AUS DER KLINIK UND POLIKLINIK FÜR PSYCHIATRIE UND PSYCHOTHERAPIE
Geschäftsführender Direktor: Univ.-Prof. Dr. Tilo Kircher
DES FACHBEREICHS MEDIZIN DER PHILIPPS-UNIVERSITÄT MARBURG

Investigating the Effects of Sex Hormones on the Female Brain – Necessary Prerequisites and a First Insight on the Influences on Gray Matter Volume

INAUGURAL-DISSERTATION
zur Erlangung des Doktorgrades der Naturwissenschaften
dem Fachbereich Medizin der Philipps-Universität Marburg
vorgelegt von

Verena Schuster aus Nauroth/Westerwald
Marburg, 2020

Angenommen vom Fachbereich Medizin der Philipps-Universität Marburg
am: 26.11.2020

Gedruckt mit Genehmigung des Fachbereichs.

Dekanin: Univ.-Prof. Dr. Denise Hilfiker-Kleiner

Referent: Univ.-Prof. Dr. Andreas Jansen

Korreferentin: Dr. Estrid Jakobsen

for Martin

"There is a crack, a crack in everything
that's how the light gets in."

Leonard Cohen

CONTENTS

Abbreviations

ACC	Anterior Cingulate Cortex
ADHD	Attention Deficit Hyperactivity Disorder
AMAP	Adaptive maximum a posteriori
BOLD	Blood Oxygen Level Dependent
COC	Combined Oral Contraceptives
CNS	Central Nervous System
CSF	Cerebrospinal Fluid
DBM	Diffusion-Based Morphometry
Dis	“Dots-in-space”
DTI	Diffusion Tensor Imaging
EPI	Echo Planar Imaging (EPI)
ER	Estrogen Receptor
ERα	Estrogen Receptor alpha
ERβ	Estrogen Receptor beta
FCA	Functional Cerebral Asymmetry
FFA	Fusiform Face Area
Fig.	Figure
fMRI	Functional Magnetic Resonance Imaging
FoV	Field of View
FSH	Follicle Stimulating Hormone
fTCD	Functional Transcranial Doppler Sonography
FWE	Family Wise Error
FWHM	Full Width Half Maximum
GABA	Gamma-Aminobutyric Acid

GLM	General Linear Model
GM	Gray Matter
GnRH	Gonadotropin-Releasing Hormone
ICC	Intraclass Correlation Coefficient
LH	Luteinizing Hormone / Left Hemisphere
LI	Lateralization Index
LT	Landmark Task
MCA	Middle Cerebral Arteries
MFG	Middle Frontal Gyrus
MLM	Multi-level linear modeling approach
MNI	Montreal Neurological Institute
MRI	Magnetic Resonance Imaging
MR	Mental Rotation
NC	Natural Cycle
NMDA	*N*-methyl-D-aspartate receptor
OC	Oral Contraceptives
OLS	Ordinary least-square regression
PBT	Projection-Based Thickness
PPA	Parahippocampal Place Area
RBM	ROI-Based Morphometry
RH	Right Hemisphere
ROI	Region Of Interest
SBM	Surface-Based morphometry
SHBG	Sex hormone-binding globulin
STG	Superior Temporal Gyrus
TR	Repetition Time
TE	Echo Time

TIV	Total Intracranial Volume
TPM	Tissue Probability Map
VBM	Voxel-Based Morphometry
VTE	venous thromboembolic event
WM	White Matter

List of Figures

List of Tables

1

Introduction

Within this very first part of this doctoral thesis, the aim is to provide an introduction to the influence of sex hormones on the human brain. Accordingly, at first, a general overview of male and female sex hormones and the consequence of taking "the pill" is given. Subsequently, the effect of female sex hormones, in particular, on brain function and brain structure is highlighted. At the end of this chapter, previous studies' limitations investigating this broad and complex topic are summarized.

1 Introduction

1.1 What makes the little Difference: Female and Male Sex Hormones in the Human Body

More than 100 years ago, in June 1905, Ernest Starling, a brilliant experimentalist and professor of physiology at the University College London, UK, introduced a term, that revolutionized and catalyzed science: *Hormone*. He defined the word deriving from the Greek word *hormân* meaning to set in motion, excite, stimulate, as a chemical messenger, which speeds from cell to cell along with the bloodstream, coordinating the activities and growth of different parts of the body (Starling & Royal College of Physicians of London, 1905; Tata, 2005). Nowadays, about 50 different hormones have been identified in the human body, all of them being unique concerning their origin, their target and their mediated effects. Nevertheless, the concentration of different hormones differs between humans, especially these of gonadal hormones between men and women.

At present, however, differences between men and women are a highly sensitive issue and of great interest to contemporary gender research. Scientists pursue the goal of determining factors, that potentially contribute to differences between these sexes and have already identified a broad range of causes in social issues e.g., access to education, resulting in different behavioral patterns (Hyde, 2014; Wood & Eagly, 2002). However, social aspects alone cannot contribute to and explain all observed differences between men and women, particularly biological differences (Confer et al., 2010).

Here, life sciences come into play and illuminate many components that contribute to sex differences:

At the very beginning of life, it is the fertilization of a female egg by a male sperm which constitutes not only the beginning of a new life but also the determination of the sex. Whereas the ovum always provides an X chromosome, it is the sperm that either contains an X or a Y chromosome and therefore regulating the resulting sex. Each individuum's cell contains 23 chromosomes pairs, 22 pairs are autosomes, and one pair of sex chromosomes.

If the pair of sex chromosomes have an X and a Y chromosome (XY), this results in a male mammal, whereas female mammals typically have two X chromosomes (XX). During the next seven weeks, female and male fetuses develop equally, before the sexual differentiation then is initiated.

The key players, therefore, are gonadal hormones: If a Y chromosome is present, its sex-determining region (SRY) gene produces SRY protein, which in turn binds to DNA

and triggers the development of the gonad into testes. These then begin to produce and secret anti-müllerian hormone, testosterone and dihydrotestosterone guiding the development of the male phenotype.
In the absence of a Y chromosome, the development of a female phenotype processes (MacLaughlin & Donahoe, 2004). In summary, just seven weeks after conception, a tiny human being with a size of only about 20 mm firstly gets in touch with sex hormones, which presence and effects will accompany it at any time after that for the rest of its life.

1.1.1 A brief Overview of Male Sex Hormones

Focusing on men, the most prominent male sex hormone, summarized as androgens (*Greek: man-maker*), is beyond doubt testosterone.
At the very beginning, it is the hypothalamus, producing gonadotropin-releasing hormone (GnRH), fulfilling its function in the adjacent pituitary gland, which in turn secrets follicle-stimulating hormone (FSH) and luteinizing hormone (LH). These hormones targets are the testicle, with LH triggering testosterone production in the *Leydig cells* and FSH, accompanying testosterone, which stimulates sperm production throughout adult life. However, testosterone offers other diverse functions: it is responsible for the growth of the genitals at puberty, the typically increased muscle mass in males, the distinctive more resonant voice and stronger bones but also contributes to specific male behavior traits, e.g. aggressiveness, sex drive, and sexual performance (Ngun et al., 2011). As mentioned above, testosterone production begins very early in life around the seventh week of embryonic development. Levels are low during childhood but rise during puberty to fulfill their typical functions to a maximum at about 17 years, remaining high for the following two to three decades or, in some men, during life, before they slowly decline, averaging about 1 % per year (Harvard Health, 2008). Summarized, most healthy men experience a relatively stable level of sex hormones after puberty for the rest of their lives (Alonso & Rosenfield, 2002; Handelsman et al., 2018; Ober et al., 2008) (see Figure 1).

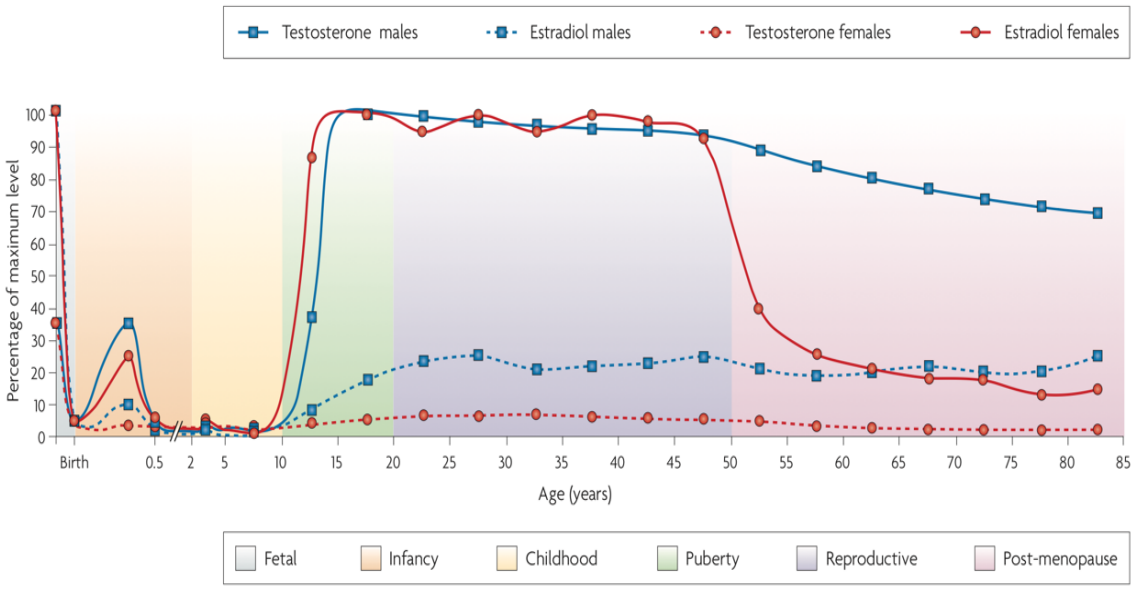


Figure 1 *Approximate mean sex-steroid levels in plasma in males and females.*

Variation in steroid levels is shown as a percentage of the maximum mean testosterone and the maximum mean estradiol across the life stages shown. The figure does not show diurnal, cyclic (female) or possible seasonal fluctuations. Female estradiol levels refer to the mean for the mid-follicular phase of the menstrual cycle; estradiol production transiently increases about fivefold during the pre-ovulatory and luteal phases of the menstrual cycle. Note the drop in the levels of all sex steroids at birth and the transient 'minipuberty' in early infancy. Free testosterone in men falls more with ageing (to approximately 50% of the maximum in 80-year-old men) than the total testosterone, which is shown here (Ober et al., 2008).

This figure is reprinted with permission from Springer Nature for reuse in thesis / dissertation.

1.1.2 A brief Overview of Female Sex Hormones

Quite the opposite holds true for women:

With the advent of puberty, sex hormone levels increase immensely, and mean values are comparatively high for the following more than 30 years, known as reproductive years (Mihm et al. 2010) before they relatively abruptly decline during the transition to reproductive senescence, commonly known as the menopause (Bale & Epperson, 2015) (see Figure 1). However, women do not exclusively experience substantial alterations of sex hormones during their lifetime, but, with a focus on the reproductive years, in an around monthly manner, known as the menstrual cycle. The menstrual cycle is tightly

governed by a complex interplay and feedback mechanisms of endocrine, paracrine and autocrine factors integrating the action of the hypothalamus, pituitary, and ovary and endometrium, with the aim to provide a fertilizable egg and its implantation. The average menstrual cycle (see Figure 2) (interval from the first day of menses to begin of next menses) is 28 days in duration and can simply be divided into two stages: the follicular phase, starting on the first day of menstrual bleeding and lasting to ovulation around cycle day 14, and the luteal phase, ranging from ovulation to the first day of menstrual bleeding around cycle day 28 (Mihm et al. 2010).

However, menstrual cycle length and hence the different phases' length are highly variable, ranging from 25 to 34 days and changes from menarche to menopause, impeding the correct prediction of the cycle phase.
The follicular phase, ranging from the first day of menses until ovulation, is characterized by measurable lower basal body temperature and the rise of follicle-stimulating hormone (FSH). The higher-level control is the hypothalamus, which in turn, secrets gonadotropin-releasing hormone (GnRH) pulses every 1-1.5h, stimulating the pituitary gland to produce luteinizing hormone (LH) and FSH (Barbieri, 2014), which act on different cells of the ovarian follicle. LH stimulates theca cells to produce androstenedione, whereas FSH triggers the synthesis of aromatase, which in turn catalyzes the conversion of androstenedione to estradiol. When a critical concentration of estradiol is passed, it causes positive feedback in the hypothalamus, resulting in an increase in GnRH secretion and an LH peak. This LH peak then initiates the ovulation and the beginning of the second half of the menstrual cycle, the luteal phase, named after the corpus luteum in which the follicle is transformed. The corpus luteum starts producing progesterone (P_4), which prepares the endometrium for the nidation of the fertilized egg. If the egg is not fertilized, the corpus luteum stops secreting progesterone and decays into the corpus albicans. With the decrease of progesterone, LH production inhibition is eliminated, laying the foundation for the next cycle. The functional endometrium necrotizes and the menstrual bleeding and consequently, the next menstrual cycle begins.

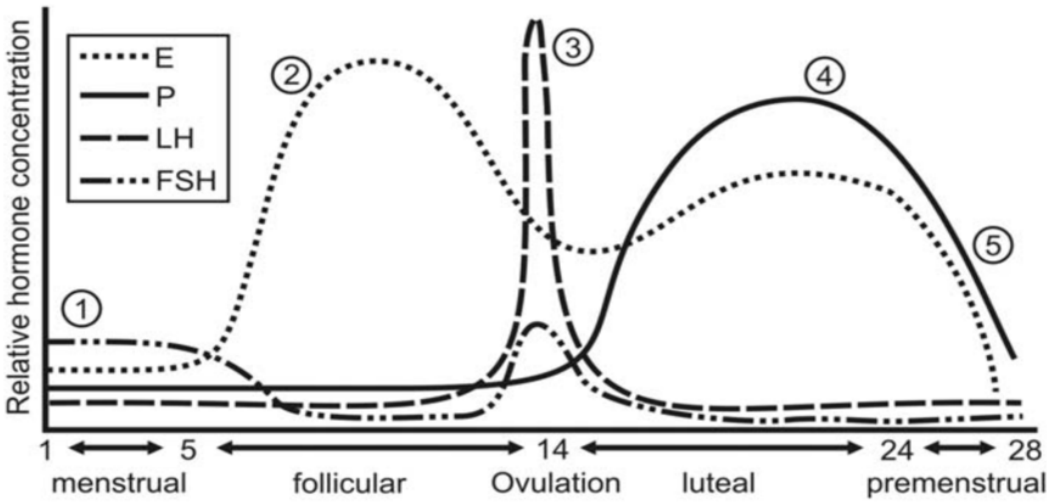


Figure 2 *The Menstrual Cycle.*

Schematic overview of fluctuations in sex hormones and gonadotropin levels during a 28-day menstrual cycle. During the menstrual phase (1) (cycle days 1-5), concentration of E and P is lowest. During the follicular phase (2) (cycle days 6-13), E levels increase with a maximum one day before ovulation (3). Ovulation in initiated due to an increase of LH on day 14. Afterwards, E level drops slightly. During this luteal phase, E and P are secreted by the luteinized cells. About 7 to 8 days postovulatory, E level approaches its second maximum together with P. P level reaches its peak at around cycle day 22 (midluteal phase, 4). Levels of E and P fall rapidly between cycle day 24 and 28 (premenstrual phase, 5) and a new cycle begins (Weis & Hausmann, 2010).

This figure is reprinted with permission from SAGE publishing at no cost for use of content in a Master's Thesis and/or Doctoral Dissertation.

1.1.3 The Effect of the "Pill"

Additionally, women experience furthermore hormonal conditions: for example, if the egg is fertilized and pregnancy occurs, or during the breastfeeding period. Whereas the latter states are a relatively short period of time during women's reproductive years, another influencing factor on the hormonal state is present in a long time period, sometimes longer than a decade: the use of hormonal contraceptives. Worldwide, more than 100 million people use oral contraceptives (OCs) (Petitti 2003). In Germany, alone more than half of the women between 18 and 49 years use OCs as a method of choice for contraception (Heßling & Bundeszentrale für Gesundheitliche Aufklärung, 2011).

OCs look back to a rapid development: The elucidation of the ovary's anatomy and physiology occurred around the turn of the century and the identification of the hormone progesterone as a key player for pregnancy and its isolation and characterization of its chemical structure was realized in the 1930s. In 1944, a group of German scientists reported that a daily doses of 20 mg progesterone is sufficient to suppress the ovulation in women (Bickenbach, 1944; Ludwig, 2011), and in the early 1950s, the chemist Carl Djerassi and colleagues synthesized norethindrone, a chemically modified gestagen, which survived the entero-hepatical passage, which is intended to be the first oral contraceptive (Djerassi, 1992, 2006).

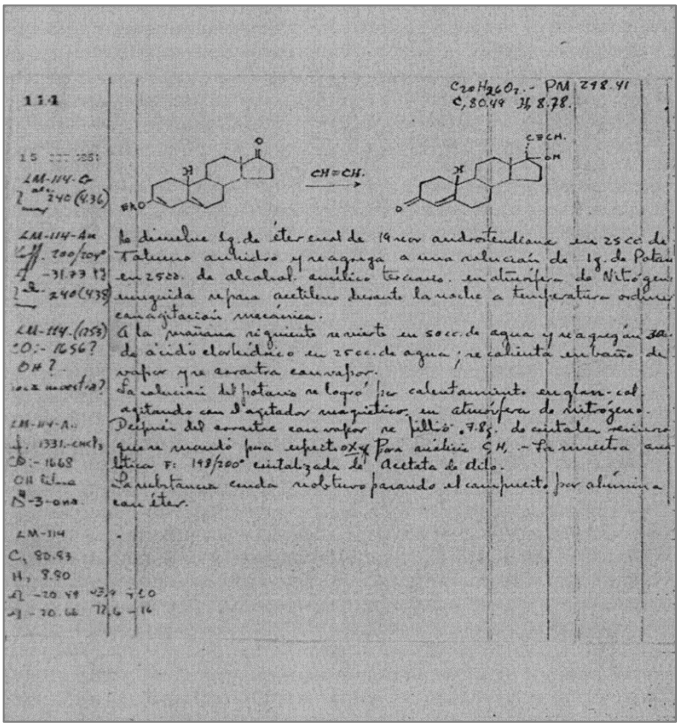


Figure 3 *The chemical birthday of the Pill.*

Page 114 (October 15, 1951, the chemical birthday of the pill) from laboratory notebook of Luis Miramontes, who was supervised by Carl Djerassi. This page shows the entire chemical and spectroscopic characterization of norethindrone *(Djerassi, 1992)*.

After a large American scale study in the late 1950s, the "Food and Drug Administration" legalized the first contraceptive named Enovid in the USA in 1960. Just one year later, the pill "Anovlar" was approved and available in Europe. Since then, OCs heavily influences our society as summarized by Liao and colleague:

> *'Throughout history of medicine, thousands of drugs have been developed, but only one has been influential enough to earn the title of simply, the pill.'* (Liao & Dollin, 2012)

Talking about “the pill”, one must distinguish between the progesterone-only pill and combined oral contraceptives (COCs). Whereas the first one plays a minor role, COCs are regularly used by a third of Germany’s 20 million women of reproductive age (Wiegratz & Thaler, 2011). They consist of a combination of estrogen, a low dose of ethinylestradiol is used in nearly all preparations, and progestogen, which type varies in efficacy and thus concentration. COCs are divided into three different generations, based on their containing progesterone. However, first-generation progestins are no longer available due to their increasing risk for developing breast cancer (Dalton, 1981). The contraceptive effect is mostly based on the suppression of FSH and LH secretion; following this leads to the arrest of follicular maturation, the absence of the LH surge, and thus the suppression of the ovulation. Additionally, progestogens act on cervical mucus, tubular function and the endometrium, impeding implantation. The majority of COCs are taken for 21 consecutive days immediately followed by a seven-day pill-free interval, during which the withdrawal bleeding takes place.

Although COCs are the contraception of choice and widely used, they are recently getting increased criticism regarding undesirable side effects.

Most commonly known are cardiovascular interactions with a three- to five-fold higher thrombosis risk, particularly venous thromboembolic events (VTEs), in women, using COCs (Trenor et al., 2011), making them one of the most common risk factors of VTEs in young women. Whereas the composition of COCs has changed over decades, characterizes as different pill generations, with an overall reduction in estrogen levels and incorporation of new progestins, a meta-study revealed a still 4.3-fold increased risk of VTE of the use of second- and third-generation (Baratloo et al., 2014).

An additional side-effect, rapidly gaining attention over the last years, is the increased potential risk of developing depression (Skovlund et al., 2016), and decreasing general well-being due to the intake of COCs (Zethraeus et al., 2017). Toffoletto and colleagues investigated the influence of endogenous and exogenous sex hormones in cortical and subcortical regions involved in emotional and cognitive processing in a systemic review (Toffoletto et al., 2014), rising the suspicion that sex hormones do not only act in the periphery, e.g. gonads, but also in the brain.

1.2 The Influence of Female Sex Hormones on Brain Function

It has become increasingly evident that the functions of sex hormones extend well beyond reproduction (Brinton et al., 2008). Already in 1964, Young and colleagues hypothesized in an article published in Science the role of sex hormones in the central nervous system:

'*Few biochemists have been attracted to the problem, but it is they who must clarify the mechanisms of hormonal action in organizing the tissues of the central nervous system during development and in bringing behavior to expression in the adult. They may be helped in such a search by the circumstance that cellular elements in the genital tract, which differentiate and are activated under the influence of these same hormones, are at present more accessible for histophysiological study than those in tissues of the central nervous system'* (Young et al., 1964)

Nowadays, it is already known that sex hormones generally determine the biological phenotype but also psychological traits and vulnerability to many psychiatric and neurological diseases (Cosgrove et al., 2007).
For instance, attention deficit hyperactivity disorder (ADHD), Autism and Dyslexia are only a few examples of diseases men are more prone to than women. In contrast, women are more vulnerable to develop, for instance, post-traumatic stress disorder, underlining the systemic effect and action of sex hormones beyond the reproductive system.

To better understand the effect of sex hormones on the brain, one must understand their underlying modes of action, beginning with the binding to their specific receptors. Focusing on estrogens, one has to distinguish between four naturally biosynthesized estrogens in women. Estrone (E1) is predominant during menopauses and derives the weakest action, whereas estradiol (E2) shows the most potent effects and is present during the menstrual cycle, before pregnancy and menopause. Estriol (E3) is the predominant estrogen during pregnancy, and Estetrol (E4) is also only present during pregnancy (Raghava et al., 2017). Estrogens produce their effects after binding to one of two to date known distinct intracellular estrogen receptors (ER): estrogen receptor alpha (ERα) and estrogen receptor beta (ERβ). ERs are part of steroid receptors and act as ligand activation transcription factors, resulting in the modulation of gene transcription (Tsai & O'Malley, 1994). ERα and ERβ are vastly distributed in the brain, whereas both receptors have overlapping expression patterns in most human

brain parts: Both receptors are present in the cerebral cortex, but with different concentration patterns in distinct cortical layers. Whereas ERα expression dominates in the hypothalamus and the amygdala with only low accumulation of ERβ, the opposite applies for the entorhinal cortex, thalamus and hippocampal regions, which are one of the most abundant ERβ expressing areas (Österlund & Hurd, 2001). This distinct expression of both receptors assigning both receptor types particular roles, with ERα being involved in the modulation of neuronal populations with autonomic and reproductive neuroendocrine functions, emotional processing, affective and motivational behaviors, and ERβ modulating cognition, non-emotional memory and motor functions. Additionally, estrogens, particularly estradiol, provide neuroprotective effects in the central nervous system due to attenuation of neuroinflammation and neurodegeneration (Raghava et al., 2017).

