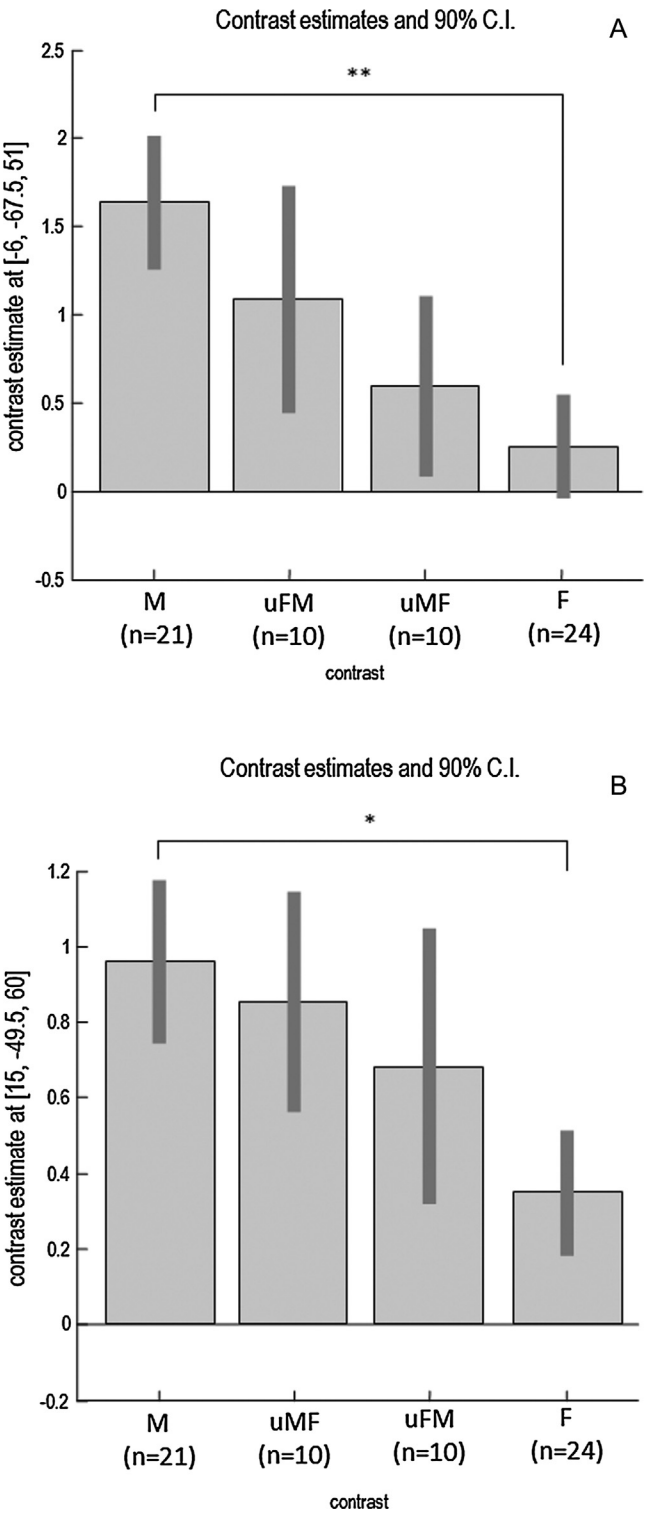

Figure 3 Contrast estimates and 90% confidence intervals (dark gray bars) for task load in the control groups and the untreated adolescents with GD in (A) the left precuneus (MNI coordinates: −6, 67.5, 51) and (B) the right precuneus (MNI coordinates: 15, −49.5, 60). A and B represent the voxels that demonstrated the greatest sex difference in the control groups. **P*(FWE) < .05, ***P*(FWE) < .01. M = control boys, F = control girls, uMF = untreated male adolescents with GD, uFM = untreated female adolescents with GD.

girls. Likewise, suppressed FMs showed lower ROI activation than control boys. These results were not found in the untreated adolescents with GD.

Thus, the present results indicate that the observed atypical sexual differentiation of ToL related brain activation in the untreated individuals with GD was not (solely) due to pre-natal organizing effects. Interestingly, a recent review by Steensma et al. (2013) suggested that the period of adolescence seems to be crucial for the development of a non-normative gender identity. Pubertal hormones might be needed to activate the sex-atypical ToL related brain activations in adolescents with GD, whereas sex-atypical activations are no longer induced when pubertal hormones are suppressed by GnRHa, leading to sex-typical activation.