Another sex hormone associated with neuroprotective effects besides its well-studied role in regulating reproduction and female sexual behaviors is progesterone (P_4). Additionally, in its role as a neurosteroid, it is involved in neuroplasticity (Baudry et al., 2013), neurogenesis (Bali et al., 2012) and neuroinflammation (Giatti et al., 2012). Because of its diverse effects, it is not surprisingly that progesterone receptors are broadly expressed throughout the brain.

1.2.1 Lateralization of cognitive functions and the hypothesis of progesterone mediated interhemispheric decoupling

The role of progesterone in the brain has already been discussed almost 20 years ago with regard to cognitive brain functions. Hausmann and Güntürkün described the effect of progesterone on brain lateralization of cognitive functions and postulated the *progesterone mediated interhemispheric decoupling hypothesis* (Hausmann & Güntürkün, 2000). To better understand the importance of this rather bulky and abstract named theory, it is necessary to take one large step back to France in the mid-19th century. Here, the scientist Paul Pierre Broca was thrilled by the idea that one distinct brain region located within the third convolution of the left frontal lobe, next to the lateral sulcus, is responsible for speech production (Broca, 1861b, 1861a, 1861c). In 1861 he performed an autopsy on Louis Victor Leborgne's brain. The patient was simply known as "Tan" due to his inability to speak any other words than "tan" clearly. Broca hoped to find a physiological explanation for Tan's disability and found a lesion in the frontal lobe of the left hemisphere. In eleven further aphasic patients, Broca performed brain

autopsies and repeatedly detected brain lesions within the mentioned area in each of his patients. In 1865 he published his results (Broca, 1865) and has since then been revolutionizing not only the understanding of language organization in the brain but also hemispheric specialization of cognitive functions in general.

The general concept of functional asymmetry and lateralization of cognitive function to either one of the two hemispheres is a basic principle of the organization of the human brain. Different cognitive functions e.g., language, spatial attention, and memory are distributed differently across the brain's two hemispheres, known as hemispheric specialization. With the advent of the development of modern brain imaging techniques, in particular functional transcranial Doppler sonography (fTCD) and functional magnetic resonance imaging (fMRI), their study has been made more widely feasible and enabled non-invasive studies addressing the issue of the hemispheric specialization of cognitive functions in large cohorts of healthy participants as well as patients (Stephan et al., 2007). For measuring brain function, the nowadays method of choice is fMRI, as it is non-invasive and without any known side effects.

Thanks to these improvements of methodology, researchers were able to show that various cognitive functions are mainly located in one hemisphere; for example, in most individuals language functions are lateralized to the left hemisphere (Frost, 1999). In contrast, visuospatial functions are processed and executed by the right hemisphere (Fink et al., 2000). However, these studies also highlighted that for all these processes, the degree of lateralization is variable not only between subjects but also within subjects. For example, language, as a typically left-hemispheric localized cognitive function, an atypical right-hemispheric or bilateral form of language lateralization has been observed in up to 10% of the human population (Frost, 1999; Jansen, et al., 2005; Knecht, Dräger, et al., 2000; Pujol et al., 1999). Visuospatial attention functions, a predominantly right-hemispheric cognitive function, is again subject to marked variability across subjects (Fink et al., 2000; Kinsbourne & Bruce, 1987; Mesulam, 1981).

Besides the hemispheric specialization of various cognitive tasks, these functions, especially language and visuospatial processes, have also shown to be lateralized to a sex-specific manner. These studies found more pronounced functional cerebral asymmetries (FCAs) in men compared to women (Hausmann & Güntürkün, 1999; Shaywitz et al., 1995) and consequently highlight the role of sex hormones for inter- and intraindividual variations in FCAs. FCAs in general are a simple model to investigate functional connectivity in the brain, especially between left and right cerebral hemispheres, referring to the relative differences in many neural functions and cognitive processes (Hausmann, 2017). In men, FCAs tend to be stable and more robust, whereas they greatly vary in women with an overall more symmetrical or bilateral pattern

(Hausmann & Güntürkün, 1999; Hausmann & Güntürkün, 2000). However, this is not the case over women's entire lifespan.

After menopause, and during menses, FCAs are comparable to those of men, highlighting the role of gonadal hormones, especially progesterone, in modulating lateralization patterns, leading us back to the progesterone-mediated interhemispheric decoupling hypothesis. Hausmann and Güntürkün examined the effect of gonadal hormones on FCAs in three distinct groups:

a) *young cycling women were tested twice, once during a low state of steroid hormones (menses) and once during a high state of steroid hormones (midluteal phase);*
b) *young men and*
c) *postmenopausal women.*

All participants performed a prototypical left (lexical decision) and two prototypical right-hemispheric visual half-field tasks (figural comparison and face discrimination). They found significant interactions between the menstrual cycle phase and visual half-field, indicating that sex hormones modulation of lateralization patterns acts independently of task or hemisphere. Additionally, they found remarkably stable lateralization patterns for all functions in postmenopausal women, identical to those of men. The authors conclude that FCAs seem to be hormonally modulated by a global mechanism: In general, both hemispheres work as partially independent systems, with each processing stimuli simultaneously. These simultaneous and independent processing requires control mechanisms to coordinate and control the outputs from both hemispheres. One coordinating key mechanism is the interhemispheric inhibition across the corpus callosum, which determines FCAs (Chiarello & Maxfield, 1996). The corpus callosum consists in large parts of excitatory glutamatergic pyramidal neurons' fibers and only to a small amount of inhibitory gamma-Aminobutyric acid-ergig (GABAergig) fibers (Hughes & Peters, 1992).

Although the amount might be smaller, the longer-lasting effects of callosal activation are inhibitory and can be induced pharmacologically, resulting in an attenuation of non-*N*-methyl-D-aspartate receptor (non-NMDA) glutamate receptors and a reduction of short excitatory and longer-lasting inhibitory influence (Kawaguchi, 1992).

The same effects are revealed by physiological doses of progesterone (Smith et al., 1987). Thus, high progesterone levels, as they occur naturally during the luteal phase of the menstrual cycle, can reduce transcallosal inhibition and thus lead to a functional decoupling of the hemispheres and hence to a temporary reduction of FCAs (Fig. 4). In summary, this leads to an overall functional hemispheric decoupling and, thus, to a temporal reduction in functional asymmetry. The authors conclude that steroid

fluctuations during the menstrual cycle modify cerebral asymmetries to a certain extent, with the decrease of sex hormones (during menses and after menopause), stabilizing cerebral asymmetries and an increase (during midluteal phase) leading to reduced lateralization.

Even more interesting, during low hormonal phases (menstruation), female asymmetries are similar to that of men, and to women after menopause.

Whereas the described hypothesis is almost 20 years old, it has received some empirical support from different studies with various designs and techniques (Hausmann, 2005; Weis et al., 2008; Weis & Hausmann, 2010).

Since then, studying women during different cycle phases, as the menstrual cycle serves with most dramatic hormonal changes within short time periods, has become a major tool to investigate the influence of sex hormones on cognitive behavior and its underlying functional brain organization.

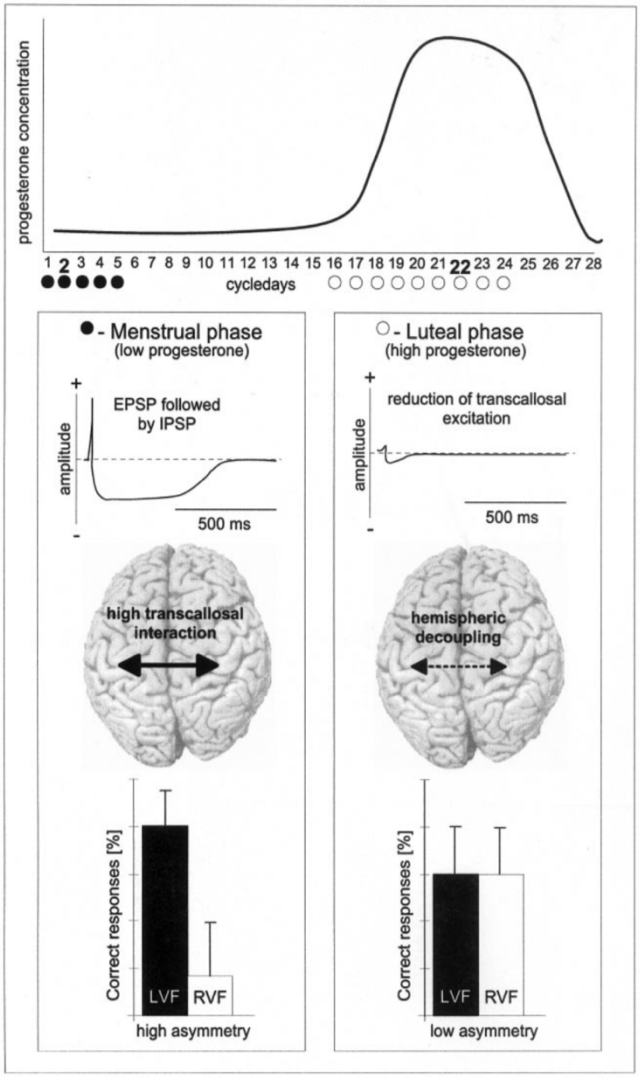


Figure 4 *The hypothesis of progesterone-mediated interhemispheric decoupling.*

It is assumed that a cycle-dependent increase in progesterone concentration during the luteal phase decreases glutamatergic non-NMDA and increases GABA receptor activation. These effects lead, via a decrease of transcallosal neuronal activation, to hemispheric decoupling, resulting in reduced FCAs (Hausmann & Güntürkün, 2000).

This figure is reprinted with permission from Elsevier, Neuropsychologia for reuse in thesis / dissertation.

1.2.2 The "Pill" and its Influence on Brain Functions

To make it even more complex and complicated, as described above, women do not only experience dramatic hormonal changes during the monthly menstrual cycle, but also during pregnancy, menopause, and of course, due to the intake of hormonal contraceptives. Whereas pregnancy and menopause are negligible to most existing MRI studies, due to ethical problems and too old age, the latter point is from enormous interest to the community. As worldwide more than 100 million people use hormonal contraceptives in the form of OCs as a method of choice for contraception (Petitti, 2003), in Germany, it is more than half of the female population in their childbearing years between 18 and 49 years (Heßling & Bundeszentrale für Gesundheitliche Aufklärung, 2011). Exactly the age group neuroscientists are interested in. Whereas the number of existing neuroimaging studies investigating the influence of naturally occurring sex hormones on brain functions is already limited, it is even worse for studies examining the effect of synthetic hormones, particularly the "pill" on the brain. Surprisingly little is known about the ways OCs affect their users' brain function. It is already well acknowledged for a decade that OCs show an altered mate preference compared to non-OCs users (Alvergne & Lummaa, 2010) and different brain activation patterns while watching erotic stimuli (Abler et al., 2013). Whereas these studies shed the first light into this by now neglected topic to the scientific community, they are, of course, also entertaining and not only interesting to professional scientists. However, systematical investigations of OCs dependent changes in "classical" robustly lateralized brain functions e.g., language, visuospatial attention, using fMRI, are still rare. Rumberg and colleagues showed increased activation in right-hemispheric task-specific areas in OC users compared to non-users during a word generation task (Rumberg et al., 2010). Furthermore, Pletzer and colleagues found masculinized brain activation patterns in numerical tasks (Pletzer & Kerschbaum, 2014), which cognitive demands can be related to spatial abilities (Hubbard et al., 2005). Further brain function differences between OC users and non-users are described during resting state (Petersen et al., 2014), and, for example, for reward- and face processing (Bonenberger et al., 2013; Marecková et al., 2014).

1.3 The Influence of Sex Hormones on Brain Structure

Brain structure differences between men and women have been described in several studies (Lentini et al., 2012; Ruigrok et al., 2014; Witte et al., 2010) with on average larger brains in males than in females and for example, larger gray matter (GM) volumes in amygdalae, hippocampi and temporal pole and orbitofrontal gyri in men, whereas women show larger thalami, precuneus, right insula cortex and right anterior cingulate gyrus (Ruigrok et al., 2014).

1.3.1 The Menstrual Cycle and Female Gray Matter

However, comparable to many fMRI studies examining differences between male and female brains, women's hormonal states have often not been considered. The first results on this topic were reported ten years ago by Protopopescu et al., (2008)., who compared women in their late follicular phase (high estradiol, low progesterone) and midluteal phase (medium estradiol, high progesterone). They found increased GM in the right anterior hippocampus and decreased values in the right globus pallidus and putamen in the late follicular phase (Protopopescu et al., 2008). These results were confirmed by Lisofsky and colleagues (Lisofsky et al., 2015) and, additionally, in a longitudinal single subject study (Barth et al., 2016).

Partly supporting results were described by Pletzer et al., who described slightly larger GM volumes in the right parahippocampal/fusiform gyrus during their early follicular phase (low estradiol and progesterone) compared to their midluteal phase (medium estradiol and high progesterone levels) (Pletzer et al., 2010). Further described brain regions, which are affected by different menstrual cycle phases, are the right middle frontal gyrus (MFG), the right anterior cingulate cortex (ACC), and the left insula. These regions showed larger volumes during the pre-ovulatory phase compared to midluteal phase (De Bondt et al., 2013a; De Bondt et al., 2016). Increased volumes in the left MFG but opposite results for the ACC, were reported by Protopopescu et al.(2008).

Using a sample size of 55 women to assess menstrual cycle-dependent effects, a recent study corroborated a significant pre-ovulatory estradiol-driven increase of bilateral hippocampal GM volumes and a significant progesterone-dependent increase of GM volumes of the right basal ganglia in the midluteal phase (Pletzer et al., 2018). Further information concerning overall structural changes across different hormonal states within a woman's life, including brain maturation, puberty, menstrual cycle, OC intake,

pregnancy and menopause is available within a recently published systemic review by Rehbein and colleagues (Rehbein et al., 2020).
Summarizing the above-described results, the hippocampus, basal ganglia and insula are possible targets of structural changes due to sex hormones fluctuations during the menstrual cycle, accompanied by trend findings in parahippocampal and fusiform regions, the ACC and MFG.

1.3.2 The Influence of "the Pill" on Brain Structure

The research field on the influence of the pill on brain structure is even younger. Less than 10 years ago, Pletzer et al. published a first exploratory study on this topic (Pletzer et al., 2010). They reported larger GM volumes in, inter alia, prefrontal cortex, ACC, parahippocampal and fusiform gyri and cerebellum in women using OCs. However, neither the "pill's generation" nor the chemical combination of the OC was taken into account. That these issues hold a strong influence with regard to brain structure, particularly GM volumes, could already be confirmed in a follow-up study conducted by the same authors a few years later. Pletzer and colleagues compared the effect of androgenic and antiandrogenic OCs on the structure of female brains (Pletzer et al., 2015a). Whereas antiandrogenic OCs lead to larger GM volumes compared to women with a natural cycle in bilateral fusiform gyri, the fusiform face area (FFA), parahippocampal place area (PPA) and the cerebellum, users of androgenic OCs displayed significantly smaller brain regions in the bilateral middle and superior frontal gyri.

1.4 What is the challenge of investigating the effect of sex hormones on the brain?

Despite the early research interest in this topic in general over 80 years ago (Frank, 1931), it is highly surprising that only a handful of scientists worldwide are actively examining the effect of sex hormones on the brain, no matter if they occur physiologically during a woman's lifespan (e.g., puberty, pregnancy, menopause), or how the application of hormonal contraceptives influence and manipulate their effects. One should think that nowadays, elaborated neuroimaging methods are feasibly available and could easily shed light on this intriguing topic, affecting millions of women worldwide. Thus, it might sound

surprising at first glance; however, scientists have been aware of methodological issues and challenges since the early 70s. Here, Sommer reviewed 33 existing publications investigating the effect of the menstrual cycle on cognition and perceptual motor behavior and found no evidence. However, she concluded that this result might be due to methodological problems, with which researchers are confronted by investigating the menstrual cycle (Sommer, 1973) since then. This issue might be an explanation for the relatively small amount of neuroimaging studies. Additionally, these studies, which are reviewed to a significant part by Sundström Poromaa & Gingnell (2014), are not consistent concerning the obtained results and the described hormonal effects. Not only is it very time consuming, but also the exact determination of hormonal state, e.g., from collected blood or salvia samples, has often not been state of the art yet. For example, in Protopopescu and colleagues' study, the menstrual cycle phase definition was very lenient, as it did not include hormone analyses (Protopopescu et al., 2008).

Furthermore, infrastructural issues occur, for example, MRI measurement appointments need to be reserved in advance and are often not spontaneously available; blood- or salvia samples need to be pretreated and correctly stored. These additional barriers and challenges might also explain the usually meager number of initially investigated subjects, e.g. Pletzer et al. (2010) only included 14 subjects and the high drop-out numbers, caused by later correction of the cycle phases .

Concerning the studies investigating the effects of OC, a further difficulty occurs: OCs are available in various combinations of synthetic hormones, particularly with regard to their androgenic modes of actions. However, that these issues hold a strong influence with regard to brain structure could already be confirmed in a follow-up study by Pletzer and colleagues. As written above, they examined the effects of androgenic and antiandrogenic OCs, resulting in opposed impacts on brain structure. However, again, they do not control for the exact hormone derivatives in the combined preparations (Pletzer et al., 2015a).

However, Pletzer and Kerschbaum have already summarized that more systemic research is needed to *"reveal the true nature of OC-dependent effects on cognition as well as the impact of synthetic steroids on neuronal correlates"* (Pletzer & Kerschbaum, 2014).

Accordingly, previous study results have to be considered with reservations, as different cycle phases were compared, relatively small sample sizes were examined, and exact hormonal determination is not present in all studies. Past studies relied on self-reports, which were rather unspecific and accompanied by high drop-out rates. Other studies did not determine hormonal concentrations at all; consequently, a data collection on the requested cycle phase cannot be guaranteed. Focusing on the effect of OCs on brain

structure, former studies did neither control for the pill “generation” nor the exact chemical combination. Additionally, different analysis pipelines were applied, using different brain parcellations, thus impeding the comparability of the yielded results.

With particular regard on brain function, fMRI studies, examining the effects of OCs and also the impact of female sex hormones in general, on visuospatial attention are still lacking. This derives quite an interesting question, as spatial tasks are yet known to favor men. Women, however, demonstrate menstrual cycle dependent variations of cognitive abilities and blood-oxygen-level dependent (BOLD)-response to cognitive tasks and score lower on mental rotation and further spatial tasks during the late follicular- and luteal phase (Pletzer et al., 2010; Schöning et al., 2007), in which estrogen and progesterone levels are high, compared to menstrual and early follicular phase when these hormones are lowest. So, the question arises, what effect does the suppression of these hormones due to OCs have on the cognitive performance and the BOLD-responses from the underlying neural correlates? To investigate this question, suitable fMRI paradigms are a necessary prerequisite. Regarding visuospatial processing, these are still missing, despite they are an essential prerequisite for investigating the effects of physiological fluctuations of female physiological sex hormones and of OCs on brain function, measured by their influence on the obtained FCAs.

2

Aims of the Thesis

Investigating the Effect of Sex Hormones on the Female Brain – Necessary Prerequisites and a First Insight on their Effects on Gray Matter Volume

The second chapter of this dissertation introduces the aims of the here presented thesis and how they address the gaps and limitations of previous research outlined in the previous chapter. It furthermore provides an overview of the projects included in this thesis and their content.

2 AIMS OF THE THESIS

This thesis aims to investigate the effect of sex hormones on the human brain. In particular, it asses the effects of endogenous and synthetic female sex hormones as they occur during the childbearing years of women.

As results of previous studies have to be considered with reservations, as different cycle phases were compared, rather small sample sizes were examined, and exact hormonal determination is not present in all studies as well as the application of varying analysis pipelines impeded the comparability of the yielded results, it is essential to establish optimal prerequisites in order to obtain robust and reliable data. Thus, an elaborate study design has been developed to investigate the effect of female sex hormones on the brain. As not the whole body of data is included in this thesis, I will introduce the scope of the study in the next paragraph, as the development of the study design, including the recruitment and examination of the included subjects (also containing the collecting of blood samples) and the analyses of sex hormones concentrations, the MRI measurements protocols, the applied fMRI paradigms, but also obtaining of funding and the positive ethic committee's vote was part of this thesis. However, from the first idea of this study to the analysis and interpretation of all obtained data would go far beyond the scope of this thesis.

Two groups of young women were included in the study:

Women with a natural menstrual cycle were invited to a personal visit three times. On the first visit, they were screened for the study's inclusions criteria. They were informed about the study's general procedure and aims and were given a thermometer and urine ovulation test stripes. They were instructed to protocol their cycle via measurement of their basal body temperature and identifying their ovulation using urine ovulation test stripes up to six months prior to the first data collection. They monthly informed us about their cycle lengths in order to organize the MRI appointments within the envisaged time windows. These were executed during the early follicular phase/menstruation (low concentrations of female sex hormones) and the midluteal phase (high concentration of female sex hormones).

We also included women, taking a defined combination and concentration of an OC. They were contacted via phone and screened for the study's inclusion criteria and their actual intake behavior of "the pill". They were invited twice: Once during the "pill" withdrawal and once during the last week of "pill" intake.

Prior to the MRI measurements, venous blood samples were collected from all women and locally processed and stored in order to analyze progesterone and estradiol concentrations. The MRI protocol contained the survey of structural data, in particular obtained from T1 and Diffusion tensor imaging (DTI) measurements and of functional data. With regard to structural data, we were interested in the effects of physiological and synthetic female sex hormones on GM, white matter (WM), total intracranial volume (TIV) and white matter tractography (DTI). (With DTI, white matter microstructures within the central nervous system (CNS) can be revealed.)

Concerning functional data, we were interested in the effects of female sex hormones on different cognitive functions, in particular language and visuospatial functions, and face perception, representing activities that can be assigned mainly to one of the hemispheres. We were interested in which extent these functions, especially their hemispheric lateralization and thus intra- and interhemispheric connectivity, are affected by sex hormone fluctuations. In order to investigate language functions, we applied two well-established fMRI paradigms, a word generation task (adapted from Jansen et al. (2005a)) and a semantic decision task (adapted from Fernández et al. (2001), Jansen et al. (2006)). A face localizer task (Frässle et al., 2016) was slightly adapted in order to assess face processing. For visuospatial functions, paradigms are markedly less well established. Here, paradigms of choice to examine visuospatial functions have been the Landmark Task (Fink et al., 2000) and the Benton Judgement of Line Orientation Test (Benton et al., 1994). Therefore, the first part of this thesis aimed to develop and to examine potentially suitable paradigms investigating visuospatial functions thoroughly. Additionally, to the described fMRI paradigms, resting-state fMRI data is acquired to evaluate regional interactions during a resting state period (Biswal, 2012). For an overview of the entire project, please see Fig. 5.

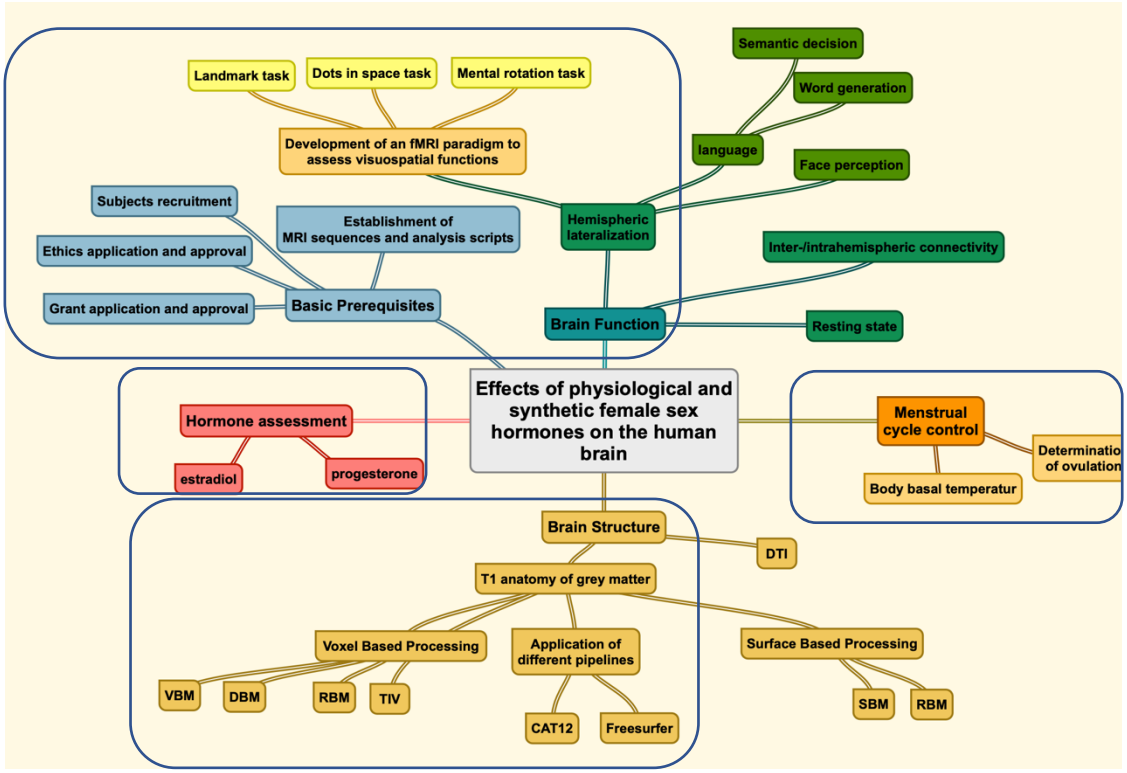


Figure 5 *Project Overview: Effects of physiological and synthetic female sex hormones on the human brain.*

Part of the PhD work was the formation of the study design, including recruitment and screening of participants, development and installation of MR sequences and fMRI paradigms, as well as the establishment of the study's infrastructure, including ethics committee vote and funding. Structural in addition to functional data, as well hormonal data from blood samples were acquired. The subprojects, examined, analyzed and described in this thesis are framed in blue. T1: T1-weighted MR image, DTI: Diffusion Tensor Imaging, VBM: Voxel-Based Morphometry, DBM: Deformation-Based Morphometry, RBM: Region-Based Morphometry, SBM: Surface-Based Morphometry, TIV: Total Intracranial Volume, CAT12: Computational Anatomy Toolbox 12.

Within this thesis, the study design was conceptualized and established, infrastructural paths were paved, e.g., external funds were raised enabling to perform the study, ethics approval were obtained, the collaboration with the main laboratory of the university hospital was initiated, fMRI paradigms, as well as analysis scripts, were programmed, subjects were intensively screened, supervised and measured. Within the framework of this dissertation, the focus is on two main aspects: One facet was the establishment of the study design per se, particularly the establishment of reliable and robust paradigms to investigate visuospatial functions (project 1), whereas the second priority was the analysis of structural GM brain data (project 2). These aims were therefore investigated separately within two different research projects and are introduced in the subsequent

paragraph. Further aspects of the overall study will be presented briefly within the closing discussion.

2.1 Aims of project 1 – In search of a robust and reliable fMRI paradigm

To examine the effect of female sex hormones on the human brain, particularly on brain function and the potential influence on FCAs and the interhemispheric connectivity, suitable, notably robust and reliable fMRI paradigms are an absolutely necessary prerequisite. “Robust” means whether the paradigm is able to activate the same lateralized network both at the group level in a specific subject population and in a certain number of subjects at the single-subject level. “Reliable” describes the linearity of results, notably whether (or not) a researcher is able to predict future fMRI results from previous ones. Furthermore, as interindividual differences are present not only on the hormonal level but also in brain activation patterns, a thorough analysis of hemispheric lateralization necessitates systematic analyses of these inter-individual differences. For this, experimental paradigms must provide robust hemispheric lateralization measures not only at the group level but also in individual subjects (Brandt et al., 2013; Jansen et al., 2006; Lohmann et al., 2004).

As robust paradigms already exist to investigate typically left-hemispheric lateralized cognitive functions, e.g., language processing, and have already applied to females during different hormonal states (Rumberg et al., 2010), this is not the case for typically right-hemispheric lateralized function, e.g., visuospatial functions. However, this is essential for investigating the effects of female sex hormones on functional brain connectivity, measured by their influence on the obtained FCAs.

Project 1 aimed to establish an fMRI paradigm that robustly and reliably evokes right-hemispheric dominance of fMRI activation patterns both at the group and single-subject level. Therefore, we focused on visuospatial processing as a “typical” right-lateralized cognitive process of the human brain, in line with an extensive body of previous literature, e.g. (Fink et al., 2000; Kinsbourne & Bruce, 1987; Mesulam, 1981). Specifically, three paradigms (“dots-in-space” task, mental rotation task, and Landmark task) that had been used frequently in imaging studies to determine hemispheric dominance were compared and their respective utility to provide *robust* and *reliable* estimates of right-hemispheric lateralization was evaluated. This is very important, as the evoked changes in activation patterns due to hormonal differences are expected to be very small. A suitable, reliable, and robust paradigm for testing right-hemispheric function is essential for the general

understanding of the human brain's function, especially to understand the interaction of the hemispheric dominance of different brain functions such as language and spatial attention and to which extent these are dependent to hormonal changes. To our knowledge, the influence of female sex hormones on the interaction of the hemispheric dominance of different brain functions has yet not been investigated systematically, as suitable fMRI paradigms examining right hemispheric functions are still not available by now.

Therefore, the aim of the first project was to establish a paradigm that fulfills these strict requirements. This paradigm should then be applied to an fMRI test battery, constituting further reliable and stable paradigms to investigate FCAs and their dependency on hormonal fluctuations.

Project 1 was split into two studies. In the first study, three frequently used paradigms for assessing visuospatial processing in 15 subjects were compared and evaluated their utility to robustly detect right-lateralized brain activity on a single-subject level. In the second study, the test-retest reliability of the so-called Landmark task, the fMRI paradigm that yielded the most robust results in study 1, was then assessed in 20 participants. To investigate the test-retest reliability of the imaging paradigms, all subjects in study 2 underwent the identical experiment twice in two separate sessions.

2.2 Aim of project 2 – Shedding Light in Regional Grey Matter Differences in Women – How do Menstrual Cycle Phases and the Intake of 'The Pill' shape the Female Brain?

The aim of project *2* focused on the effect of female sex hormones, as they occur naturally during the menstrual cycle or due to the intake of OCs, on the structure of the female brain. At present, the data situation is conflicting and inconsistent. Reasons for this are twofold:

First, there is a limited number of existing studies examining this topic by now, as menstrual cycle research is costly and time-consuming per se.

The serious difficulty of precisely obtaining data from women with a natural cycle on the exact days of the menstrual cycle to investigate a distinct cycle phase is a major challenge. Additionally, the exact combination of synthetic hormones of OCs has been neglected in a vast amount of studies.

Second, different data analysis pipelines relying on different brain parcellations were used, reducing the comparability of different studies' results. GM volume is widely analyzed using a method named voxel-based morphometry (VBM). This approach

implements the voxel-wise estimation of the local amount or the local volume of specific brain tissue (e.g., GM, white matter (WM), cerebral spinal fluid (CSF)). De Bondt and colleagues evaluated the reproducibility using two slightly different software packages for GM volume analysis in the widely used analyses program SPM8 and SPM12, finding, again, inconsistent results (De Bondt et al., 2016). In 2016 a new toolbox for VBM analysis, namely the computational anatomy toolbox (CAT12), was introduced by Gaser and colleagues, which uses completely different segmentation approaches than SPM12 does. Additionally, further analysis steps differ, e.g., interpolation, spatial normalization (http://www.neuro.uni-jena.de/hbm2016/GaserHBM2016.pdf).
More than 600 published studies using this newly established method are listed by *Google Scholar* by now (21.09.2020, search items “cat12” and “vbm”).

Until now, the obtained mixed and partly inconsistent results of the effects of either naturally fluctuating female sex hormones or exogenously applied hormones might be attributable to different study designs and different methods of analysis.
In project 2, we, therefore, aimed to shed light on this rapidly developing research field by establishing an elaborated study design to investigate the effects of female sex hormones on the brain. This study design is then applied to examining the effects of different concentrations of female sex hormones during the menstrual cycle and the influence of synthetic hormones derived from “the pill” in a strictly controlled cohort of the study. We investigate the effect of sex hormones as they occur naturally during two controlled opposed cycle phases (early follicular-and midluteal phase) and as they are suppressed by the intake of one defined hormonal OC on GM volume during the pill-free week and during pill-intake. At this point, it needs to be clarified that not the direct effect of sex hormones is analyzed within project 2, but, with regard to women with a natural cycle, the effect of the two opposed cycle phases, which are entirely governed by distinct levels of the female sex hormones estradiol and progesterone. Focusing on the investigated women taking OCs, here, the influence of female sex hormones was indirectly analyzed by examining the effect of the active OC intake or the withdrawal, respectively. In order to preserve the reading flow, the term “effects of sex hormones” is therefore equally used.

Furthermore, two state-of-the-art methods, specifically designed to process structural MRI data, notably, the above described CAT12 toolbox as well as the software package Freesurfer were applied to investigate potential volume changes and to allow a direct comparison of the possible obtained differences. The latter analysis package, Freesurfer, was previously known as a gold-standard, with regard to surface analysis,

for which it is the most widely used (Nakamura et al., 2011); however, it also offers tools to analyze brain volumetric measures which are highly esteemed by the scientific community.

2.3 Summary of Aims

In summary, this thesis aimed to fill two missing puzzle pieces in order to shed light into and support the rapidly growing research field of the effect of sex hormones on the human brain in general, but in particular, the effect of endogenous and synthetic female sex hormones as they occur during the childbearing years of women. Aims of this thesis are an indirect and a direct contribution to this topic:

With project 1, the indirect contribution, I pursuit for the first missing puzzle piece in order to investigate hemispheric lateralization of cognitive functions and its potential dependency on sex hormones: a robust and reliable fMRI paradigm to investigate right-hemispheric brain activation induced by visuospatial processing not only on the group- but on single-subject level. This is the prerequisite for future studies investigating potentially small effects on brain function, as they are caused e.g., by hormonal fluctuations.

With project 2, I aimed to establish an elaborated study design and detangle the effects of female sex hormones during defined hormonal states on brain structure. Hormonal states of interest were in this thesis, the early follicular- and the midluteal phase of the menstrual cycle. These phases demonstrate the highest differences in the female sex hormone progesterone, which has been described as a potential cause for FCA changes during the menstrual cycle. Additionally, this thesis included the hormonal states of the intake of a distinct pill and its withdrawal.

This project constitutes a direct contribution to this research area. It aims to deliver the missing puzzle pieces to which extent structural changes depend on female sex hormones and the role of the applied analysis routine. This thesis aims to promote future studies with a well-established study design and further illuminate the role of female sex hormones on brain structure.

3

Project One

Comparison of fMRI Paradigms assessing visuospatial Processing: Robustness and Reproducibility

Within the third chapter of the here presented thesis, the first project is presented. In a first study, three frequently used paradigms for assessing visuospatial processing were therefore compared and evaluated regarding their utility to robustly detect right-lateralized brain activity on a single-subject level. In a second study, the test-retest reliability of the so-called Landmark task – the paradigm that yielded the most robust results in study 1 was assessed.

3 PROJECT ONE:

COMPARISON OF FMRI PARADIGMS ASSESSING VISUOSPATIAL PROCESSING: ROBUSTNESS AND REPRODUCIBILITY

The content of Project 1 has been published as:

***Schuster V**, Herholz P, Zimmermann KM, Westermann S, Frässle S, Jansen A (2017) Comparison of fMRI paradigms assessing visuospatial processing: Robustness and reproducibility. PLoS ONE 12(10): e0186344. https://doi.org/10.1371/journal.pone.0186344*

Author Contributions

Conceptualization:	Verena Schuster, Stefan Frässle, Andreas Jansen.
Data curation:	Verena Schuster, Peer Herholz, Stefan Frässle.
Formal analysis:	Verena Schuster, Peer Herholz, Stefan Frässle.
Funding acquisition	Andreas Jansen.
Investigation:	Verena Schuster, Peer Herholz.
Methodology:	Verena Schuster, Andreas Jansen.
Project administration:	Verena Schuster, Peer Herholz.
Resources:	Kristin M. Zimmermann, Stefan Westermann, Stefan Frässle, Andreas Jansen.
Software:	Peer Herholz, Kristin M. Zimmermann, Stefan Westermann, Stefan Frässle, Andreas Jansen.
Supervision:	Stefan Frässle, Andreas Jansen.
Validation:	Verena Schuster, Peer Herholz, Andreas Jansen.
Visualization:	Verena Schuster, Peer Herholz.
Writing – original draft:	Verena Schuster, Stefan Frässle.
Writing – review & editing:	Verena Schuster, Peer Herholz, Kristin M. Zimmermann, Stefan Frässle, Andreas Jansen

3.1 Introduction

Hemispheric specialization is a fundamental principle of human brain organization and describes the fact that different cognitive or executive processes are distributed differently across the two hemispheres of the brain. Such functional asymmetries between the hemispheres have been known since the mid-19th century (P. Broca, 1861a, 1861b, 1861c), and their study has been made more widely feasible with the development of modern brain imaging techniques. In particular, functional transcranial Doppler sonography (fTCD) and functional magnetic resonance imaging (fMRI) for the first time enabled non-invasive studies of the hemispheric lateralization of cognitive functions in large cohorts of both patients and healthy subjects (Stephan et al., 2007). Using these methods, researchers have mapped the hemispheric lateralization of various cognitive functions and highlighted that for all these processes, the degree of lateralization is subject to inter-individual variability. For instance, while most people show left-hemispheric dominance for language, an atypical right-hemispheric or bilateral form of language lateralization has been observed in up to 10% of the human population (Frost, 1999; Jansen, et al., 2005b; Knecht, Dräger, et al., 2000; Pujol et al., 1999; Springer et al., 1999). On the contrary, visuospatial attention is lateralized predominantly to the right hemisphere, again subject to marked variability across subjects (Fink et al., 2000; Kinsbourne & Bruce, 1987; Mesulam, 1981).

A thorough analysis of hemispheric lateralization necessitates systematic analyses of these inter-individual differences. For this, experimental paradigms must provide robust measures of hemispheric lateralization not only at the group level, but also in individual subjects (Brandt et al., 2013; Jansen et al., 2006; Lohmann et al., 2004). This refers to the test-theoretical concept of test-retest reliability (or within-subject stability) and has been studied for various lateralized cognitive processes, such as language (Gorgolewski et al., 2013; Jansen et al., 2006), face processing (Frässle et al., 2016), motor processing (Frässle, Stephan, et al., 2015; Kristo et al., 2014; Weiss et al., 2013) and declarative memory (Brandt et al., 2013). However, the reliability of imaging paradigms assessing a typical right-lateralized cognitive function, i.e., visuospatial processing, has received considerably less attention so far.

The aim of the present study was therefore to establish a paradigm that *robustly* and *reliably* evokes right-hemispheric dominance of fMRI activation patterns both at the group and single-subject level (note that robustness and reliability are distinct test-theoretical concepts and a definition of both metrics is given in the Methods section). Here, we focused on visuospatial processing as a “typical” right-lateralized cognitive

process, in line with an extensive body of previous literature, e.g. (Fink et al., 2000; Kinsbourne & Bruce, 1987; Mesulam, 1981) (but see the Discussion for a critical review of this assumption). Specifically, we compared three paradigms ("dots-in-space" task, mental rotation task, and Landmark task), that had been used frequently in imaging studies to determine hemispheric dominance and evaluated their respective utility to provide *robust* and *reliable* estimates of right-hemispheric lateralization. In what follows, we briefly introduce the different paradigms.

First, we used the "dots-in-space" task, which has been described as a "robust and reliable method for investigating laterality of visuospatial skills" in a previous fTCD study (Whitehouse & Bishop, 2009). In the "dots-in-space" task, participants have to memorize the location of circles, which are randomly distributed on a black screen. This primarily assesses spatial memory skills, which was found to engage a right-hemispheric brain network (Awh & Jonides, 2001), comprising areas in the ventrolateral frontal cortex, occipital cortex, parietal cortex and premotor cortex (Smith et al., 1996). However, until now, this paradigm has not been translated into an fMRI setting.

Second, the mental rotation task was implemented, testing a second aspect of visuospatial abilities, namely spatial orientation. The mental rotation task has been used previously both in fTCD (e.g., Dorst et al., (2008)) and fMRI studies (e.g., Hattemer et al., (2011)), however, yielding contradictory results with regard to the evoked lateralization. For example, Dorst and colleagues found right-lateralized activation in the majority (72.4 %) of participants (Dorst et al., 2008). In contrast, Hattemer and colleagues did not observe significant lateralization to the right hemisphere using either fTCD or fMRI, but bilateral activation in the middle and superior frontal gyrus, the insular cortex, thalamus, mesencephalon and cerebellum (Hattemer et al., 2011).

Third, we tested the Landmark task, which is often used in fMRI studies that investigate spatial attention, representing another important aspect of visuospatial abilities. The Landmark task originates from a clinical setting, where it served as a bedside test of hemi-spatial neglect. The underlying cognitive processes involved in the Landmark task are summarized in Ciçek et al. (Cicek et al., 2009) and include spatial judgments, sustained attention and object-based spatial processing, leading to an activation of the dorsal attention network.

Hence, all three fMRI tasks involve somewhat distinct aspects of visuospatial abilities (e.g., spatial memory, spatial orientation, spatial attention), but have all been linked to an increased recruitment of right-hemispheric brain regions. As the aim of the present study was to establish an fMRI paradigm that robustly and reliably evokes right-hemispheric lateralization in the human brain, we compared these different right-

hemispheric lateralized cognitive processes under the umbrella term of visuospatial processing.

The present study was divided into two parts. In the first part ("study 1"), we focused on the comparison of the above-mentioned imaging paradigms for studying visuospatial processing: the "dots-in-space" task (adapted from Whitehouse et al., (2009)), the mental rotation task (Dorst et al., 2008) and the Landmark task (Fink et al., 2000). The aim of study 1 was to test which of these three paradigms was able to robustly map right-hemispheric dominance. In the second part ("study 2"), we then assessed the test-retest reliability of the paradigm that yielded the most robust results in study 1.

3.2 Methods

3.2.1 Subjects

Sixteen subjects (6 men, mean age: 24.7 ± 2.5 years) participated in study 1 (comparison of imaging paradigms). Notably, one subject had to be excluded from study 1 because of uncomfortableness in the scanner, yielding 15 remaining subjects. In study 2 (assessment of test-retest reliability), 20 subjects (10 men, mean age: 25.0 ± 2.2 years) participated. To investigate the test-retest reliability of the imaging paradigms, all subjects in study 2 underwent the identical experiment twice in two separate sessions. The time interval between sessions ranged from 5 to 8 days (mean time interval: 6.9 ± 0.2 days). All 36 subjects were right-handed, had completed the equivalent of a high school degree ("Gymnasium") and were native German speakers. None had any history of medical, neurological or psychiatric illnesses or brain pathology. All subjects had normal or corrected to normal vision. Each gave informed written consent prior to participation. The study conformed with the Declaration of Helsinki and was approved by the local ethics committee of the Medical Faculty of the University of Marburg (file reference 85/13).

3.2.2 Experimental paradigms

In study 1, subjects performed three different tasks, which are well established for testing hemispheric lateralization during spatial processing (see below for a detailed description of each task): "dots-in-space" task, adapted from (Whitehouse & Bishop, 2009), mental rotation task (Dorst et al., 2008), and Landmark task (Fink et al., 2000). The order of these tasks was pseudorandomized and counterbalanced across subjects. The aim of study 1 was to identify which of these paradigms evoked robust right-hemispheric

lateralization both at the group and single-subject level. In study 2, we then evaluated the test-retest reliability for the most robust paradigm (as quantified by our pre-defined criteria of robustness described below). All paradigms were implemented and displayed using the Presentation® Software package (Version 14.1, http://neurobs.com). Prior to the experiment, subjects practiced each task outside the MR-scanner to ensure that they had understood the instructions. During the fMRI measurements, responses were reported by pushing a button on an MR-compatible response box, which was located on the left and right thigh.