The GnRHa treated adolescents with GD even appeared to have *exaggerated* sex-typical activation of the ROIs. The suppressed FMs showed a significantly smaller activation of the right precuneus than Fs and the suppressed MFs showed a greater left RLPFC activation than Ms. Furthermore, the suppressed groups showed significant sex differences in every ROI, including ROIs that were not significantly different in the control groups. Interestingly, pre-pubertally administrated GnRHa was also found to modulate the development of cognitive functioning in sheep in a sex-specific manner (Wojniusz et al., 2011). Finally, additional factors might have played a role in the more prominent activation of the RLPFC in suppressed MFs. It is possible that this increase in left RLPFC activity reflects a greater effort of the suppressed MFs in performing the ToL task since they had the lowest IQ scores and made more errors than any other group.

In conclusion, our results suggest that there are no detrimental effects of GnRHa on EF. In addition, we have shed some light on another concern that has been raised among clinicians: whether GnRHa treatment would push adolescents with GD in the direction of their experienced gender. We found no evidence for this and if anything, we found that puberty suppression even seemed to make some aspects of brain functioning more in accordance with the natal sex.

Role of the funding source

The funding sources did not play a role in any component of this study.

Conflicts of interest

The authors report no biomedical financial interest or potential conflicts of interest.

Acknowledgements

The authors thank Thomas D. Steensma for his help during several stages of the study and Paul F.C. Groot for his help with the data analysis. This work was supported by an educational grant from Ferring BV, Hoofddorp, and by a VICI grant (453-08-003) from the Dutch Science Foundation (Nederlandse Organisatie voor Wetenschappelijk Onderzoek) to J. Bakker. J. Bakker is a senior research associate of the Belgian Fonds National de la Recherche Scientifique.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.psyneuen.2015.03.007.

References

Achenbach, T.M., Edelbrock, C.S., 1983. Manual for the Child Behavior Checklist and Revised Child Behavior Profile. University of Vermont, Department of Psychiatry, Burlington, VT.

Albert, D., Steinberg, L., 2011. Age differences in strategic planning as indexed by the Tower of London. Child Dev. 82, 1501–1517, http://dx.doi.org/10.1111/j.1467-8624.2011.01613.x.

American Psychiatric Association, 2000. Diagnostic and Statistical Manual of Mental Disorders, fourth ed. American Psychiatry Association, Washington, DC.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, fifth ed. American Psychiatric Publishing, Arlington, VA.

Asato, M.R., Sweeney, J.A., Luna, B., 2006. Cognitive processes in the development of TOL performance. Neuropsychologia 44, 2259–2269, http://dx.doi.org/10.1016/j.neuropsychologia.2006.05.010.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. Neuroimage 38, 95–113, http://dx.doi.org/10.1016/j.neuroimage.2007.07.007.

Baker, S.C., Rogers, R.D., Owen, A.M., Frith, C.D., Dolan, R.J., Frackowiak, R.S., Robbins, T.W., 1996. Neural systems engaged by planning: a PET study of the Tower of London task. Neuropsychologia 34, 515–526.

Berglund, H., Lindström, P., Dhejne-Helmy, C., Savic, I., 2008. Male-to-female transsexuals show sex-atypical hypothalamus activation when smelling odorous steroids. Cereb. Cortex 18, 1900–1908, http://dx.doi.org/10.1093/cercor/bhm216.

Boghi, A., Rasetti, R., Avidano, F., Manzone, C., Orsi, L., D'Agata, F., Caroppo, P., Bergui, M., Rocca, P., Pulvirenti, L., Bradac, G.B., Bogetto, F., Mutani, R., Mortara, P., 2006. The effect of gender on planning: an fMRI study using the Tower of London task. Neuroimage 33, 999–1010, http://dx.doi.org/10.1016/j.neuroimage.2006.07.022.

Bourgeois, J.P., Goldman-Rakic, P.S., Rakic, P., 1994. Synaptogenesis in the prefrontal cortex of rhesus monkeys. Cereb. Cortex 4, 78–96.

Casey, B.J., Trainor, R.J., Orendi, J.L., Schubert, A.B., Nystrom, L.E., Giedd, J.N., Castellanos, F.X., Haxby, J.V., Noll, D.C., Cohen, J.D., Forman, S.D., Dahl, R.E., Rapoport, J.L., 1997. A developmental functional MRI study of prefrontal activation during performance of a Go-No-Go task. J. Cognit. Neurosci. 9, 835–847, http://dx.doi.org/10.1162/jocn.1997.9.6.835.