"Dots-in-space" task: The "dots-in-space" task used in the present study was based on a spatial memory task originally developed for fTCD (Whitehouse & Bishop, 2009). Subjects had to memorize the location of a number of red dots randomly interspersed with a larger number of white dots, presented on a black background (Fig. 1). The dots were randomly distributed across the screen and not aligned in rows or columns to prevent verbal encoding strategies. The task was divided into two parts: an encoding phase and a retrieval phase. Subjects were shown 20 different arrangements of white and red dots ("target stimuli"). Each target stimulus was shown three times in a pseudorandomized order. Ten of the target stimuli consisted of 17 white and 9 red dots ("difficult condition"), whereas the remaining target stimuli consisted of 23 white and 3 red dots ("easy condition"). During the encoding phase, each stimulus was presented for 5 s and then followed by a blank screen with an inter-stimulus interval (ISI) of 1 s. Subjects were instructed to memorize the location of the red dots.
Functional images were acquired only during the retrieval phase, which consisted of 18 blocks (Fig. 2a). Six blocks belonged to the difficult condition, six blocks to the easy condition (i.e., test conditions), and six blocks to a control condition (see below). Again, the order of conditions was pseudorandomized. Each block started with a black screen (3 s) followed by an instruction screen that indicated the task condition of the upcoming trial (5 s). After the instructions, six stimuli ("test stimuli") were presented. Each test stimulus was shown for 2 s followed by a jittered ISI (average length 1 s; range: 0.6 s to 1.4 s) and a black screen after the sixth stimulus (5 s). This resulted in a total block length of 30 s. During the retrieval phase, subjects were asked to decide whether the presented test stimulus was familiar or not. Subjects responded by pressing a button with the index finger (yes) or middle finger (no) of their right hand on the MR-compatible response box. During the control condition, subjects were instructed to decide whether a stimulus contained exactly one red circle or not. Stimuli in this condition consisted of either only white dots or white dots and exactly one red dot randomly located on the black background (Fig. 6a). The total length of the paradigm was 9 minutes 10 seconds.

Mental rotation task: The mental rotation task used in the present study was based on a spatial orientation task that had also been originally introduced for fTCD (Dorst et al., 2008; Hattemer et al., 2011). Subjects were presented with pairs of three-dimensional images of transparent cubes with 1, 2 or 3 cables inside showing the same object from two different perspectives (Fig. 6b). During the activation condition (high spatial processing load), subjects were presented with pairs of identical cubes, however, the right-sided cube was always seen from different perspectives. Subjects were asked to decide whether the right-sided cube showed the left-sided cube seen from the left, right, back, top or bottom. According to the perspective in which the right cube was presented, answers were indicated by pressing either the right thumb (perspective: left), index finger (bottom), middle finger (top), ring finger (back) or the little finger (right). During the control condition (low spatial processing load), either pairs of identical cubes shown from the same perspective or two different cubes were presented. Subjects had to decide whether the presented cubes were identical or not. Subjects responded by pressing a button with the index finger (same) or middle finger (different) of their right hand. The paradigm consisted of 10 control and 10 activation blocks that were presented in a pseudorandomized order (Fig. 7b). In each block, 4 stimuli were shown for 5 s each. Before each block, an instruction screen informed the subjects about the condition of the upcoming block (5 s). Each block was followed by a baseline period of 20 s, resulting in a total block length of 45 s. The total length of the paradigm was 15 minutes.

Landmark task: The Landmark task has been used frequently to study spatial attention by means of functional imaging (Fink et al., 2000; Flöel et al., 2005; Jansen et al., 2006; Jansen et al., 2005a) The paradigm consisted of two conditions. In the activation condition (high spatial processing load), subjects had to decide whether a horizontal line was correctly bisected by a crossing vertical line or not. In the control condition (low spatial processing load), subjects had to decide whether a horizontal line contained a transection mark (irrespective of the position of that mark) or not (Fig. 6c). Eight activation and eight control blocks were presented in an alternating order. Each block lasted 20 s and contained 11 stimuli that were presented for 0.6 s followed by an ISI of 0.9 s. Stimuli were presented in the four corners of the screen while subjects had to fixate the center of the screen. This prevented subjects from solving the task by simply attending a single point as the center of all lines without the need to engage in spatial processing. Each block was preceded by an instruction screen displayed for 1.5 s, informing subjects about the condition of the upcoming block. The total length of the paradigm was 5 minutes 34 seconds.

We used three different versions of the Landmark task (version A in study 1, version B and C in study 2). We slightly adapted the task in study 2 for the following two reasons:

First, although the Landmark task was the most robust paradigm in study 1 according to our criteria (see below), the BOLD signal difference between activation and control condition was relatively small, as expressed, e.g., by low t-values. One reason for this might have been that the activation task was too simple, as indicated by the behavioral results. We therefore made the task more difficult in version C by using more demanding visual stimuli in the activation condition (Fig. 6c). Second, we adapted both the control stimuli and the instructions (see below) to make our version of the Landmark task more similar to those used in recent work (Cai et al., 2013; Gorgolewski et al., 2013). Additionally, consistent with previous studies (Cicek et al., 2009; Fink et al., 2000), proper fixation of the subjects was explicitly controlled in versions B and C by online visual inspection of the recorded traces of the direction of eye gaze using an MRI-compatible infrared-sensitive camera (EyeLink 1000, SR Research, Osgoode, ON, Canada). Specifically, a qualitative screening of the eye-tracking data was performed to identify subjects that poorly fixated the central cross during the experiment (e.g., performing saccades to the presented stimulus) and should thus be excluded. Importantly, no such cases were observed and thus all subjects were included in the subsequent analyses. Note that eye-tracking data in the present study did not enter any further analysis.

In version A, we used horizontal lines with or without a transection mark as control stimuli (Fig. 6c left and middle) and subjects had to decide whether the transection mark was present or not (irrespective of the position of that mark). In the activation condition, subjects had to decide whether the horizontal line was transected left or right from the middle or whether the vertical line crossed the horizontal line on the left or right side. For both conditions, subjects reported their decision with either their right index (“right side” or “transection mark is present”) or middle finger (“left side” or “no transection mark”). In versions B and C, all stimuli contained a vertical line. In the control condition, the vertical line crossed the horizontal line in half of the images, whereas the vertical line was above or beneath the horizontal line for the other half (Fig. 6c right). Subjects had to decide whether the vertical line crossed the horizontal line or not. In the activation condition, subjects had to decide whether the horizontal line was correctly bisected or not. For both conditions, subjects reported their decision using both hands. Specifically, they indicated their answer by pressing both index (“correctly bisected” or “vertical line transects”) or middle fingers (“not correctly bisected” or “vertical line does not transect”) simultaneously.

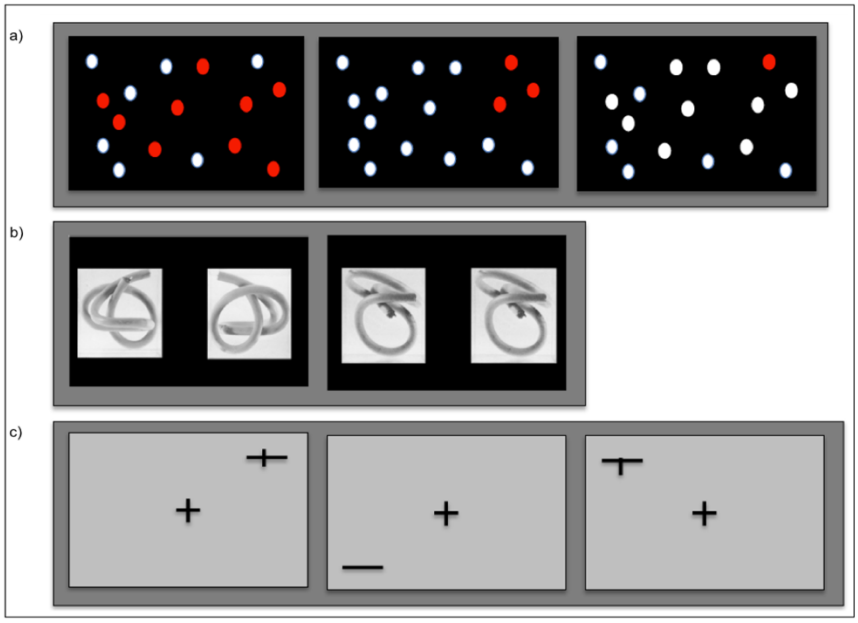


Figure 6 *fMRI-Stimuli.*

a) "Dots-in-space" task: Activation (left, middle) and control-condition (right) of the "dots-in-space" task. In the activation condition subjects were asked whether they have seen the same arrangement of red dots during the encoding part. In the control condition, they were asked to decide whether there was a red dot or not.

b) Mental rotation task: Activation (left) and control-condition (right) of the mental rotation task. In the activation condition subjects were asked to indicate, whether the figure on the right, showed the figure on the left from behind, the bottom, the top, the left, or the right by pressing the respective finger of their right hand. In the control condition, they were asked to decide whether the two figures were identical or not. All stimuli were taken from Stumpf and Fay (1983), scanned in and corrected for contrast and brightness differences.

c) Landmark task: Activation (left) and control-condition of version A (middle) and version B and C (right) of the Landmark task. During the activation condition, subjects were asked to decide whether the horizontal line was transected left or right from the middle (version A) or whether the line was bisected correctly (version B and C). The horizontal line, which measured 200 pixels (13.48 cm), appeared 0.6 s in one of the four corners of the screen. In versions A and B, the vertical line was centered either exactly in the middle of the horizontal line or slightly deviated to the left or the right. Distances of 15, 30 and 45 pixels (resulting in 1.01, 2.02, and 3.03 cm lengths and visual angles of 0.241°, 0.482° and 0.723° respectively) were used to shift the vertical line to either side. Version C was characterized by smaller distance variations. Distances to the middle of the horizontal line were 12, 25 and 37 pixels, resulting in 0.809, 1,685 and 2.493 cm and visual angles of 0.193°, 0.402° and 0.505° respectively. In the control condition of version A (middle) subjects were asked to decide, whether a transecting line was present, whereas in the control condition of version B and C, subjects had to decide whether the vertical line transected the horizontal one or not (right figure). Answers were indicated with the index- and middle finger of the right hand (version A) or of both hands (version B and C).

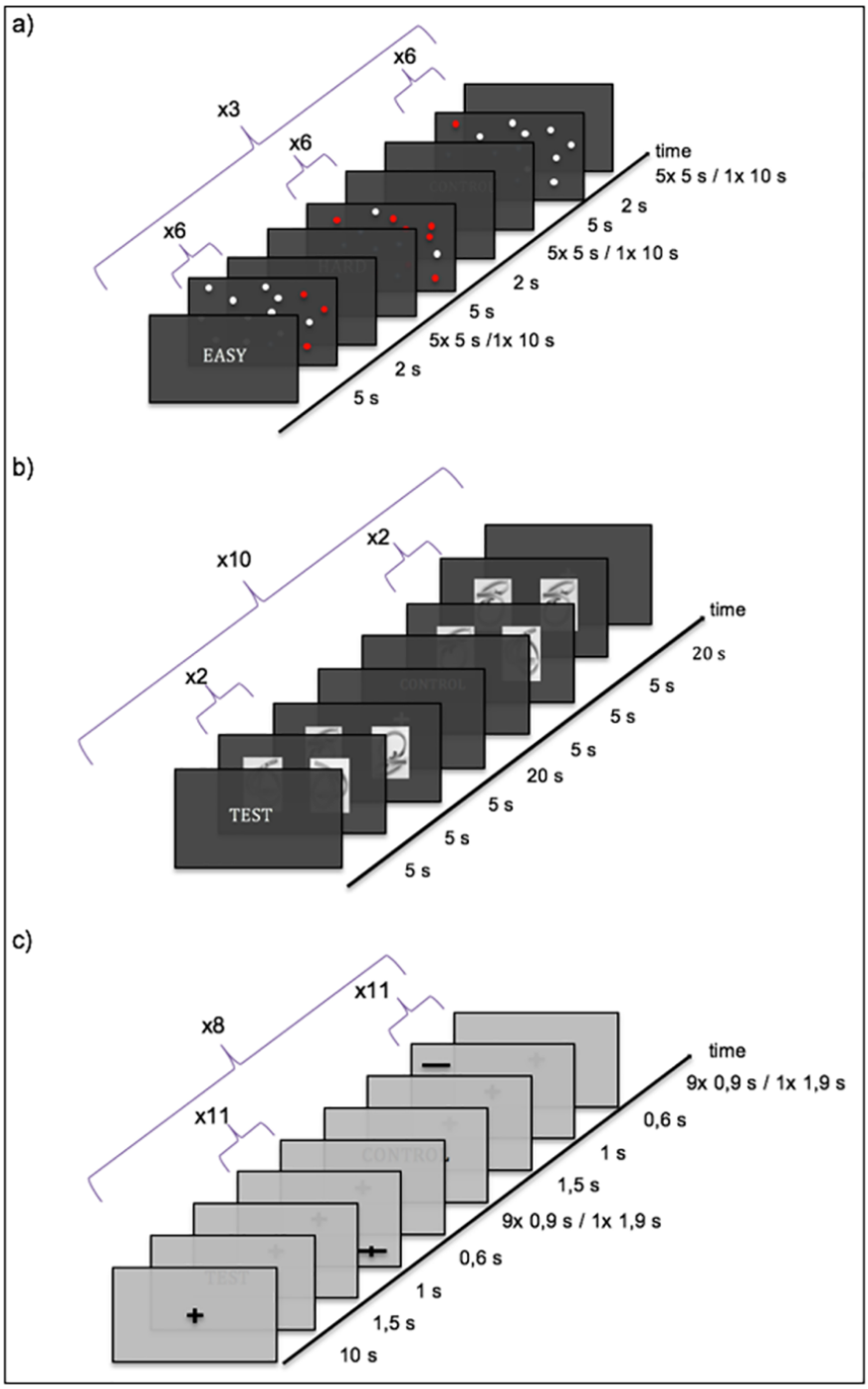


Figure 7 *Schematic representation of the experimental procedure.*

a) "Dots-in-space" task: Each experimental trial began with an introduction screen (5 s) indicating the following condition (easy, hard, control). Participants were asked, whether they have seen the same arrangement of dots during the encoding part or not (easy, hard), or whether the arrangement contained a red dot (control) by pressing a button with the index- or middle finger of their right hand, respectively. The paradigm consisted of three blocks of each condition, with six stimuli within each block.

b) Mental rotation task: Each experimental trial began with an introduction screen (5 s) indicating the following condition (test, control). Participants were asked, to decide whether the picture on the right side showed the identical cube on the left side seen from the left, right, back, top or bottom. Answers

were indicated by pressing either their right thumb (from the left), index finger (from the bottom), middle finger (from the top), ring finger (from the back) or the little finger (from the right). During the control condition (low spatial load), either pairs of identical cubes shown from the same perspective or two different cubes were presented. Subjects had to decide whether the presented cubes were the same or not. Subjects gave their answer by pressing a button with the index finger (same) or middle finger (different) of their right hand on a MR-compatible response box. The paradigm consisted of 10 control and 10 activation blocks that appeared in pseudorandomized order. In each block 4 stimuli were shown. Each stimulus was presented for 5 s.

c) Landmark task: The paradigm began with a fixation screen (10 s) prior the experimental conditions. Participants were instructed to fixate the cross during the whole experiment.

Each experimental trial began with an introduction screen (1.5 s) indicating the following condition (test, control). In version A participants were asked, whether the horizontal line, appearing in one of the four corners of the screen, was intersected left or right from the middle point (test) or whether a transection mark was present or not (control). Answers were indicated by pressing a respective button with their right hand. In version B and C participants were asked to decide whether the horizontal line was bisected correctly or not. In the control condition, they were asked to indicate, whether the horizontal line was transected by the vertical one or not. Answers were indicated by pressing the respective button with both hands, simultaneously.

3.2.3 MRI data acquisition

Subjects were scanned on a 3-Tesla TIM-Trio MR Scanner (Siemens Medical Systems) with a 12-channel head matrix receive coil at the Department of Psychiatry and Psychotherapy, University of Marburg. Functional images were acquired using a T_2*-weighted echo planar imaging (EPI) sequence sensitive to the Blood Oxygen Level Dependent (BOLD) contrast. Slices covered the whole brain and were positioned transaxially parallel to the anterior-posterior commissural line (AC-PC). In study 1, the following parameters were used: matrix size 64×64 voxels, FoV = 210 mm, 30 slices (ascending), slice thickness 4.5 mm (10% gap), TR = 1600 ms, TE = 30 ms, flip angle 90°. For the “dots-in-space” task (study 1), we used slightly different parameters (FoV = 192 mm, 35 slices (ascending), slice thickness 4 mm (10% gap), TR = 2150 ms). In total, 208 functional images were collected during the “dots-in-space” task, 215 scans for the Landmark task (version A) and 569 images for the mental rotation task. The initial images were excluded from further analyses in order to remove the influence of T1 stabilization effects.

In study 2, we aimed to optimize the acquisition sequence for the Landmark task (version B and C) in order to boost the relatively low t-statistics observed in study 1. Specifically, we used a sequence that had previously been shown to provide high BOLD sensitivity and – more importantly for the goal of study 2 – excellent test-retest reliability of BOLD

activation for a face perception paradigm (Frässle et al., 2016). The following scanning parameters were used: matrix size 64×64 voxels, FoV = 192 mm, 30 slices (descending), slice thickness 4 mm (15% gap), TR = 1450 ms, TE = 25 ms, flip angle 90°. In total, 222 functional images were collected for each subject.

3.2.4 MRI data analysis

All fMRI data were analyzed using the standard routines and templates from the software package SPM8 (v4290; www.fil.ion.ucl.ac.uk/spm) in MATLAB 7.7.0.471 (R2008b) (The MathWorks, Inc.). Functional images were realigned, normalized (using the standard SPM EPI-Template), resampled to a voxel size of 2×2×2 mm^3, smoothed with a 5-mm isotropic Gaussian kernel, and high-pass filtered (cut-off period 128 s). After pre-processing, statistical analysis was performed in a two-stage, mixed-effects procedure. At the single-subject level, BOLD responses were modelled in a General Linear Model (GLM) using boxcar functions convolved with the canonical hemodynamic response function from SPM8 (Friston et al., 1995; Worsley & Friston, 1995). For the "dots-in-space" task, we modelled four conditions (i.e., control, easy, hard, and baseline; instructions were not modelled). For the mental rotation task, we modelled three conditions (i.e., high spatial processing load, low spatial processing load, and fixation baseline; instructions were not modelled). For the Landmark task, we modelled two conditions (i.e., activation and control; instructions were not modelled). Additionally, the six realignment parameters were included as nuisance regressors in each design matrix to control for movement-related artifacts. For each paradigm and subject, contrast images were computed by contrasting activation and control conditions. More specifically, the following linear contrasts were calculated for each subject: "dots-in-space" task: "easy + hard > 2*control"; mental rotation task: "high spatial processing load > low spatial processing load"; Landmark task: "activation > control". At the group level, individual contrast images for each paradigm were entered into separate one-sample t-tests. The anatomical localization of activated brain regions was assessed both by the SPM anatomy toolbox (Eickhoff et al., 2005) and the WFU-Pickatlas (Maldjian et al., 2003).

3.2.4.1 Study 1: Comparison of imaging paradigms

In study 1, we tested whether the three visuospatial processing paradigms were able to robustly determine right-hemispheric dominance not only at the group level, but also at the individual-subject level. Here, we defined four (subjective) criteria for characterizing

robustness of right-hemispheric activation. These criteria assessed whether the paradigm activated a right-lateralized network both at the group level in a typical subject population (criterion a and b) and in a certain number of subjects at the single-subject level (criterion c and d):

a. At the group level, the paradigm had to induce brain activity in a fronto-parietal network typically associated with spatial processing (Fink et al., 2000; Jansen et al., 2005a) at a significance level $p < 0.001$, cluster threshold $k = 20$.
b. At the group level, the paradigm had to evoke right-hemispheric lateralization of brain activity, as indicated by a lateralization index $LI < -0.4$ (Wilke & Schmithorst, 2006), in core regions of the above-mentioned network (i.e., in frontal or parietal regions-of-interest (ROIs)).
c. At the single-subject level, the paradigm had to induce brain activity in the right-hemispheric frontal and parietal ROIs at the significance level $p < 0.001$ uncorrected, cluster threshold $k = 20$, in more than 30 % of all subjects.
d. At the single-subject level, at least 30% of the subjects had to show right-hemispheric lateralization of the brain activation pattern ($LI < -0.4$) in both the frontal and parietal ROI.

According to our previous experience with functional imaging tasks assessing spatial processing (e.g., Flöel et al., 2005; Jansen et al., 2004; Jansen et al., 2005a) we expected that the strength of activity (in terms of t-values) would be rather low at the individual subject level. We therefore opted at this point for a liberal whole-brain threshold for the single-subject analyses (i.e., 30%) to not exclude paradigms that might provide reliable but weak (in terms of t-values) measures of hemispheric dominance in some subjects.

Analysis 1: Brain activation at the group level: Brain activation patterns for each paradigm were first analyzed at the group level. One-sample t-tests were calculated separately for each paradigm, based on the contrast images from the single-subject analysis.

ROI masks definition: ROI masks were defined using an approach based on Jansen et al., (2006). As visuospatial processing is subserved by a large neurocognitive network, with different regions within this network showing different extents of lateralization, the LI is unlikely to reveal a consistent pattern of lateralization across these key regions of the brain. Therefore, constructing precise ROIs that capture the brain activity is of great importance. For this, two approaches, with their own advantages and disadvantages,

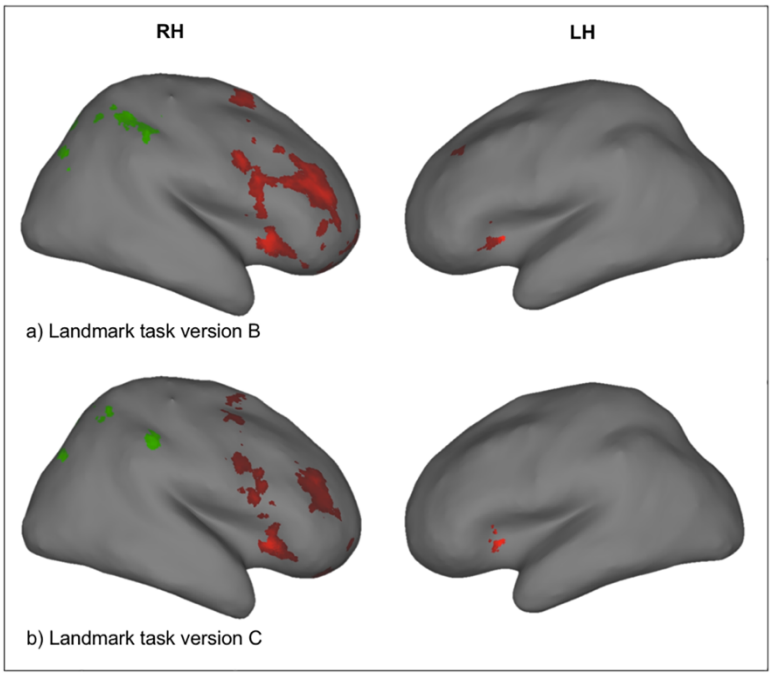
are commonly used: ROIs can be defined either functionally, based on the pattern of activation, or anatomically, based on anatomic knowledge. Obvious disadvantages of the latter are that pure anatomical definitions might include areas that are not engaged by the task or, vice versa, might exclude activation of interest that lies outside the chosen ROI (e.g., due to inaccuracies in the normalization procedure). Additionally, macroscopic landmarks are rarely reliable indicators of cytoarchitectonic borders (Amunts et al., 1999). On the other hand, functionally defined ROIs are also subject to limitations as they are typically derived from data of a pilot study investigating the functional activation in another group of subjects or within the same cohort with similar paradigms (Adcock et al., 2003; Fernández et al., 2003). Consequently, they might include regions outside the actual area of interest – thus, often necessitating additional masking.

To accommodate for the respective weaknesses of anatomically and functionally defined ROIs, we here applied a combined approach, which is based on both anatomical and functional constraints, for the definition of frontal and parietal ROIs as described in Adcock et al., (2003) and Jansen et al., (2004):

For the frontal ROI, group level activation patterns were thresholded as follows: "dots-in-space" and mental rotation task with $p < 0.001$ uncorrected, $k = 50$; Landmark task version A, B and C with $p < 0.01$ uncorrected, $k = 50$ (version A) and $k = 20$ (Version B and C), reflecting the functional constraint. We then applied anatomical constraints based on prior anatomical knowledge. Specifically, we used the frontal lobe mask as given by the WFU-Pickatlas as an inclusive mask to differentiate between regions of interest and regions of no interest. For each paradigm, we then created the frontal mask as a combination of the activated voxels surviving the defined thresholds within the anatomical landmarks."

For the parietal ROI, the same procedure as for the frontal ROI was used. For each paradigm, parietal ROIs were defined by creating masks of the group level activation pattern ("dots-in-space" and mental rotation task with $p < 0.001$ uncorrected, $k=50$; Landmark task version A with $p < 0.001$, $k=100$; Landmark task version B and C with $p < 0.01$, $k=20$) within the anatomical constraints of the parietal lobe as given by the WFU-Pickatlas. The specific thresholds were chosen to ensure that masks were roughly similar in size for the different paradigms. Note that, in study 2, the conjunction of the group level activation patterns of session 1 and 2 was used to define frontal and parietal ROIs (Fig. 8). In total, we created 10 masks – that is, a frontal and a parietal one for each paradigm.

Figure 8 *Landmark Task ROIs.*



Frontal (red) and parietal (green) ROIs are shown for Landmark task version B (a) and C (b). Frontal and parietal ROIs were defined by creating a mask, resembling the conjunction of the group level activation patterns of session 1 and 2 within the frontal lobe and parietal lobe, respectively, as given by the WFU-Pickatlas ($p < 0.01$ uncorrected, $k = 20$). RH = right hemisphere, LH = left hemisphere.

The degree of hemispheric lateralization was quantified by the lateralization index (LI), which is given by the formula

$$LI = \frac{A_L - A_R}{A_L + A_R},$$

where A_L and A_R refer to measures of fMRI-activity for equal ROIs within the left (L) and right (R) hemisphere, respectively. Several approaches have been established to calculate the LI (for a discussion, see Jansen et al., (2006)). We here applied the bootstrapping approach implemented in the SPM8 LI-toolbox, which is the current gold standard (Wilke & Schmithorst, 2006). The bootstrap approach uses 20 thresholding intervals with equally sized steps from 0 to the maximum t-value in the investigated region. At each threshold 100 bootstrap, resamples with a resample ratio of $k = 0.25$ were generated for each side from all the voxels in the investigated ROIs. From these resamples, all 10,000 possible LI combinations were calculated. A trimmed mean is then computed by only considering the central 50% of data points to exclude statistical or artefactual outliers and, thus, enhance the stability of the estimate. In a last step, a weighted mean LI is calculated by weighting the LIs with the respective thresholds, with higher thresholds receiving higher weights. The LI values range from −1 to +1. Positive values indicate a left-hemispheric dominance and negative values indicate a right-

hemispheric dominance. We masked out the midline (+/- 5 mm) to avoid flow artifacts in the large draining veins, as proposed by Wilke and Schmithorst (Wilke & Schmithorst, 2006). LIs for the group level BOLD patterns were calculated from the activation in the defined frontal and parietal masks for each paradigm.

Analysis 2: Brain activation in single subjects: The ROIs resulting from each paradigm's group analysis were used to investigate the activation strength in individual subjects. Therefore, single-subject activation maps were screened for activation in the respective ROIs at a fixed significance threshold ($p < 0.001$ uncorrected, cluster threshold $k = 20$). Individual LIs were also calculated in these ROIs, resulting in two indices (one frontal and one parietal LI) per subject for each paradigm.

3.2.4.2 Study 2: Test-retest reliability

In study 2, we assessed the test-retest reliability of both the activation patterns and the lateralization of the spatial processing network. Notably, we restricted our analyses to the Landmark task since this was the only task that fulfilled all criteria of robustness described above (see Results). First, we quantified the reliability of the activation patterns by computing intra-class correlation coefficients (ICCs) for each voxel using the ICC toolbox extension within SPM (Caceres et al., 2009). We then assessed the reliability of the brain lateralization of the spatial attention network. As a measure of the test-retest reliability of the *degree of lateralization*, we computed an ICC (two-way mixed model with absolute agreement using SPSS; IBM SPSS Statistics for Macintosh, version 22.0) for the LIs in the frontal and parietal ROI, respectively. As a measure of the test-retest reliability of *hemispheric dominance* (i.e., left, right), we determined the percentage of subjects in which categorical decision on the dominant hemisphere was consistent across measurements. Since the exact thresholds for partitioning left-dominance, right-dominance and bilateral activation are somewhat arbitrary, we repeated our analyses for three different specifications to account for this issue:

(i) *Three categories; left dominance for $LI>0.4$, right dominance for $LI < -0.4$, bilateral activation for $|LI| \leq 0.4$*

(ii) *Three categories; left dominance for $LI > 0.2$, right dominance for $LI < -0.2$, bilateral activation for $|LI| \leq 0.2$*

(iii) *Two categories; left dominance for $LI > 0$, right dominance for $LI \leq 0$*

3.3 Results

3.3.1 Study 1: Comparison of imaging paradigms

The "Dots-in-space" task: Behavioral data: Mean hit rate in the easy condition was 73.0 ± 14.4%, in the difficult condition 63.0 ± 17.0%, and in the control condition 98.5 ± 2.5%. A one-way ANOVA for repeated measures revealed an expected, significant main effect of difficulty ($p < 0.001$) across conditions. Imaging data: At the group level, brain activity for the linear contrast "easy + difficult > 2*control" ($p < 0.001$ uncorrected, $k = 20$) was found in a fronto-parietal network, in the anterior cingulate cortex and in the occipital lobe (Fig. 4a, Table 1). At the group level, brain activity in the parietal cortex was right-lateralized (LI = -0.42) and bilateral to right-lateralized in the frontal cortex (LI = -0.24). At the single-subject level, brain activity at $p < 0.001$ uncorrected, $k = 20$, was found in 15/15 subjects in the right frontal ROI and in the right parietal ROI. Brain activity was right-lateralized (LI < -0.4) for 7/15 subjects in the frontal ROI and for 6/15 subjects in the parietal ROI. However, only three subjects showed right-lateralized activation when looking at both ROIs simultaneously, whereas the remaining subjects showed bilateral LIs in either the frontal and/or parietal ROIs.

As suggested by one of our reviewers, we next performed correlation analyses between performance levels (i.e., hit rates) and measures of BOLD activation and hemispheric lateralization in order to address whether the observed inter-individual differences in brain activity and lateralization were related to the behavior (i.e., task performance) of individual subjects. We found a positive correlation between performance levels of the "dots in space" task and the number of activated voxels (at a statistical threshold of $p<0.001$ uncorrected, $k=20$, in the right frontal ROI (Spearman Rho correlation coefficient: $\rho = 0.528$ ($p = 0.043$) and $\rho = 0.559$ ($p = 0.030$) for the easy and hard condition of the task, respectively). On the contrary, no significant correlations were found for the right parietal ROI ($\rho = 0.455$ ($p = 0.088$) and $\rho = 0.361$ ($p = 0.186$) for the easy and hard condition, respectively). Additionally, we tested whether for a relation between task performance and lateralization strength of BOLD activations by computing correlations between the hit rates of the easy and hard condition and the LIs of frontal and parietal ROIs. However, no significant correlations between these variables were observed (see Supplementary Table S2).

Mental rotation task: Behavioral data: Mean hit rate in the easy condition was 97.2 ± 7.3%, in the difficult conditions 16.2 ± 10.2%, thus showing the expected effect of difficulty across conditions ($p < 0.001$), but also pointing to the extremely challenging nature of the activation condition. Subjects were essentially performing at chance with no significant difference from chance level (20%, $p = 0.168$). Note that, despite the low hit rate in the difficult condition, subjects reported to have actively engaged in the task rather than purely guessed the orientation of the cubes, as assessed by debriefing after the experiment. Imaging data: At the group level, brain activity for the linear contrast "high spatial processing load > low spatial processing load" ($p < 0.001$ uncorrected, $k = 20$) was found in a fronto-parietal network and in the anterior cingulate cortex (Fig. 4b, Table 1). However, in contrast to our expectations, the brain activity was left-lateralized in both the parietal (LI = 0.44) and the frontal cortex (LI = 0.60). At the single-subject level, brain activity at $p < 0.001$ uncorrected, $k = 20$, was found in 14/15 subjects in frontal and parietal ROIs in both hemispheres. Brain activity was right-lateralized (LI < -0.4) for only one subject in both the frontal and parietal ROI, but left-lateralized in both ROIs in nine subjects. We again tested for a brain-behavior interaction by computing correlations between the task performance and measures of brain activity. No significant correlation between performance levels (i.e., hit rates) of the task and the number of activated voxels (at a statistical threshold of $p<0.001$ uncorrected, $k=20$) in the right frontal ($\rho = 0.468$; $p = 0.092$) or right parietal ROI ($\rho = 0.257$; $p = 0.376$) were detected. Similarly, no significant correlations were observed between task performance and the LIs (see Supplementary Table S2).

Landmark task (version A): Behavioral data: Mean hit rate in the activation condition was 86.1 ± 15.9% and in the control condition 87.3 ± 17.8%. Hence, in contrast to the "dots-in-space" task and the mental rotation task, we did not observe a significant difference in the hit rate between activation and control condition ($p = 0.781$). Imaging data: At the group level, brain activity for the linear contrast "activation > control" ($p < 0.001$ uncorrected, $k = 20$) was found in a fronto-parietal network and the anterior cingulate cortex (Fig. 4c, Table 1). Brain activity was right-lateralized in the parietal cortex (LI = -0.63) and bilateral in the frontal cortex (LI = 0.03). At the single-subject level, brain activity at $p < 0.001$ uncorrected, $k = 20$, was found in 6/15 subjects in the right frontal ROI and in 7/15 subjects in the right parietal ROI. Furthermore, 5/15 subjects showed brain activity in both regions in the right hemisphere. Brain activity was right-lateralized (LI < -0.4) for 6/15 subjects in the frontal ROI (two subjects showed left-lateralized activation with LI > 0.4) and for 9/15 subjects in the parietal ROI (one subject showed left-lateralized activation with LI > 0.4). Five subjects displayed right-lateralized activation in both ROIs and one subject was left-lateralized in the frontal and parietal ROI. As for

the other two tasks, we assessed correlations between the task performance and measures of brain activity in order to test for brain-behavior interactions. However, no significant correlations between performance levels (i.e., hit rates) of the Landmark task and the number of activated voxels (at a statistical threshold of p<0.001 uncorrected, k=20) in the right frontal (ρ = -0.213; p = 0.686) or right parietal ROI (ρ = -0.273; p = 0.554) were detected. Similarly, no significant correlations were found between task performance and the LIs (see Supplementary Table S2).

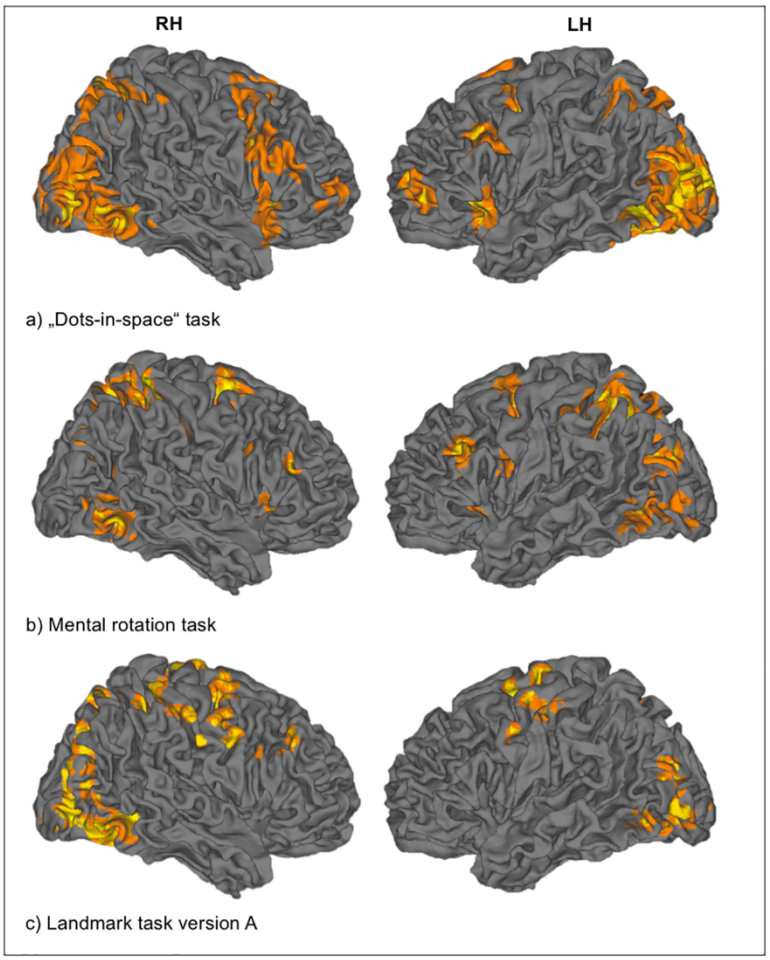


Figure 9 *Group activation patterns.*

Group activation patterns evoked by the “dots-in-space” task, the mental rotation task and the Landmark task version A (p < 0.001 uncorrected, k = 20). RH = right hemisphere, LH = left hemisphere. A detailed overview of the activated brain areas can be found in the Supplementary section (S1).

3.3.1.1 Evaluation of paradigms

All tasks showed, at the group level, brain activity in a fronto-parietal network related to spatial processing (criterion a). This is in line with the previous studies that had utilized the three tested paradigms (Cicek et al., 2009; Kinsbourne & Bruce, 1987; Whitehouse et al., 2009). This activity was right-lateralized for the Landmark task and the "dots-in-space" task, but not for the mental rotation task (criterion b). Given that we aimed to establish a paradigm that induces right-hemispheric lateralization, the latter task did not fulfill one important requirement. At the single-subject level, all three paradigms fulfilled the criterion of "robust activation in core-regions" (criterion c). For each task, we found activity in the frontal and parietal cortex (at least in the dominant hemisphere) at a specified threshold ($p < 0.001$ uncorrected). Furthermore, right-lateralized brain activity at the single-subject level in the frontal or parietal cortex, or both, was observed for the "dots-in-space" task and the Landmark task in a subset of the subjects. This suggests that both tasks are, in principle, suited for studying right-hemispheric lateralization. However, the Landmark task yielded slightly higher ratios and was the only paradigm that fulfilled criterion d (but note, this wouldn't have been the case when using a slightly more conservative threshold, e.g., > 40%). Furthermore, it was the most time-efficient paradigm in the sense that it provided equally (or even slightly more) robust results, while lasting only 5 minutes (as compared to 9 minutes for the "dots-in-space" task). Hence, we chose the Landmark task as the most promising paradigm for efficiently and robustly assessing right-hemispheric dominance related to spatial processing in single subjects. Note that this was also in line with the current view, suggesting the Landmark task to be the state-of-the-art paradigm for inducing right-hemispheric lateralization. A summary of the results can be found in Table 1.

Critically, although inducing right-hemispheric activation most consistently across the tested paradigms, the Landmark task in its present form showed somewhat weak overall activation strength (as illustrated by the fact that only 5/15 subjects showed significant activation in both ROIs at the pre-specified threshold). This represents a limiting factor when being interested in single-subject effects. For study 2, which addressed the test-retest-reliability of our selected (Landmark task) paradigm, we therefore aimed to induce stronger activations by optimizing the task and acquisition sequence (see Discussion).

Table 1. *Study participants' characteristics, study 1.*

Age, sex and the lateralization indies for the ”dots-in-space” (Dis), the mental rotation (MR) and the Landmark task version A (Lt). The additional column shows the number of voxel surviving the threshold (p<0.001 uncorrected, k=20) in the respective right-hemispheric ROI.

(m = male; f = female; LI = lateralization index; ROI = region of interest)

Subject ID	age	sex	LI Dis (frontal ROI)	active voxel (p<.001 unc.; k=20)	LI Dis (parietal ROI)	active voxel (p<.001 unc.; k=20)	LI MR (frontal ROI)	active voxel (p<.001 unc.; k=20)	LI MR (parietal ROI)	active voxel (p<.001 unc.; k=20)	LI Lt (frontal ROI)	active voxel (p<.001 unc.; k=20)	LI Lt (parietal ROI)	active voxel (p<.001 unc.; k=20)
1.1	24	f	0,10	1167	-0,35	1254	0,90	148	0,82	25	-0,60	455	-0,47	205
1.2	25	f	0,05	157	-0,37	346	0,68	443	0,89	-	0,59	-	-0,17	-
1.3	20	f	-0,39	1687	-0,53	895	0,67	442	-0,06	607	-0,38	20	0,28	-
1.4	24	m	-0,42	1300	-0,46	1403	0,49	1121	0,58	915	-0,46	61	-0,50	60
1.5	21	f	-0,16	1236	-0,12	2098	0,57	1280	0,65	935	-0,44	-	-0,73	-
1.6	26	f	-0,43	1171	-0,31	1916	0,25	187	0,49	969	0,17	-	-0,44	92
1.7	26	m	-0,57	1189	-0,30	902	0,57	925	0,58	690	0,45	-	0,55	-
1.8	25	m	0,34	62	-0,21	117	0,77	-	0,31	50	-0,41	-	0,22	-
1.9	23	f	-0,47	57	0,07	1105	0,66	796	0,55	801	-0,22	-	-0,49	61
1.10	32	m	-0,84	547	-0,71	348	-0,44	1228	-0,78	1032	-0,06	160	0,29	62
1.11	25	m	0,13	166	-0,41	514	0,66	948	0,70	926	0,00	-	-0,13	-
1.12	24	f	-0,31	2103	-0,48	2402	-0,07	808	-0,06	1066	-0,65	284	-0,54	420
1.13	24	f	-0,64	801	-0,11	532	0,57	907	0,53	763	-0,34	162	-0,56	363
1.14	26	m	0,23	161	-0,28	679	0,62	1487	0,51	1495	-0,39	-	-0,67	-
1.15	24	f	-0,55	170	-0,73	516	0,22	381	0,01	388	-0,63	-	-0,86	-

3.3.2 Study 2: Test-retest reliability

In study 2, we assessed the test-retest reliability for two adapted versions of the Landmark task (version B and C). Consistent with study 1, we here first describe the behavioral results, followed by the characteristics of the activation patterns at the group and single-subject level for both versions of the Landmark task (3.2.1). We then report the results of the reliability analyses (3.2.2).

3.3.2.1 Comparison of both versions of the Landmark task

Landmark task (version B): Behavioral data: Mean hit rate in the activation condition was 75.8 ± 10.9% in session 1 and 78.1 ± 9.1% in session 2. Mean hit rate in the control condition was 84.6 ± 12.3% in session 1 and 90.9 ± 8.4% in session 2. Hit rates in the activation and control conditions were significantly different in both sessions ($p < 0.001$), suggesting that the revised version of the Landmark task indeed induced performance

differences. Imaging data: At the group level, brain activity for the linear contrast "activation > control" was found in a fronto-parietal network and the anterior cingulate cortex (Fig. 10a, Table 2). At the group level, brain activity was right-lateralized both in the parietal cortex (session 1: LI = -0.88, session 2: LI = -0.86) and in the frontal cortex (session 1: LI = -0.85, session 2: LI = -0.90). At the single-subject level, brain activity at $p < 0.001$ uncorrected, $k = 20$, in the right frontal ROI was found in 13/20 subjects in session 1 and in 16/20 subjects in session 2. In the right parietal ROI, we found brain activity in 10/20 subjects in session 1 and in 10/20 subjects in session 2. For both session 1 and 2, 9/20 subjects showed brain activity in both regions. Brain activity was right-lateralized (LI < -0.4) for 16/20 subjects in session 1 and for 17/20 subjects in session 2 in the frontal ROI. For the parietal ROI, 14/20 subjects were right-lateralized for both session 1 and 2. 12 subjects displayed right-lateralized activation in both ROIs for both session 1 and 2, and only one subject was left-lateralized in the frontal and parietal ROI (see Table 2) in session 1. Hence, for this version of the Landmark task, right-hemispheric lateralization at the single-subject level could be detected in the majority of people in both ROIs (60%), and in almost all people in at least one of the regions (85% for session 2). This represents a notable improvement in robustness as compared with the initial version (A) of the Landmark task due to the optimized task and acquisition sequence.

Landmark task (version C): Behavioral data: Mean hit rate in the activation condition was 72.2 ± 8.5 % in session 1 and 76.2 ± 9.5 % in session 2. Mean hit rate in the control condition was 82.2 ± 23.1 % in session 1 and 90.6 ± 11.5 % in session 2. Activation and control conditions were significantly different in session 1 ($p = 0.048$) and session 2 ($p < 0.001$), again suggesting that the revised version induced performance differences. Imaging data: At the group level, brain activity for the contrast "activation > control" was found in a fronto-parietal network and the anterior cingulate cortex (Fig. 10b, Table 3). Brain activity was right-lateralized both in the parietal cortex (session 1: LI = -0.83, session 2: LI = -0.74) and in the frontal cortex (session 1: LI = -0.80, session 2: LI = -0.90). At the single-subject level, brain activity at $p < 0.001$ uncorrected, $k = 20$, in the right frontal ROI was found in 15/20 subjects in session 1 and in 13/20 subjects in session 2. For the parietal ROI, 8/20 subjects showed activation in session 1 and 10/20 subjects in session 2. Furthermore, 7/20 subjects showed brain activity in both ROIs in session 1 and 10/20 subjects in session 2. Brain activity in the frontal ROI was right-lateralized (LI < -0.4) for 15/20 subjects in session 1 and in session 2. For the parietal ROI, brain activity was right-lateralized for 10/20 subjects in both sessions. Additionally, 9/20 subjects displayed right-lateralized activation for both ROIs in session 1 and 8/20 subjects in

session 2. Only one subject was left-lateralized in the frontal and parietal ROI (see Table 3) in session 2.