Cohen-Kettenis, P.T., Gooren, L.J., 1999. Transsexualism: a review of etiology, diagnosis and treatment. J. Psychosom. Res. 46, 315–333.

Dagher, A., Owen, A.M., Boecker, H., Brooks, D.J., 1999. Mapping the network for planning: a correlational PET activation study with the Tower of London task. Brain 122 (Pt. 1), 1973–1987.

De Luca, C.R., Wood, S.J., Anderson, V., Buchanan, J.A., Proffitt, T.M., Mahony, K., Pantelis, C., 2003. Normative data from the CANTAB. I: development of executive function over the lifespan. J. Clin. Exp. Neuropsychol. 25, 242–254, http://dx.doi.org/10.1076/jcen.25.2.242.13639.

Giedd, J.N., 2008. The teen brain: insights from neuroimaging. J. Adolesc. Health 42, 335–343, http://dx.doi.org/10.1016/j.jadohealth.2008.01.007.

Gizewski, E.R., Krause, E., Schlamann, M., Happich, F., Ladd, M.E., Forsting, M., Senf, W., 2009. Specific cerebral activation due to visual erotic stimuli in male-to-female transsexuals compared with male and female controls: an fMRI study. J. Sex. Med. 6, 440–448, http://dx.doi.org/10.1111/j.1743-6109.2008.00981.x.

Gooren, L., Delemarre-van de Waal, H., 1996. The feasibility of endocrine interventions in juvenile transsexuals. J. Psychol. Hum. Sex. 8, 69–74, http://dx.doi.org/10.1300/J056v08n04_05.

Hoekzema, E., Schagen, S., Kreukels, B., Veltman, D.J., Cohen-Kettenis, P., de Waal, H.D., Bakker, J., 2015. Regional volumes and spatial volumetric distribution of grey matter in the gender dysphoric brain. Psychoneuroendocrinology 55, 59–71, http://dx.doi.org/10.1016/j.psyneuen.2015.01.016.

Houk, C.P., Lee, P.A., 2006. The diagnosis and care of transsexual children and adolescents: a pediatric endocrinologists' perspective. J. Pediatr. Endocrinol. Metab. 19, 103–109.

Huizinga, M., Dolan, C.V., van der Molen, M.W., 2006. Age-related change in executive function: developmental trends and a latent variable analysis. Neuropsychologia 44, 2017–2036, http://dx.doi.org/10.1016/j.neuropsychologia.2006.01.010.

Huttenlocher, P.R., 1979. Synaptic density in human frontal cortex – developmental changes and effects of aging. Brain Res. 163, 195–205.

Juraska, J.M., Sisk, C.L., DonCarlos, L.L., 2013. Sexual differentiation of the adolescent rodent brain: hormonal influences and developmental mechanisms. Horm. Behav. 64, 203–210, http://dx.doi.org/10.1016/j.yhbeh.2013.05.010.

Korte, A., Lehmkuhl, U., Goecker, D., Beier, K.M., Krude, H., Grüters-Kieslich, A., 2008. Gender identity disorders in childhood and adolescence: currently debated concepts and treatment strategies. Dtsch. Arztebl. Int. 105, 834–841, http://dx.doi.org/10.3238/arztebl.2008.0834.

Kranz, G.S., Hahn, A., Kaufmann, U., Küblböck, M., Hummer, A., Ganger, S., Seiger, R., Winkler, D., Swaab, D.F., Windischberger, C., Kasper, S., Lanzenberger, R., 2014. White matter microstructure in transsexuals and controls investigated by diffusion tensor imaging. J. Neurosci. 34, 15466–15475, http://dx.doi.org/10.1523/JNEUROSCI.2488-14.2014.

Kreukels, B.P.C., Cohen-Kettenis, P.T., 2011. Puberty suppression in gender identity disorder: the Amsterdam experience. Nat. Rev. Endocrinol. 7, 466–472, http://dx.doi.org/10.1038/nrendo.2011.78.

Lazeron, R.H., Rombouts, S.A., Machielsen, W.C., Scheltens, P., Witter, M.P., Uylings, H.B., Barkhof, F., 2000. Visualizing brain activation during planning: the tower of London test adapted for functional MR imaging. AJNR Am. J. Neuroradiol. 21, 1407–1414.