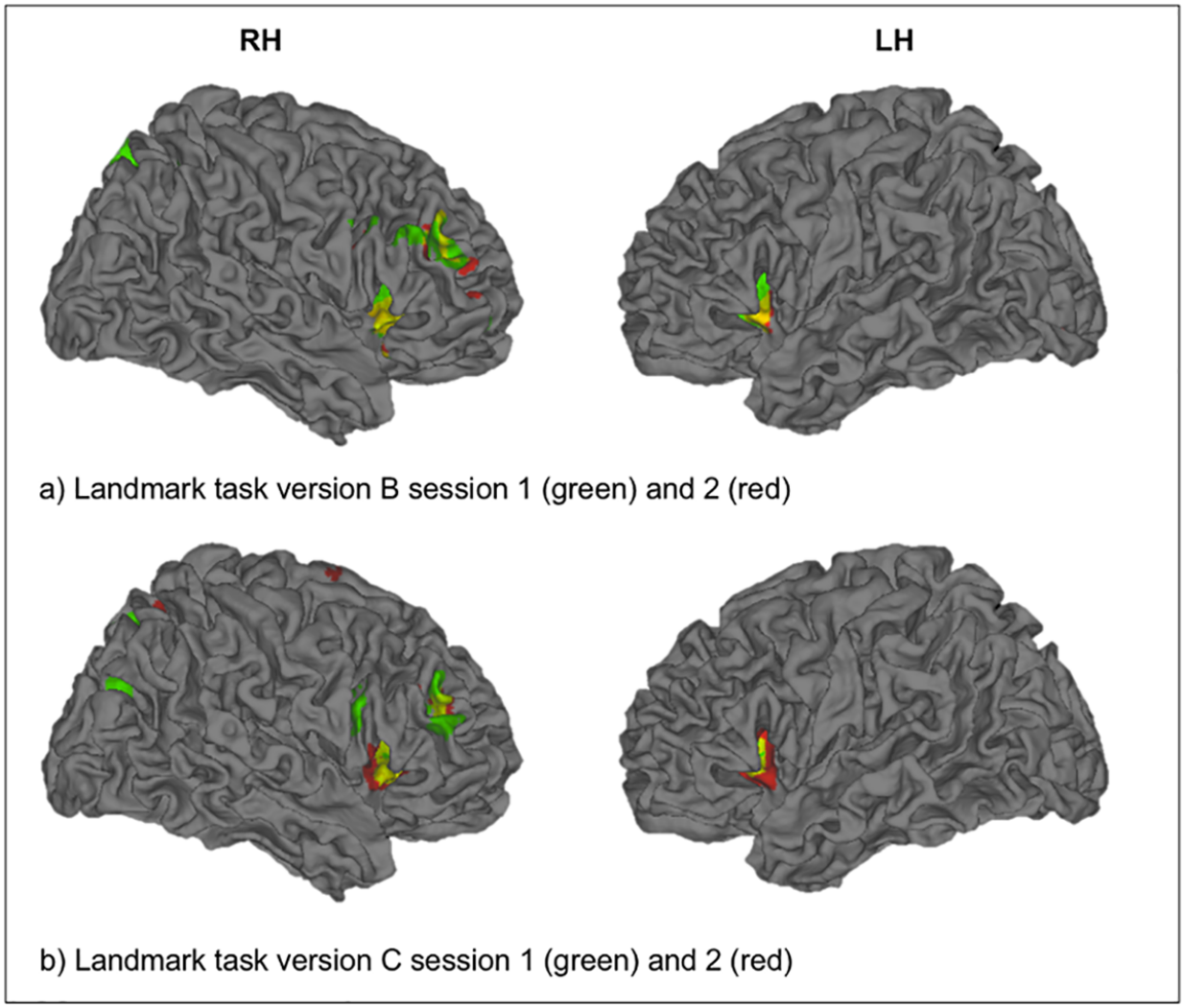


Figure 10 *Group activation patterns evoked by the Landmark task version B and C.*
Activation of the first (green) and second session (red) are plotted together ($p < 0.001$ uncorrected, $k = 20$) for version B (top) and version C (bottom), respectively. Yellow colored regions indicate voxels that were active in both sessions. RH = right hemisphere, LH = left hemisphere.

Table 2. *Study participants' characteristics, study 2 and lateralization indices for the Landmark Task version B (LtB).*

The additional column shows the number of active voxel surviving the threshold (p<0.001 uncorrected, k=20) in the respective right-hemispheric ROI.

Subject ID	age	sex	LI LtB (frontal ROI) Session I	LI LtB (frontal ROI) Session II	active voxel (p<.001 unc.; k=20) session I	active voxel (p<.001 unc.; k=20) session II	LI LtB (parietal ROI) Session I	LI LtB (parietal ROI) Session II	active voxel (p<.001 unc.; k=20) session I	active voxel (p<.001 unc.; k=20) session II
2.1	27	f	-0,89	-0,62	646	664	-0,49	-0,51	125	115
2.2	25	m	-0,64	-0,88	-	289	-0,84	-0,40	-	71
2.3	24	m	-0,31	-0,72	688	45	-0,69	-0,59	251	-
2.4	30	f	-0,85	-0,82	110	1465	-0,93	-0,74	-	228
2.5	25	m	-0,79	-0,68	257	-	-0,65	-	-	-
2.6	25	f	-0,61	-0,19	1156	76	-0,24	-0,81	244	-
2.7	24	f	-0,71	-0,75	399	1177	-0,50	-0,45	34	233
2.8	25	f	-0,85	-0,80	1266	-	-0,89	-	29	-
2.9	24	f	-0,71	-0,78	1227	512	-0,02	0,01	101	153
2.10	21	f	-0,80	-0,71	357	1263	-0,45	-0,13	33	179
2.11	25	f	-0,88	-0,72	-	449	-0,08	-0,64	-	30
2.12	22	m	-0,03	-0,64	-	-	-0,13	-0,63	-	29
2.13	29	m	-0,42	-0,48	1505	305	-0,42	-0,44	106	-
2.14	25	f	0,55	-0,02	52	23	0,50	-	-	-
2.15	23	m	-0,81	-0,53	54	46	-0,77	-0,54	-	-
2.16	27	m	-0,70	-0,81	192	412	-0,77	-0,94	26	-
2.17	26	m	-0,82	-0,85	-	-	-0,82	-0,91	-	-
2.18	27	f	-0,88	-0,16	-	105	-	-0,82	-	34
2.19	22	m	-0,22	-0,82	-	140	-0,69	-0,37	-	-
2.20	24	m	-0,83	-0,60	-	640	-0,55	-0,68	71	149

Table 3. *Lateralization indices for the Landmark task version C (LtC).*

The additional column shows the number of active voxels surviving the threshold (p<0.001 uncorrected, k=20) in the respective right-hemispheric ROI.

Subject ID	LtC (frontal ROI) Session I	LtC (frontal ROI) Session II	active voxel (p<.001 unc.; k=20) Session I	active voxel (p<.001 unc.; k=20) Session II	LtC (parietal ROI) Session I	LtC (parietal ROI) Session II	active voxel (p<.001 unc.; k=20) session I	active voxel (p<.001 unc.; k=20) session II
2.1	-0,77	-0,76	61	503	-0,19	-0,84	46	120
2.2	-0,84	-0,91	25	-	-0,35	-0,06	-	-
2.3	-0,40	-0,73	-	163	-0,82	-0,65	33	-
2.4	-0,82	-0,73	1132	373	-0,69	-0,1	198	110
2.5	-0,82	-0,58	32	-	-0,69	-0,35	-	-
2.6	-0,59	-0,04	178	96	-0,88	-0,36	-	-
2.7	-0,75	-0,75	25	529	-0,21	-0,26	-	72
2.8	-0,87	-0,77	290	655	-0,99	-0,68	-	177
2.9	-0,69	-0,80	580	626	0,03	-0,03	295	98
2.10	-0,59	-0,79	479	65	0,26	-0,55	72	-
2.11	-0,55	-0,76	183	-	-0,58	-0,49	62	-
2.12	-0,48	-0,08	-	-	-0,92	0,50	-	-
2.13	-0,34	-0,17	334	268	-0,22	-0,69	-	133
2.14	0,60	0,40	58	876	0,71	0,64	-	85
2.15	-0,67	-0,52	23	-	-0,73	-	-	-
2.16	-0,28	-0,69	-	-	-	-0,64	-	-
2.17	-0,85	-0,76	-	506	-0,29	-0,85	-	88
2.18	0,01	-0,41	1370	584	-0,56	-0,58	322	101
2.19	-0,32	-0,67	453	-	0,07	-0,1	72	-
2.20	-0,51	0,05	-	40	-0,41	-0,81	-	52

3.3.2.2 Test-Retest Reliability

Test Retest Reliability was assessed in a four-step procedure: While the first two analyses steps were performed with regard to the whole brain BOLD activation patterns, the third and fourth analysis step concerned the test-retest reliability of the obtained LIs:

1) Qualitative analysis of the group activation overlap
2) Test-retest reliability of voxelwise activation strength across subjects
3) Test-retest reliability of the degree of lateralization
4) Analysis of consistency of the categorical classification of hemispheric dominance

We first provide a qualitative analysis of consistency of activation patterns across the different measurements by inspecting the overlap of brain activation patterns at the group level. These overlaps show that for both versions (i.e., B and C) of the Landmark task, a comparable frontal network was activated in session 1 and session 2 (Fig. 10).
Second, with regard to the test-retest reliability of the activation patterns, we computed ICCs for each voxel using the ICC toolbox extension within SPM (Caceres et al., 2009). Whole-brain joint probability distributions showed an association between t-values and ICCs (Fig. 11). According to established conventions, we classified test-retest reliability as "poor" for ICC ≤ 0.4, "fair" for 0.4 < ICC ≤ 0.6, "good" for 0.6 < ICC ≤ 0.8, and "excellent" for ICC > 0.8 (Caceres et al., 2009; Fliessbach et al., 2010). For both versions of the paradigm, ICCs were higher for voxels showing strong activation (high t-values) or "deactivation" (high t-values for the opposite contrast). For the Landmark task version B, the median ICC for the whole brain was 0.21, for the activated network 0.21, for the frontal ROI 0.30, and for the parietal ROI 0.23 (Fig. 11a, b). For the Landmark task version C, the median ICC for the whole brain was 0.22, for the activated network 0.29, for the frontal ROI 0.31, and for the parietal ROI 0.29 (Fig. 11c, d). All these values indicate poor reliability, suggesting that the BOLD signal in individual voxels was not very reliable.

Notably, however, the main goal of study 2 was to determine whether measures of hemispheric lateralization (i.e., degree of hemispheric lateralization and the categorical classification of hemispheric dominance) were reliable across multiple sessions. In a third step, we therefore focused on these quantities. With regard to the test-retest reliability of the degree of lateralization, the average measures for ICCs for the LIs in the frontal (ICC = 0.57) and in parietal ROI (ICC = 0.54) suggested fair reliability of this measure in Landmark task version B. In Landmark task version C, average measures for ICCs were excellent (ICC = 0.81) and poor to fair (ICC = 0.33) in the frontal and parietal ROI, respectively.

In a fourth step, we tested the consistency of the categorical classification of hemispheric dominance across both sessions. When using a conservative threshold of |LI| > 0.40, 70.0% of subjects showed consistently right-hemispheric dominance in the frontal and 62.5% in the parietal ROI (version B). For version C, these percentages were reduced, with 60.0% and 27.8% of subjects showing consistently right-hemispheric dominance in the frontal and parietal ROI, respectively. As mentioned above, the chosen threshold of |LI| > 0.40 is rather conservative and, in fact, various studies in the literature have used more liberal thresholds for identifying left- or right-hemispheric lateralization (Deppe et al., 2000; Jansen et al., 2004). In a more exploratory analysis, we used a more liberal criterion for hemispheric dominance (i.e., |LI| > 0.2) and found that 80.0% of subject showed right-hemispheric dominance in the frontal and 75.0% in the parietal ROI (version B). Again, these values were reduced for the version C with right-hemispheric dominance in the frontal ROI (70.0%) and parietal ROI (55.6%). When further softening the criterion for hemispheric dominance by considering only two categories (i.e., left, right), 95.0% of subjects showed right-hemispheric dominance in the frontal and 93.8% in the parietal ROI across both sessions in Landmark task version B. This percentage was marginally smaller (85.0 % in the frontal, 72.2% in the parietal ROI) in Landmark task version C. These results again highlight to utility of the revised versions of the Landmark task for studying right-hemispheric lateralization using fMRI – at least for the liberal scenario when making binary decisions between left- or right-hemispheric dominance of the brain activation pattern.

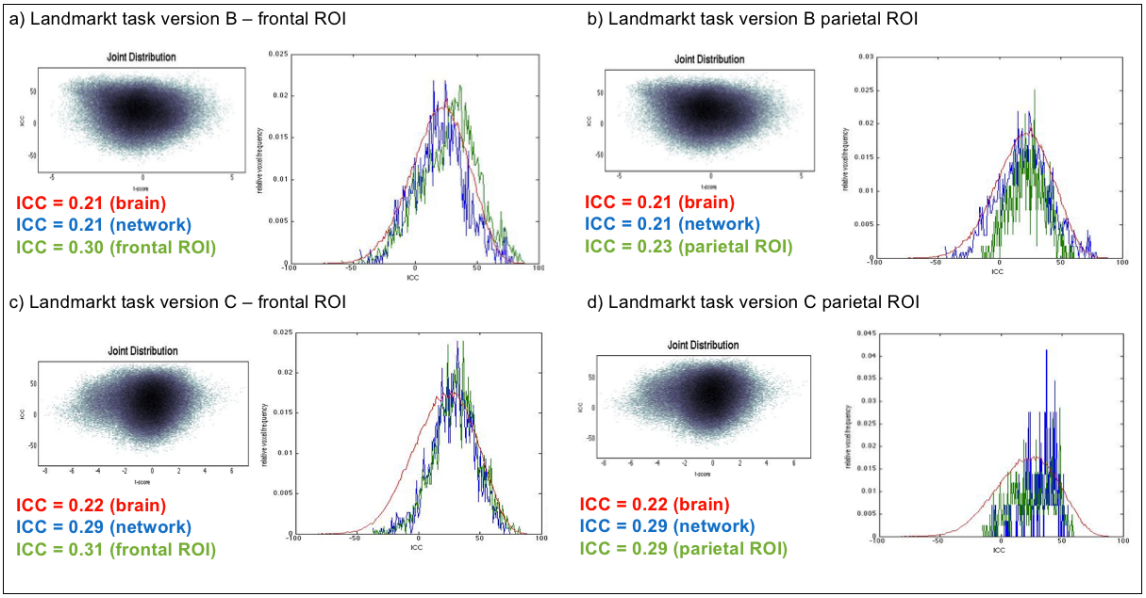


Figure 11 *ICC Results.*

Left (top and bottom): Joint probability distribution of voxel-wise t-values and associated median ICC values. Right (top and bottom): ICC frequency distributions for the whole brain (red), for the

voxels in the activated network (blue) and for the defined frontal and parietal ROI (green). The "activated network" was defined based on the results from the first measurement. Voxels were classified as active if they had t-values t>3.5794 (corresponding to $p < 0.001$ uncorrected, k=20). Both diagrams are presented for the Landmark tasks version B (a, b) and C (c, d), for the frontal and the parietal ROI, respectively.

3.4 Discussion

In the present study, we aimed to establish an fMRI paradigm that *robustly* and *reliably* evokes right-hemispheric lateralization of brain activity in the human brain. To this end, we first compared the suitability of three different paradigms on visuospatial processing: (i) the "dots-in-space" task (adapted from Whitehouse et al., (2009)), (ii) mental rotation task (Dorst et al., 2008), and (iii) Landmark task (Fink et al., 2000). These tasks had been used frequently in imaging studies to assess right-hemispheric lateralization in the human brain. In study 1, we evaluated the utility of these different paradigms to induce right-hemispheric lateralization in the spatial processing network. It is worth highlighting that, despite the fact that it is generally accepted that visuospatial processing induces right-hemispheric lateralization in the majority of subjects, studies have also reported notable variability across the population. Specifically, atypical (left-hemispheric) lateralization of visuospatial attention has also been described in a non-negligible proportion of subjects. Such inter-individual variability has not only been observed for visuospatial attention, but for any lateralized cognitive function, including language (Frost, 1999; Jansen et al., 2007; Knecht, Deppe, et al., 2000; Knecht, Dräger, et al., 2000; Pujol et al., 1999; Springer et al., 1999). Hence, one would not expect that the tested visuospatial processing paradigms evoke right-hemispheric lateralization in each and every subject. This makes it somewhat challenging to devise a principled criterion to assess the "robustness" of the above-mentioned paradigms at the single-subject level. We have tried to accommodate for this inherent inter-individual variability by using rather liberal "robustness" criteria at the single-subject level in study 1.

Under these (subjective) criteria, we compared the "robustness" of the different paradigms. First, the mental rotation task failed, as it revealed a left-hemispheric lateralization of the activation pattern both at the group level, as well as in most of the individual subjects. Second, the "dots-in-space" task and the Landmark task performed more or less similar. More precisely, the Landmark task yielded slightly higher ratios of right-hemispheric dominance at the single-subject level than the "dots-in-space" task and passed all four criteria (but note that this would have not been the case for slightly more

conservative criteria c and d). In an additional analysis, we asked whether the marked inter-subject variability observed in the brain activation patterns during all tasks was related to the behavior of individual subjects. While we found a significant correlation between performance levels (i.e., hit rates) and the number of activated voxels in the right frontal ROI for the “dots-in-space” task (at an uncorrected threshold not accounting for multiple comparisons), no other correlation reached significance. This suggests that a potential link between behavioral performance and the measures of brain activity tested in the present study was rather weak and thus could not sufficiently explain the variability in right-hemispheric lateralization observed across subjects. These differences are thus likely to relate to other factors such as distinct cognitive strategies (Dunlosky & Kane, 2007), inter-subject variability in the BOLD signal (Harrington et al., 2006; Salli et al., 2001; Tjandra et al., 2005), or variability in neuroanatomy (Juch et al., 2005).

Overall, we focused on the Landmark task (instead of the “dots-in-space” task) for study 2 because of its slight benefit in performance and the considerably higher efficiency (i.e., the Landmark task provided similarly robust results while taking only half the scan time). However, one has to carefully consider the limitations of the choice of fMRI paradigms for study 1. Here, we compared three commonly used paradigms, in a form in which they already existed in the literature (Landmark task and mental rotation task) or by closely adapting a paradigm designed for fTCD (“dots-in-space” task) for the use in fMRI. All three paradigms are however subject to several limitations (e.g., low performance in the mental rotation task, suboptimal implementation of the Landmark task in study 1), which might have confounded our selection of the most suitable and robust paradigm for study 2. In what follows, we highlight some of these issues to make the reader aware of the limitations of the present study.

With regard to the mental rotation task, subjects showed strikingly low behavioral performance levels during the activation condition, which aggravates the interpretation of the hemispheric lateralization results. The left-lateralized brain activity in the frontal and parietal ROIs is in contrast to the right-hemispheric lateralization reported by Dorst et al. (Dorst et al., 2008) and the bilateral patterns observed by Hattemer and colleagues (Hattemer et al., 2011). Several explanations might account for these differences. First, given the above-mentioned poor performance levels, we cannot exclude the possibility that subjects did not correctly engage in the mental rotation task (despite the fact that they reported active participation during the debriefing). To address this confound, forthcoming studies testing the utility of the mental rotation task should strive for a more graceful difficulty level to ensure adequate behavioral performance. Another potential source of variation between the present study and Dorst et al. (Dorst et al., 2008) lies in the imaging method. While the present study used fMRI, Dorst and colleagues used

fTCD. Notably, fTCD is sensitive primarily to major cerebral arteries, most commonly the middle cerebral arteries (MCA), which are insonated through the transtemporal window (Bleton et al., 2016). The MCA mainly supplies the lateral surface of the hemisphere with the exception of the superior parietal lobe, the inferior temporal lobe and the occipital lobe. Consequently, fTCD has relatively poor spatial resolution because signals can only be captured from brain regions that are supplied by the MCA. This results in rather crude estimates of hemispheric lateralization. Additionally, while fTCD might be suitable for studying brain areas completely supplied by the MCA or widely distributed network, brain regions lying outside the supply area of the MCA will not be captured at all. Given these marked differences between fTCD and fMRI, differences in the hemispheric lateralization results are to be expected to some degree – especially at the single-subject level where we observed high variability of the fronto- and parietal activation patterns.

An additional limitation of the present study is the small difference in the MRI acquisition parameters across the three paradigms (i.e., dots-in-space, mental rotation, and Landmark task). This might have resulted in somewhat different SNRs of the acquired data. Similarly, the differences in cognitive and behavioral characteristics of the tasks themselves (see Introduction) could have also biased our endeavor for identifying the most robust paradigm of right-hemispheric lateralization during visuospatial processing. Finally, active and control conditions for each of the three visuospatial processing tasks differed in behavioral demands, as well as basic sensory and motor aspects, leading to differences between the two conditions arguably unrelated to the mechanisms underlying hemispheric lateralization.

Overall, these limitations suggest that future refinements of the utilized paradigms are needed for a more thorough investigation of the right-hemispheric lateralization in visuospatial attention. Having said this, the present study still makes a valuable contribution because it provides the (to our knowledge) first comparison of the robustness of currently established paradigms for assessing right-hemispheric dominance during visuospatial processing. As such, the present findings could serve as a benchmark against which future developments aiming to improve the robustness and reliability of “state-of-the-art” paradigms can be compared.

In study 2, we investigated the test-retest reliability of two (optimized) versions of the Landmark task. Test-retest reliability is an important test-theoretical property because, in fMRI, scientists are often confronted with poor SNR at the single-subject level (Bennett & Miller, 2010) while, at the same time, being interested in inter-individual differences (e.g., in hemispheric lateralization) since they might offer deeper insights into the mechanisms underlying a cognitive task. Here, we find the reliability of the brain activation strength in single voxels to be poor, as indicated by low ICCs. This speaks to

the above-mentioned poor SNR levels of the BOLD signal when only looking at single voxels individually. Having said this, the hemispheric dominance and the degree of hemispheric lateralization (as measured with the LI), which depicted the most important criteria in the present study, could be identified with reasonably high reliability across the two sessions. Specifically, we showed that for a binary classification of hemispheric dominance (i.e., left vs. right), right-hemispheric lateralization could be reliably detected across both sessions in the vast majority of subjects for the Landmark task version B (i.e., > 90 % in the frontal and parietal ROI). Additionally, we found fair to good ICCs (ICC > 0.5) both in frontal and parietal ROIs for this version of the task. For version C, we report slightly lower reliabilities for the categorical classification of hemispheric dominance and the degree of hemispheric lateralization. We would like to stress out, that is a well-established procedure to study reliability when restricting the analysis to the activated brain network (Adcock et al., 2003; Jansen et al., 2004). In this sense, our results can be interpreted as an upper bound on the reliability of the Landmark task, against which future studies that aim to refine and improve fMRI paradigms on visuospatial attention can be compared. Hence, our results provide evidence that the Landmark task version B should be preferred to reliably characterize hemispheric dominance.

In summary, comparing established paradigms for studying brain activity related to visuospatial processing, our results suggest that the Landmark task is best suited to assess right-hemispheric dominance of brain activation patterns and should be considered as the current method of choice for studying the lateralization of visuospatial attention with fMRI. This is supported by the reasonably high test-retest reliability of the LIs and the high consistency of hemispheric dominance classification across multiple sessions. Notably, these results are also in line with the current view that the Landmark task represents the state-of-the-art paradigm for investigating visuospatial functions, and is widely used both in clinical practice (Fink et al., 2000) and neuroscientific studies on hemispheric specialization (Cai et al., 2013). Our results therefore suggest that the Landmark task, especially version B, robustly and reliably determines hemispheric dominance for visuospatial processing.

Establishing robust and reliable imaging paradigms to study hemispheric lateralization at the single-subject level will hopefully enable a more thorough understanding of the putative mechanisms (Beaumont, 1997). In particular, developing precise models that capture which factors drive hemispheric specialization in individual subjects, how lateralization processes of different cognitive functions interact with each other, and how the brain integrates processes that are lateralized to opposite hemispheres. The relevance of such interactions among lateralized processes has been suggested for

language and spatial attention (Jansen et al., 2005a; Lidzba et al., 2006; Staudt et al., 2002), for language and working memory (Axmacher et al., 2009), and for face perception and handedness (Frässle et al., 2016). While moving in the right direction, these studies have not yet provided a principled and systematic investigation of the interactions among lateralized cognitive functions.

Along these lines, it is worth highlighting that the present study was part of a larger project that aims to establish a test battery for studying hemispheric lateralization of several cognitive functions (e.g., language, visuospatial processing, face processing and memory) and their interactions using fMRI. Specifically, this project strives to map how lateralized processes within one hemisphere interact (and compete) with each other, as well as the mechanisms by which the brain integrates processes lateralized to opposite hemispheres. Importantly, to ensure sensible inference on these mechanisms, the utilized paradigms need to adequately and reliably take into account the inter-individual variability in the degree of hemispheric lateralization mentioned above. The present study speaks to this endeavor of identifying such paradigms with respect to the right-hemispheric lateralization of visuospatial processing.

In summary, we here demonstrated the utility of the Landmark task for mapping right-hemispheric lateralization, but at the same time highlight current limitations of the paradigm. Further improvements of the present paradigm, as well as the use of more sophisticated scanner hardware and/or acquisition sequences (e.g., high field MRI, multiband EPI techniques), should constitute major endeavors of forthcoming studies in order to boost the sensitivity and stability of single-subject measures of hemispheric lateralization in the human brain.

4

Project Two

Shedding Light in Regional Grey Matter Differences in Women – How do Menstrual Cycle Phases and the Intake of 'The Pill' shape the Female Brain?

Within the fourth chapter the second project is described. Here, the effect of two opposed menstrual cycle phases, notably early follicular phase/menstruation and mid luteal phase, and OC treatment during pill-intake and pill-withdrawal on brain GM volume is investigated. Additionally, the impact of the application of two commonly used processing pipelines is investigated.

4 PROJECT TWO:

SHEDDING LIGHT ON REGIONAL GRAY MATTER DIFFERENCES IN WOMEN – HOW DO MENSTRUAL CYCLE PHASES AND THE INTAKE OF 'THE PILL' SHAPE THE FEMALE BRAIN?

Schuster V, Herholz P, García Alanis JC, Schröder C, Sommer J, Jansen A.

Conceptualization:	Verena Schuster, Andreas Jansen.
Data curation:	Verena Schuster, Peer Herholz, Chantal Schröder.
Formal analysis:	Verena Schuster, Peer Herholz, José C. García Alanis, Jens Sommer.
Funding acquisition	Verena Schuster, Andreas Jansen.
Investigation:	Verena Schuster, Peer Herholz.
Methodology:	Verena Schuster, Andreas Jansen.
Project administration:	Verena Schuster.
Resources:	Verena Schuster, Andreas Jansen.
Software:	Verena Schuster, Peer Herholz, José C. García Alanis, Andreas Jansen, Jens Sommer.
Supervision:	Andreas Jansen.
Validation:	Verena Schuster, Andreas Jansen.
Visualization:	Verena Schuster, José C. Garcia Alanís, Peer Herholz.
Writing – original draft:	Verena Schuster.
Writing – review & editing:	Verena Schuster, Peer Herholz, José C. Garcia Alanís, Andreas Jansen.

4.1 Introduction

Effects of sex differences on brain structure and function have become an intensive object of research since the improvement of modern imaging techniques like MRI. Subsequently, structural and behavioral differences between males and females have been extensively investigated in healthy men and women, showing many similarities but also distinct characteristics that distinguish the male from the female brain.

In neuropathological and MRI studies, it has been consistently shown that brains of men and women significantly differ in absolute brain size (Allen et al., 2003; Cosgrove et al., 2007; Davison Ankney, 1992; Shin et al., 2005), with men showing greater brain volume than women. However, when controlling for total volume, women present a higher percentage of gray matter and men for white matter (Allen et al., 2003; Goldstein et al., 2001; Gur et al., 1999). Thus, researches uniformly agree that total brain volume is approximately 10% larger in men than in women (Allen et al., 2003; Andreasen et al., 1993; Goldstein et al., 2001; Good et al., 2001; Nopoulos et al., 2000; Paus, 2010).

Nonetheless, the obtained results have also been accompanied by inconsistent conclusions and various neuroanatomical features that differ. For example, women showed larger hippocampal and caudate volumes (Filipek et al., 1994) in frontal and medial paralimbic cortices (Goldstein et al., 2001). In contrast, men showed larger volumes in the medial frontal cortex, the amygdala, and the hypothalamus (Goldstein et al., 2001).

Following, it remains unclear whether the observed differences are characteristic for women (or men, respectively), or whether they are just typical for larger or smaller brains (Leonard et al., 2008). In order to address this pertinent question, one has to consider fundamental physiological differences, which may be responsible for the detected sex-specific characteristics. For example, the global cerebral blood flow alters between women in men (Barnes, 2017) and the concentration and composition of neurotransmitters (Staley et al., 2001), indicating that male and female brains are neurochemically unique. This underlying physiological diversity might affect brain structure and, thus, cause the outlined differences between male and female brain anatomy.

However, one fundamental key player, precipitating complete peripheric sex differences, and therefore a potential candidate for influencing brain structure is still missing: Sex hormones. Particularly, the entirely different and highly varying sex hormone concentration in women: women offer high fluctuations of their hormonal state, not only physiologically over their entire lifespan (e.g., puberty, pregnancy, menopause) but also

within concise periods, in particular during the menstrual cycle or by suppressing these endogenous hormones by using hormonal contraceptives, for example by taking oral contraceptives. The latter two described hormonal states accompany women during their whole reproductive state, causing substantial divergence in hormonal occurrence and are often neglected when performing scientific research, including women.

However, focusing on the menstrual cycle, complex endocrine principles, governed by various regulation and feedback mechanisms involving hypothalamic, pituitary, and ovary action, occur, resulting in large fluctuations of female sex hormones, inter alia progesterone and estradiol (Mihm et al., 2010). Simplified, the menstrual cycle can be divided into two different sections: the follicular phase, primarily governed by estradiol concentrations, begins on the first day of menstrual bleeding and lasts to ovulation around cycle day 14, and the luteal phase, predominantly controlled by progesterone, ranges from ovulation to the end of the cycle around cycle day 28 (e.g., one day before the first day of menstrual bleeding (Mihm et al., 2010).

Both estradiol and progesterone, in turn, are well-known to additionally offer strong neuromodulatory properties (Arélin et al., 2015; Barth et al., 2015; McEwen & Woolley, 1994). Thus, besides immense physical changes concerning the whole body, the menstrual cycle strongly affects behavioral alterations (e.g., in food intake (Arnoni-Bauer et al., 2017; Kammoun et al., 2017) and sleep quality (Van Reen & Kiesner, 2016)), functional (for a review see Sundström Poromaa & Gingnell, (2014)) and brain structural aspects, with regard to gray and white matter (for example see Luders et al., (2009). Previous behavioral and neuroimaging studies have particularly shown effects of the menstrual cycle on cognitive functions, for example, visuospatial ability (Hausmann, 2005), verbal skills (Maki et al., 2002; Rumberg et al., 2010), and emotional memory (Bayer et al., 2014).

With the advent of oral hormonal contraceptives, notably *the pill* in the 1960s, women were firstly enabled to control their hormonal states. Nowadays, OCs, particularly combined preparations containing an estrogen and a progestin, are used by over 100 million women worldwide (Petitti, 2003). Whereas metabolic side effects of OC use have already been investigated (Wang et al., 2016; Wiegratz & Thaler, 2011), it is all the more surprising, especially concerning a large number of worldwide users, that the effect of OCs on brain function and thus resulting behavior and brain structure is a relatively new research area. Only a handful of studies so far considered and investigated the consequences of the use of OCs, on brain function (Bonenberger et al., 2013; Marecková et al., 2014; Petersen et al., 2014; Pletzer & Kerschbaum, 2014; Rumberg et al., 2010).

In summary, previous studies emphasized the influence of endogenous sex hormones' fluctuations as they occur physiologically during the menstrual cycle as well as the intake

of ‘the pill’ on behavior and various cognitive functions. Additionally, recent research has shown that these different hormonal states affect brain structure, particularly GM volumes, as well. Previous results are summarized in the following section.

Influence of naturally fluctuating hormones on brain structure:

It has already been shown that natural fluctuations of female sex hormones during the menstrual cycle lead to changes with regard to brain structure. Initial findings of human brain plasticity associated with endogenous menstrual cycle were published by Protopopescu et al., (2008). They found increased gray matter in the right anterior hippocampus and decreased gray matter in the right dorsal basal ganglia during the late follicular phase compared to the late luteal phase. In 2010 Pletzer and colleagues reported larger volumes in the right fusiform/parahippocampal gyrus during the early follicular phase compared to the midluteal cycle phase (Pletzer et al., 2010).
Lisofsky et al. (2015) investigated the effect of the menstrual cycle on hippocampal volume and found increased hippocampal volumes in the late follicular phase relative to the early follicular phase. However, no changes were obtained between the follicular and the luteal phase. Pletzer and colleagues investigated the effect of three different cycle phases, namely the early follicular, the pre-ovulatory, and the mid-cycle phase (Pletzer et al., 2018) on GM volume. Using a large sample of 55 subjects, each measured three times to assess potential menstrual cycle-dependent changes with sufficient power, they detected larger GM volumes during the pre-ovulatory phase in the left and right hippocampus and within the right basal ganglia during mid-cycle. They could not find any other brain area being affected by menstrual cycle changes.

Influence of OCs on brain structure:

In 2010 Pletzer and colleagues published a first exploratory study investigating the influence of OC intake on gray matter (Pletzer et al., 2010). They reported larger gray matter volumes in several brain areas (e.g., prefrontal cortex, anterior cingulate cortex, parahippocampal and fusiform gyri, cerebellum) in women taking OCs. However, neither the type (notably, “generation”) nor the exact chemical combination and hormone concentration of estrogens and progestins of the OC was taken into account. That these issues hold a strong influence with regard to the brain structure, mainly gray matter volumes could already be confirmed in a follow-up study: In 2015, Pletzer and colleagues investigated the consequences of the use of so-called androgenic and anti-androgenic OCs, referring to their receptor binding properties and thus their ability to stimulate male

characteristics (Bitzer & Simon, 2011), resulting in opposed effects on brain structure (Pletzer et al., 2015a). Whereas anti-androgenic OCs lead to larger gray matter volumes compared to women with a natural cycle, users of androgenic OCs displayed partly smaller brain regions in specific brain areas. However, the authors did not control for the exact hormone derivatives in the combined preparations.

Inconsistencies and heterogeneity of results between these and further studies (Bondt et al., 2013a; Hagemann et al., 2011) might be attributable to (1) different study designs and (2) different analyzing methods:

Regarding the first point, sample sizes were often small, and the examined cycle phases in women with a natural cycle were often designated on the participants' oral statements and not verified by reviewing hormone concentrations. Additionally, the exact hormonal combination and type of hormonal OC were not taken into account, which, however, do exert a profound influence on brain structure (Pletzer et al., 2015a). Even though these studies have controlled for androgenicity levels, they did not regulate the exact combination of estrogen- and progestogen components and condensed women using different anti-androgenic progestins and other androgenic progestins with either strong or weak androgenic properties (Bondt et al., 2016; Pletzer et al., 2015a).

Focusing on the second point, different data analysis pipelines for structural images and furthermore, different brain parcellations were used, reducing the comparability of various studies' results.

Until now, the obtained heterogeneous and partly inconsistent results of the effects of either naturally fluctuating female sex hormones or exogenously applied hormones might be caused by the above-described reasons.

Aim of our study:

Bearing the above described two methodological weaknesses previous studies were confronted with in mind, the present study aimed to further shed light on this rapidly developing research area by investigating the effect of the menstrual cycle and OC intake on brain structure in a strictly controlled cohort of subjects using two commonly applied analysis pipelines.

In particular, the effect of menstrual cycle was explored during two opposed cycle phases with the highest differences in progesterone levels (low-hormonal levels during menstruation synonymous with the early follicular phase and high hormonal levels during the midluteal phase) and the effect of one defined hormonal OC was examined during the pill-free week and pill-intake.

For women with a spontaneous menstrual cycle, the oral designation of the cycle phase, as it was performed in previous studies ,e.g., Pletzer et al. (2010), was not sufficient. Yielded results from these studies might not be conclusive, as it cannot be assured that data was collected within the striven cycle phases. Thus, the control of the participants' menstrual cycles for 3-4 months previous to the first MRI appointment using basal temperature tracking and ovulation tests was an essential requirement of the described study. At the time of appointment, additional blood samples in order to analyze progesterone and estradiol concentrations and thus to verify that measurements were performed during the envisaged cycle phases were collected.

For women included in the OC group, only defined oral contraceptives with similar hormonal preparation were allowed. Here the combined OC containing ethinylestradiol (20 µg) as an estrogen component and levonorgestrel (100 µg) as a progestogen compound was the medication of choice. This combination is considered one of the first-line choices and the 'gold standard', as they show a relatively low risk of venous thromboembolism (Stewart & Black, 2015). With regard to the OC's androgenicity, a combination containing levonorgestrel as a progestogen analogue is defined to have androgenic effects. Taking the precise combination of the investigated OC into account was from high importance, as previous studies have already shown, that different androgenicity characteristics of OCs influence their effect on the brain structure (De Bondt et al., 2016; Pletzer et al., 2015a, 2019a).

Furthermore, within this study, we were interested in the effect of the application of different analysis streams on yielded potential regional GM differences. Therefore, two state-of-the-art methods investigating the impact of the menstrual cycle and oral contraceptive on GM volume were performed. GM volume is widely analyzed using voxel-based morphometry (VBM). This approach implements the voxel-wise estimation of the local amount, or the local volume of specific brain tissue (eg., GM, white matter (WM), cerebral spinal fluid (CSF)). In 2016 Gaser and colleagues introduced the new CAT12 toolbox (http://www.neuro.uni-jena.de/hbm2016/GaserHBM2016.pdf) as an extension for VBM analysis, allowing various analysis of the brain structure. Since then, the toolbox has been extended and continuously improved and now provides multiple morphometric methods, notably region- or label-based morphometry.

Another frequently used software package for analyzing structural but also functional and diffusion MRI data is Freesurfer, which includes a set of software tools for the study of cortical and subcortical anatomy (http://surfer.nmr.mgh.harvard.edu/).

Both analysis packages were applied using a region-based approach on the same acquired data sets. This approach is chosen based on the considerable heterogeneity of previous study results in order to include any potential effect of the above described

hormonal stages on regions of interest (ROIs) provided by each analysis package. Only these ROIs that have been strongly affected by menstrual cycle phases or the intake of OCs in previous studies (Rehbein et al., 2020) were included in order to evaluate these preceding findings. ROI-based approaches increase spatial information of structural data, as they are based on native space, and thus, no smoothing is required. Additionally, they enable useful summary statistics and analysis of the directly yielded mean inside ROI values within elaborate statistical models, particularly linear mixed effect models (LMM). LMMs allow the examination of even small sample sizes, the consideration of inter-individual differences, and the combination of many factors (effect of groups and hormonal phase).

In summary, this study aims to investigate

(i) *potential brain volume differences between two different cycle phases, notably the early follicular (menstruation) and the midluteal phase in women with a natural cycle;*

(ii) *potential brain volume differences between the active pill-intake and the pill-withdrawal in women taking a defined androgenic OC;*

(iii) *potential brain volume differences across both groups, notably NC women and women taking OCs;*

(iv) *and to descriptively compare the obtained results yielded by two commonly used volume-based analysis streams: ROI-based morphometry within the toolbox CAT12, and the volume-based stream using Freesurfer.*

4.2 Methods

4.2.1 Subjects

Two groups of women were included in the study. The first group comprised naturally cycling women (NC group). The second group consisted of women that took OC (OC group). They were recruited by distributing flyers and bulletins in public places and advertisements through the student- and staff- mailing list of the University of Marburg, Germany.

Naturally cycling women were invited for a short interview. They did not have to use any hormonal medication for at least six months prior to inclusion into the study and had to report a regular cycle duration (at most ± 2 days variation from individual cycle length). They were informed about the study design, filled in a brief questionnaire about their menstrual cycle (duration, date of last menstruation, current cycle day), and were introduced to the basal body temperature measurement method. The women were instructed to track their basal body temperature three to four months before the first MRI measurement. They were asked to send their temperature curves monthly to the research team. Basal body temperature was measured orally, immediately after waking up, so that the increase of 0.2 to 0.5 °C in the second half of the cycle due to increased progesterone levels could be detected. This information was used to identify the individual time of ovulation in order to schedule the subsequent MRI measurements. Additionally, women were equipped with ovulation tests to determine the luteinizing hormone (LH) concentration in their urine, which climaxes 24 to 36 hours before ovulation. They were instructed to use them on defined days before the MRI measurement and inform us about a positive testing. Each woman was tested once during menstruation (between days 2 and 5 after the onset of menstruation) and once during the midluteal phase (midcycle) (between days 5 and 9 after a positive ovulation test). We used a counterbalanced study design. Half of the NC group women were measured first during menstruation, the other half first during the midluteal phase.

Women who were taking OC performed a short telephone screening. They had to use a defined androgenic OC, containing a progestin-estradiol combination of 100 µg levonorgestrel and 20 µg ethinylestradiol. The OC had to be taken for more than six months prior to inclusion in the study. They were informed about the study design and gave all necessary information about their OC-intake, including the duration of intake, the start and end days of their blisters in order to arrange two appointments for MRI measurements. During the intake of the 12^{th} to 21^{st} pill of the blister, one measurement took place, the other measurement on days 3 to 7 of the pill-withdrawal interval. Again, the order of measurements was counterbalanced.

Thirty-eight women aged 20 to 33 years (mean age 24.5 ± 3.2 years) were initially included in the study. Twenty subjects belonged to the NC group, 18 subjects to the OC group. Inclusion criteria for both were: aged 20 – 35 years, right-handedness (Oldfield, 1971; handedness quotient ≥ 90), German as a first language, normal or corrected to normal vision, no mental and physical disorders, and CNS active medication or drugs. Two subjects from the NC group had to be excluded from the final analysis. One subject got pregnant before the second data set was acquired; the other subject was measured during the wrong cycle phase, as indicated by low hormone concentration. The final sample, therefore, comprised 36 subjects (mean age 24.4 ± 3.3 years). Eighteen subjects belonged to the NC group, 18 subjects to the OC group. The mean duration of the pill-intake was 6.3 ± 3.6 years in the OC group. Subjects' characteristics are summarized in Table 1.

After they were apprised in detail about the experimental setup and the study procedure, all subjects provided written informed consent. The study confirmed to the Declaration of Helsinki and was approved by the local ethics committee of the Medical Faculty of the University of Marburg (file reference 37/14). Subjects within the NC group received an allowance of 150 Euros; these within the OC group received 50 Euro.

4.2.2 Hormone assessment

A peripheral blood sample was collected at the MRI-examination for the determination of progesterone and estradiol concentrations. Samples were centrifuged for 10 minutes at 2500 g, 4 °C, the supernatant was collected and stored at -20 °C before analysis. Parameters were measured using an immunoassay system (UniCel ® DxI800 Access ® Immunoassay-System) at the central laboratory of the university clinic. The results of hormone concentrations were then used as a confirmation for the cycle phases and further analysis.

4.2.3 MRI data acquisition

MRI data were acquired at the University of Marburg (Department of Psychiatry) using a 3-Tesla MRI scanner (Siemens TIM Trio, Erlangen, Germany) and a 12-channel head matrix receive coil. All subjects were assessed with a large neuroimaging battery involving structural (high-resolution T1-weighted images, diffusion-weighted images) and functional measurements. The functional imaging battery included a face perception task (adapted from Frässle et al., (2016)), a word generation task (adapted from Jansen et al., (2005a)), a semantic decision task (adapted from (Fernández et al., 2001; Jansen

et al., 2006) as well as an 8-min resting-state sequence. The order of MRI measurement was as follows: (i) high-resolution T1-weighted image, (ii) task-based functional imaging measurements (task order was varied between subjects), (iii) resting-state sequence, (iv) diffusion-weighted images. In the following, we will focus solely on the T1-weighted structural images. They were acquired using a magnetization-prepared rapid gradient-echo (3d MP-Rage) sequence in sagittal plane (176 slices), covering the whole brain (TR = 1.900 ms, TE = 2.26 ms, matrix size 256 x 256 voxels, voxel size 1 x 1 x 1 mm^3, FoV= 256 x 256 mm^2, flip angle 9°). Measurement sequences and order of measurements were equal for each subject within both sessions.

4.2.4 MRI data analysis

Various methods have been developed to assess the brain morphometry based on MRI data (for a review, see Mechelli et al., 2005). In the present study, we assessed the effect of menstrual cycle phases and OCs intake on GM brain volumes in predefined ROIs. To generalize our findings, we used two standard software packages, *CAT12* and *Freesurfer*. Both software packages were specifically designed to process structural MRI data. The preprocessing steps include skull stripping, bias field correction, GM/WM segmentation, and registration with a stereotaxic atlas. As we yielded data from two time points for each subject, we chose established analysis pipelines specifically adapted for longitudinal data. These pipelines ensured voxel-wise comparability not only across subjects but also across time points within subjects.

4.2.4.1 Cat 12 ROI analysis

We used the ROI-based morphometry approach as implemented in the *CAT12* toolbox (version 12.5 1364, http://www.neuro.uni-jena.de/cat/) of the SPM12 software (version 7487 http://www.fil.ion.ucl.ac.uk/spm/), performed in Matlab (version 9.2, https://www.mathworks.com/products/matlab.html) using default settings. Longitudinal data processing included the following steps: First, structural images from both sessions were realigned to the IXI_555 template (a template derived from 555 healthy controls of the IXI database) using inverse-consistent rigid-body registration. Second, intra-subject bias correction was performed. The resulting images were then processed individually using the standard voxel-based-processing pipeline.