Lemieux, L., Salek-Haddadi, A., Lund, T.E., Laufs, H., Carmichael, D., 2007. Modelling large motion events in fMRI studies of patients with epilepsy. Magn. Reson. Imaging 25, 894–901, http://dx.doi.org/10.1016/j.mri.2007.03.009.

Luders, E., Sánchez, F.J., Gaser, C., Toga, A.W., Narr, K.L., Hamilton, L.S., Vilain, E., 2009. Regional gray matter variation in male-to-female transsexualism. Neuroimage 46, 904–907, http://dx.doi.org/10.1016/j.neuroimage.2009.03.048.

Luders, E., Sánchez, F.J., Tosun, D., Shattuck, D.W., Gaser, C., Vilain, E., Toga, A.W., 2012. Increased cortical thickness in male-to-female transsexualism. J. Behav. Brain Sci. 2, 357–362, http://dx.doi.org/10.4236/jbbs.2012.23040.

Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19, 1233–1239.

Marshall, W.A., Tanner, J.M., 1969. Variations in pattern of pubertal changes in girls. Arch. Dis. Child. 44, 291–303.

Marshall, W.A., Tanner, J.M., 1970. Variations in the pattern of pubertal changes in boys. Arch. Dis. Child. 45, 13–23.

Newman, S.D., Carpenter, P.A., Varma, S., Just, M.A., 2003. Frontal and parietal participation in problem solving in the

Tower of London: fMRI and computational modeling of planning and high-level perception. Neuropsychologia 41, 1668–1682, http://dx.doi.org/10.1016/S0028-3932(03)00091-5.

Nielsen, F.A., Hansen, L.K., 2002. Automatic anatomical labeling of Talairach coordinates and generation of volumes of interest via the BrainMap database. Neuroimage, 16.

Ogawa, S., Lee, T.M., Kay, A.R., Tank, D.W., 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc. Natl. Acad. Sci. U. S. A. 87, 9868–9872.

Owen, A.M., Doyon, J., Petrides, M., Evans, A.C., 1996. Planning and spatial working memory: a positron emission tomography study in humans. Eur. J. Neurosci. 8, 353–364.

Rametti, G., Carrillo, B., Gómez-Gil, E., Junque, C., Segovia, S., Gomez, Á., Guillamon, A., 2011a. White matter microstructure in female to male transsexuals before cross-sex hormonal treatment. A diffusion tensor imaging study. J. Psychiatr. Res. 45, 199–204, http://dx.doi.org/10.1016/j.jpsychires.2010.05.006.

Rametti, G., Carrillo, B., Gómez-Gil, E., Junque, C., Zubiarre-Elorza, L., Segovia, S., Gomez, Á., Guillamon, A., 2011b. The microstructure of white matter in male to female transsexuals before cross-sex hormonal treatment. A DTI study. J. Psychiatr. Res. 45, 949–954, http://dx.doi.org/10.1016/j.jpsychires.2010.11.007.

Romeo, R.D., 2003. Puberty: a period of both organizational and activational effects of steroid hormones on neurobehavioural development. J. Neuroendocrinol. 15, 1185–1192.

Rowe, J.B., Owen, A.M., Johnsrude, I.S., Passingham, R.E., 2001. Imaging the mental components of a planning task. Neuropsychologia 39, 315–327.

Savic, I., Arver, S., 2011. Sex dimorphism of the brain in male-to-female transsexuals. Cereb. Cortex 21, 2525–2533, http://dx.doi.org/10.1093/cercor/bhr032.

Schall, U., Johnston, P., Lagopoulos, J., Jüptner, M., Jentzen, W., Thienel, R., Dittmann-Balçar, A., Bender, S., Ward, P.B., 2003. Functional brain maps of Tower of London performance: a positron emission tomography and functional magnetic resonance imaging study. Neuroimage 20, 1154–1161, http://dx.doi.org/10.1016/S1053-8119(03)00338-0.

Schöning, S., Engelien, A., Bauer, C., Kugel, H., Kersting, A., Roestel, C., Zwitserlood, P., Pyka, M., Dannlowski, U., Lehmann, W., Heindel, W., Arolt, V., Konrad, C., 2010. Neuroimaging differences in spatial cognition between men and male-to-female transsexuals before and during hormone therapy. J. Sex. Med. 7, 1858–1867, http://dx.doi.org/10.1111/j.1743-6109.2009.01484.x.

Simon, L., Kozák, L.R., Simon, V., Czobor, P., Unoka, Z., Szabó, Á., Csukly, G., 2013. Regional grey matter structure differences between transsexuals and healthy controls – a voxel based morphometry study. PLOS ONE 8, e83947, http://dx.doi.org/10.1371/journal.pone.0083947.

Sisk, C.L., Zehr, J.L., 2005. Pubertal hormones organize the adolescent brain and behavior. Front. Neuroendocrinol. 26, 163–174, http://dx.doi.org/10.1016/j.yfrne.2005.10.003.

Soleman, R.S., Schagen, S.E.E., Veltman, D.J., Kreukels, B.P.C., Cohen-Kettenis, P.T., Lambalk, C.B., Wouters, F., Delemarre-van de Waal, H.A., 2013. Sex differences in verbal fluency during adolescence: a functional magnetic resonance imaging study in gender dysphoric and control boys and girls. J. Sex. Med. 10, 1969–1977, http://dx.doi.org/10.1111/jsm.12083.

Spriggs, M.P., 2004. Ethics and the proposed treatment for a 13-year-old with atypical gender identity. Med. J. Aust. 181, 319–321.

Steensma, T.D., Kreukels, B.P.C., de Vries, A.L.C., Cohen-Kettenis, P.T., 2013. Gender identity development in adolescence. Horm. Behav. 64, 288–297, http://dx.doi.org/10.1016/j.yhbeh.2013.02.020.

Swaab, D.F., 2004. Sexual differentiation of the human brain: relevance for gender identity, transsexualism and sexual orientation. Gynecol. Endocrinol. 19, 301–312.

Van den Heuvel, O.A., Groenewegen, H.J., Barkhof, F., Lazeron, R.H.C., van Dyck, R., Veltman, D.J., 2003. Frontostriatal system in planning complexity: a parametric functional magnetic resonance version of Tower of London task. Neuroimage 18, 367–374.

Van Goozen, S.H.M., Slabbekoorn, D., Gooren, L.J.G., Sanders, G., Cohen-Kettenis, P.T., 2002. Organizing and activating effects of sex hormones in homosexual transsexuals. Behav. Neurosci. 116, 982–988.

Van Strien, J.W., 1992. Classificatie van links- en rechtshandige proefpersonen. Ned. Tijdschr. Psychol. 47, 88–92.

Verhulst, F.C., van der Ende, J., Koot, H.M., 1996. Handleiding voor de CBCL/4-18 (Manual for the CBCL/4-18). Erasmus University, Department of Child and Adolescent Psychiatry, Sophia Children's Hospital, Rotterdam, The Netherlands.

Viner, R., 2005. Sex on the brain: dilemmas in the endocrine management of children and adolescents with gender identity disorder. Arch. Dis. Child. 90 (Sup.2), A77–A81.

Wagner, G., Koch, K., Reichenbach, J.R., Sauer, H., Schlösser, R.G.M., 2006. The special involvement of the rostrolateral prefrontal cortex in planning abilities: an event-related fMRI study with the Tower of London paradigm. Neuropsychologia 44, 2337–2347, http://dx.doi.org/10.1016/j.neuropsychologia.2006.05.014.

Wechsler, D., 1991. Wechsler Intelligence Scale for Children, third ed. The Psychological Corporation, San Antonio.

Wechsler, D., 1997. Wechsler Adult Intelligence Scale, third ed. The Psychological Corporation, San Antonio.

Wojniusz, S., Vögele, C., Ropstad, E., Evans, N., Robinson, J., Sütterlin, S., Erhard, H.W., Solbakk, A.-K., Endestad, T., Olberg, D.E., Haraldsen, I.R.H., 2011. Prepubertal gonadotropin-releasing hormone analog leads to exaggerated behavioral and emotional sex differences in sheep. Horm. Behav. 59, 22–27, http://dx.doi.org/10.1016/j.yhbeh.2010.09.010.

Woo, T.U., Pucak, M.L., Kye, C.H., Matus, C.V., Lewis, D.A., 1997. Peripubertal refinement of the intrinsic and associational circuitry in monkey prefrontal cortex. Neuroscience 80, 1149–1158.

Zubiaurre-Elorza, L., Junque, C., Gómez-Gil, E., Segovia, S., Carrillo, B., Rametti, G., Guillamon, A., 2013. Cortical thickness in untreated transsexuals. Cereb. Cortex 23, 2855–2862, http://dx.doi.org/10.1093/cercor/bhs267.