This pipeline comprised skull-stripping of the brain, followed by the parcellation into the left and right hemisphere, subcortical areas, and the cerebellum. Furthermore, local

white matter hyperintensities were detected (to be later accounted for during the spatial normalization and cortical thickness estimation). Subsequently, a local intensity transformation of all tissue classes was performed, which is particularly helpful to reduce the effects of higher gray matter intensities in the motor cortex, basal ganglia, or occipital lobe prior to the final adaptive maximum a posteriori (AMAP) segmentation. This final AMAP segmentation step (Rajapakse et al., 1997), which does not rely on a priori information of the tissue probabilities, was then refined by applying a partial volume estimation (Tohka et al., 2004), which effectively estimates the fractional content for each tissue type per voxel. As a last default step, the tissue segments were spatially normalized to a common reference space using DARTEL (Ashburner, 2007) registrations.

For the voxel-based ROI analyses, the volume based Neuromorphometrics atlas (provided by Neuromorphometrics, Inc. (http://Neuromorphometrics.com)) was applied. The atlas was transformed in the predefined template space using the inverse non-linear deformations, regarding which a necessary normalization of the individual scans to the CAT12 default DARTEL template was conducted. Maximum probability tissue labels derived from this atlas were then used to estimate the sum of local gray matter inside the atlas defined ROIs. The total intracranial volume (TIV) was calculated for each image by summarizing GM, WM, and CSF values. This value was subsequently used as covariates for statistical analysis.

Results were summarized in one data file for each MRI volume (2 files per subject, 72 files in total). These files included GM values for each of the 140 ROIs as well as WM, CSF, and TIV values. GM and TIV values were then further statistically processed (see 4.2.5).

4.2.4.2 Freesurfer analysis

Comparable steps to the CAT12 analysis of cortical reconstruction and volumetric segmentation, described above, were performed using Freesurfer image analysis suite (http://surfer.nmr.mgh.harvard.edu/). Using the longitudinal preprocessing stream, the analysis comprised the following steps: First, the data was preprocessed using the standard preprocessing stream for cross-sectional data. This stream included motion correction, followed by a non-parametric non-uniform intensity normalization, making relatively few assumptions about the data. Afterward, a registration and transformation from the original volume to the MNI305 atlas within the Talairach space was computed and conducted. An intensity normalization of the original volume was then performed, attempting to correct for fluctuations in intensity that would otherwise complicate intensity-based segmentation. Subsequently, the skull was removed, and an unbiased

within-subject template space and image is created (Reuter & Fischl, 2011), using robust, inverse consistent registration. Thus, all time points were aligned to an unbiased common space. Since two data sets (one per session) were derived from the same subject, rigid registration with 6 degrees of freedom (translation, rotation) was sufficient to achieve a good alignment between the intensity normalized images. The registration and its inverse were used to transfer information between time points mutually and between time points in the longitudinal stream. After the registrations, the motion corrected images from all time points were mapped to the template space and averaged to produce the final template image. The images were then again further processed cross-sectionally with the standard Freesurfer stream applying common information from the within-subject template.

Processing steps included operations such as atlas registration as well as spherical surface map generation and parcellations using the Desikan/Killiany atlas (Desikan et al., 2006) based on common information from this within-subject template, significantly increasing reliability and statistical power (Reuter et al., 2012). The final segmentation was based on both a subject-independent probabilistic atlas and subject-specific measured values. These obtained labels were mapped into the common space (MNI305) to achieve voxel to voxel correspondence for all subjects. Once the individual subject's surface map was normalized to this template, the Desikan/Killiany atlas could be used to parcellate the cortex into 70 anatomically distinct regions.

The technical details of these procedures are described in prior publications (Dale et al., 1999; Fischl et al., 2001; Fischl et al., 1999, 1999, 2002; Fischl, Salat, et al., 2004; Fischl, van der Kouwe, et al., 2004; Fischl & Dale, 2000; Han et al., 2006; Jovicich et al., 2006; Reuter et al., 2010, 2012; Segonne et al., 2004).

With regard to intracranial volume, TIV values were estimated using an atlas scaling factor. Values were found to correlate with the determinant of the transform matrix used to align an image with the atlas (Buckner et al., 2004). The authors demonstrated that a one-parameter scaling factor using an atlas based spatial normalization procedure provides a reasonable TIV estimation that is equivalent to manual TIV correction.

We obtained a table for each group and each session per hemisphere, including estimated TIV values and gray matter volumes for each ROI. These values were further statistically processed (see 4.2.5).

4.2.5 Statistical Analyses

Statistical analyses were carried out using a multi-level linear modeling (MLM) approach. Similar to ordinary least-squares regression (OLS), MLM allows examining the relationship between a set of predictors (e.g., experimental conditions) and response variables (e.g., gray matter volume measures). In addition, MLM allowed us to account for the response variable's variance (e.g., unsystematic variation among individuals), which can otherwise lead to increased error terms in OLS models, diminishing their reliability and statistical power. In the present study, data was analyzed using a two-level MLM that accounted for measures nested within subjects by estimating a random intercept for each participant.

All statistical analyses were performed in the R programming environment (R 4.0.2; R Core Team 2020; https://www.R-project.org/). Prior to analysis, all categorical variables were effect (i.e., deviation) coded, and all continuous variables centred around zero. Data manipulation was performed using custom R-scripts, largely relying on diverse functions (Wickham et al., 2019). MLMs were fitted as implemented in the R package lme4 (Bates et al., 2015). For convenience, results in the following sections are presented in a standard ANOVA format. In order to better describe key main effects and interactions, the results of pairwise contrasts computed based on the estimated marginal means (cf. R package emmeans; (Lenth et al., 2020)) of the model in question were provided. All reported p-values were estimated using the R package car (Fox & Weisberg, 2019) and adjusted according to the FDR method (Benjamini & Hochberg, 1995) when necessary (e.g., multiple testing). In addition, we provide estimates of effect sizes for each of the included predictors using the semi-partial R-squared coefficient (Lj et al., 2008). Figures were created using the packages ggplot2 (Wickham, 2016) and viridis[1].

In the following sections, we describe the specific models used for the analysis of group (i.e., women with a NC vs. women taking OC) and session (i.e., M (menstruation/pill-withdrawal) vs. P (midluteal phase/pill-intake) on 1) hormone data, 2) TIV-values obtained from CAT12 and Freesurfer, and 3) gray matter volume estimates for specific ROIs obtained from CAT12 and Freesurfer.

[1] https://cran.r-project.org/web/packages/viridis/index.html

4.2.5.1 Analysis of hormone data

The aim of analyzing hormonal data was two-fold. Firstly, it served as a proof of concept, that is, whether the targeted cycle phases are accompanied by the expected hormonal patterns and whether the intake of the pill suppressed the measured endogenous sex hormones estradiol and progesterone.
Secondly, hormonal data was more thoroughly analyzed to investigate if and to what extent estradiol and progesterone concentrations differ between groups and sessions.
For this purpose, hormonal concentration was modeled as a function of hormone (2-level within-subjects categorical predictor: Estradiol vs. Progesterone), session (2-level within-subjects categorical predictor: M vs. P), and group (2-level between-subjects categorical predictor: NC vs. OC). Prior to analysis, hormone values were transformed to the logarithmic scale to achieve a more normal distribution and thus more robust model predictions. For this reason, the model estimates described below can be interpreted as magnitude-ratios.

4.2.5.2 Analysis of TIV values

Analysis of TIV data was conducted as an initial approach to explore the obtained structural data. TIV values were modeled as a function of session (2-level within-subjects categorical predictor: M vs. P), and group (2-level between-subjects categorical predictor: NC vs. OC).

4.2.5.3 Statistical Analysis of GM ROI values obtained by CAT12 and Freesurfer

Statistical analysis of GM ROI data obtained by CAT12 and Freesurfer was performed with regard to the three main research questions:

(i) *Are there any potential brain volume differences between two different cycle phases, notably the early follicular (menstruation) and the midluteal phase in women with a natural cycle?*
(ii) *Are there potential brain volume differences between the active pill-intake and the pill-withdrawal in women taking a defined androgenic OC?*
(iii) *Are there any potential brain volume differences across both groups, notably NC women and women taking OCs?*

In order to investigate question (i) and (ii), only data from women within the NC group or the OC group, respectively, was examined using a within subject design (2-level within-subject categorical predictor: M vs. P). Focusing on question (iii), a between-subject design was applied 2-level between-subjects categorical predictor: NC vs. OC).
To this end, each ROI GM value was modeled as a function of session (2-level within-subjects categorical predictor: M vs. P) and group (2-level between-subjects categorical predictor: NC vs. OC). Within both CAT12 and Freesurfer data, ROI GM values were defined as the dependent variable, with subject modeled as the intercept (random-factor), scaled age, and TIV values as predictor variables and including the interaction session*group.

4.3 Results

4.3.1 Exploratory Analysis of Hormones

In the NC group, progesterone and estradiol concentrations proceeded in accordance with a normal, natural cycle with the highest progesterone and estradiol concentration during the midluteal phase and low concentration during the early follicular phase (menstruation) (Mihm et al., 2010). For women taking OCs, progesterone values were generally small during pill-intake and pill-withdrawal, whereas estradiol concentrations were higher during pill-withdrawal.

Table 4. *Subjects' age and individual hormone concentrations.*

NC = natural cycle; OC = oral contraceptives; M = menstruation and pill-withdrawal; P = post-ovulation (midluteal phase) and pill-intake

ID	group	age [years]	progesterone P [µg/L]	progesterone M [µg/L]	estradiol P [ng/L]	estradiol M [ng/L]
1	NC	27	7.77	0.33	92.8	26.6
2	NC	26	16.86	0.31	95.24	22.81
3	NC	32	9.75	0.76	151.21	44.7
4	NC	22	7.88	0.23	231.69	42.82
5	NC	26	10.15	0	268.9	31.54
6	NC	24	24.53	0.63	393.12	28.98
7	NC	28	16.99	0.26	235.4	33.96
8	NC	27	5.78	0.09	124.66	22.17
9	NC	25	12.93	0.57	122.57	29.35

10	NC	27	8	0.28	163.2	12.96
11	NC	23	4.39	0.03	79.17	16.74
12	NC	20	5.55	0.81	101.97	24.39
13	NC	24	4.65	0.5	189.01	36.75
14	NC	23	6.01	0.1	107.96	45.03
15	NC	21	9.39	0.19	104.27	21.77
16	NC	21	4.94	0.01	56.65	19.19
17	NC	20	9.22	0.13	50.04	43.33
18	NC	22	6.43	0.19	145.6	45.19
19	OC	20	0.1	0.14	0	0
20	OC	20	0.63	0.13	0	43.71
21	OC	27	0.17	0.02	0	9.29
22	OC	33	0.23	0.49	23.98	6.51
23	OC	29	0.47	0.65	11.36	109
24	OC	26	0.23	0.33	24.58	59.44
25	OC	23	0.1	0.19	2.25	24.63
26	OC	23	0.31	0.26	3.47	14.57
27	OC	23	0.79	0.99	8.37	38.57
28	OC	21	0.35	0	11.06	17.18
29	OC	22	0.09	0.09	134.7	32.41
30	OC	24	0.18	0.3	10.55	63.16
31	OC	22	0.43	0.07	1.84	6.12
32	OC	29	0.7	0.56	0.16	16.18
33	OC	23	0.03	0.26	0	21.19
34	OC	23	0.22	0.04	0	7.29
35	OC	25	0.78	0.2	17.4	72.51
36	OC	28	0.55	0.15	15.68	19.97

Statistical analysis revealed a significant main effect of hormone ($F(1, 102) = 16.429$, $p < 0.001$, $spR^2 = 0.139$), indicating that the estradiol concentrations were overall (i.e., regardless of group and session) significantly higher than progesterone concentrations ($b = 2.38$, $SE = 0.507$, $t(102) = 4.053$, $p = 0.0001$). Furthermore, the analysis revealed a significant main effect of the variable session ($F(1, 102) = 20.177$, $p < 0.001$, $spR^2 = 0.165$), indicating that hormone concentration were generally higher at time point P (midluteal phase/pill-intake) compared to time point M (menstruation/pill-withdrawal) ($b = 0.383$, $SE = 0.082$, $t(102) = -4.492$, $p <.0001$). Analyses also revealed a significant main effect of the factor group ($F(1, 34) = 46.303$, $p < 0.001$, $spR^2 = 0.577$), indicating that the overall hormonal concentration was significantly higher in women with a natural cycle compared to the women taking OCs ($b = 8.14$, $SE = 2.51$, $t(34) = 6.805$, $p <.0001$). This effect was further explained by a significant two-way interaction between group and session ($F(1, 102) = 69.468$, $p < 0.001$, $spR^2 = 0.405$). As depicted in Figure 12, both groups showed comparable hormone concentrations at time point M (menstruation/Pill-

withdrawal) ($b_{NC / OC}$ = 1.37, SE = 0.515, t(69) = 0.845, p = 0.401), while at time point P (midluteal phase/Pill-intake), hormone concentrations were significantly higher in t NC women compared to women taking OCs ($b_{NC / OC}$ = 48.23, SE = 18.078, t(69) = 10.340, p < 0.001).

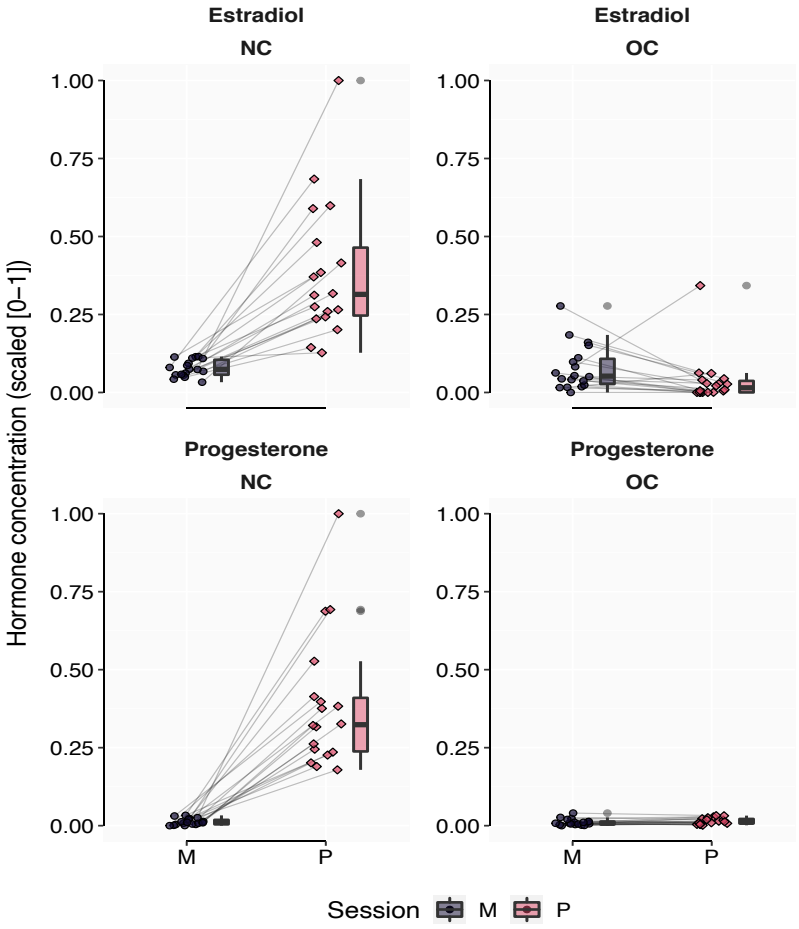


***Figure 12**. Results of hormone values across the menstrual cycle and pill-intake.*

In the upper figure, scaled estradiol values are displayed; the lower figures display progesterone concentrations. On the left, values obtained from NC women are shown in gray during menstruation (M) and in red during mid-cycle (P). Hormone concentrations yielded from women taking OCs are shown in the right half in gray during pill-withdrawal and in red during pill-intake. Overall higher hormone concentrations were detected in the NC group during midcycle; high fluctuations in both hormones' concentrations were additionally obtained only in women with a NC. Fluctuations of either progesterone or estradiol were more reduced in women taking OCs.

NC = natural cycle; OC = oral contraceptives; M = menstruation and pill-withdrawal; P = post-ovulation (midluteal phase) and pill-intake

4.3.2 Results of total intracranial volume obtained by CAT12 and Freesurfer

As depicted in Figure 13, statistical analysis revealed that neither the factor group ($F(1, 34) = 1.4213$, $p = 0.2415$, $spR2 = 0.006$), nor the factor session ($F(1, 34) = 0.2093$, $p = 0.6502$, $spR2 = 0.040$), nor the two-way interaction between group and session ($F(1, 34) = 0.0546$, $p = 0.8167$, $spR2 = 0.001$) significantly influenced the TIV-measures obtained from CAT12.

Similarly, neither the factor group ($F(1, 34) = 1.8775$, $p = 0.1769$, $spR2 = 0.052$), nor the factor session ($F(1, 34) = 1.1891$, $p = 0.2832$, $spR2 = 0.033$), nor the two-way interaction between group and session ($F(1, 34) = 1.1639$, $p = 0.2883$, $spR2 = 0.033$) significantly influenced the TIV-measures obtained from Freesurfer. Across analysis pipelines, women with a NC descriptively showed widely distributed and additionally larger TIV values compared to women taking OCs beyond both sessions.

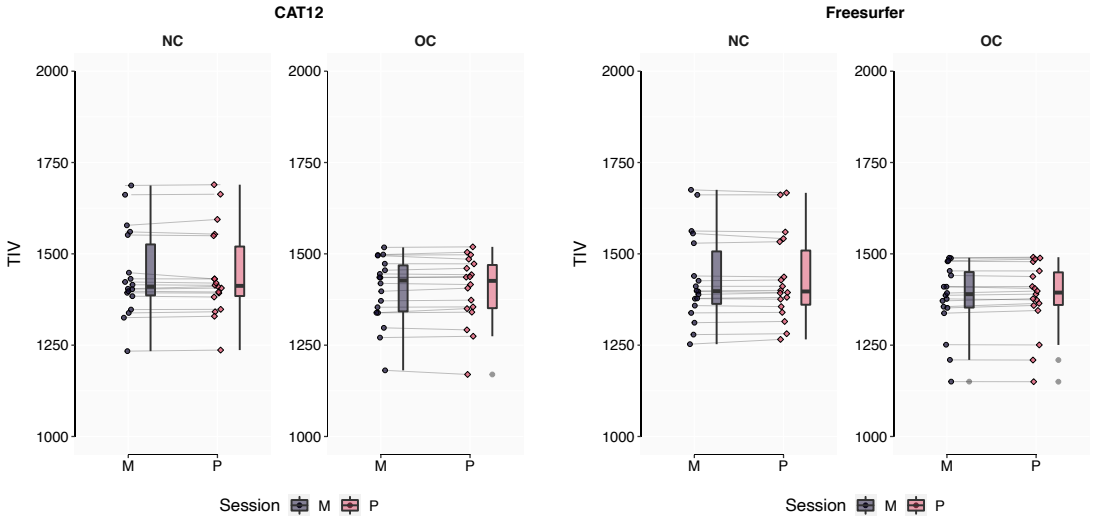


Figure 13. *TIV values obtained by CAT12 (left) and Freesurfer (right) between groups and cycle phases.*

Focusing on CAT12 data displayed in Fig. 13, left NC results obtained by women with a NC are presented on the left in gray for TIV values during menstruation (M) and in red for TIV values during the midluteal phase (P). On the right, TIV values for the OC group are shown in Fig 13, left OC. Data of pill-withdrawal is presented in gray (M), this of pill-intake in red (P). With regard to the Freesurfer results, displayed in Fig 13, right NC, results obtained by women with a NC are presented on the left in gray for TIV values during menstruation (M) and in red for TIV values during the midluteal phase (P). TIV values for the OC group are shown in Fig 13, right OC. Data of pill-withdrawal is presented in gray (M), this of pill-intake in red (P). In total, NC women displayed larger TIV values; however, significant changes between TIV values yielded from either pipeline could neither be detected between NC women and women taking OC nor between cycle phases in NC women.

NC = natural cycle; OC = oral contraceptives; M = menstruation and pill-withdrawal; P = post-ovulation (midluteal phase) and pill-intake

4.3.3 Analysis of gray matter volume ROI data

The yielded results from either CAT12 or the Freesurfer stream are presented along with the three main research questions:

Focusing on the first two questions, notably '*Are there any potential brain volume differences between two different cycle phases, notably the early follicular (menstruation) and the midluteal phase in women with a natural cycle?';* and '*Are there potential brain volume differences between the active pill-intake and the pill-withdrawal in women taking a defined androgenic OC?';* the yielded results are reported together.

CAT 12 data

Regarding the obtained results from the CAT12 stream, no significant differences were obtained between the midluteal phase and menstruation within NC women, neither in the left nor in the right hemisphere. All included ROIs showed descriptively greater GM values during the menstruation compared to the midluteal phase except the left middle temporal gyrus and the posterior insula, as well as the right superior temporal gyrus.

For women taking OCs, we found larger GM volumes during pill-intake compared to pill-withdrawal in the left and right anterior cingulate gyrus (*$p._{unc.}$ 0.0391, estimate -0.0438; t_{ratio} -2.2366*; $p._{unc.}$ 0.0460, *estimate -0-0478; t_{ratio} -2.1537;* respectively) and in the left and right posterior cingulate ($p._{unc.}$ 0.0423, *estimate -0.0196; t_{ratio} -2.1971*). The left posterior cingulate gyrus (*$p._{FDR}$ 0.0340, estimate -0.0344; t_{ratio} -3.2391*), as well as the right middle cingulate gyrus (*$p._{FDR}$ 0.0146, estimate -0.0261, t_{ratio} -3.1638*) and the right superior temporal gyrus (*$p._{FDR}$ 0.0047, estimate -0.0613, t_{ratio} -4.0018*), also showed significantly larger GM volumes during pill-intake compared to pill-withdrawal after correction for multiple testing. All included ROIs, except for the left posterior insula, yielded descriptively larger GM values during pill-intake compared to pill-withdrawal.

Freesurfer data

With regard to the data obtained from the Freesurfer stream, overall, no significant GM volume differences of investigated ROI neither between cycle phases in women with a NC nor between pill-take and pill-withdrawal in women taking OCs could be detected after correction for multiple testing. Uncorrected data revealed larger GM volumes in the right posterior cingulate gyrus during the midluteal phase compared to menstruation (*$p._{unc}$ 0.0199, estimate 0.0732; t_{ratio} 2.5689)* in women with a NC cycle;

These results are summarized in Figure 14 and Table 7 and 8 in the Appendix.

With regard to the third research question '*Are there any potential brain volume differences across both groups, notably NC women and women taking OCs?'* a significant main effect of group for the left posterior insula ($p_{unc.}$ *0.0237, estimate 0.1388,* t_{ratio} *2.3749)* and the left middle cingulate gyrus *(*$p_{unc.}$ *0.0391, estimate 0.2759,* t_{ratio} *2.1515)* was obtained through the CAT12 stream with larger GM values in women with a NC compared to women taking OCs. No significant differences were detected within the right hemisphere.

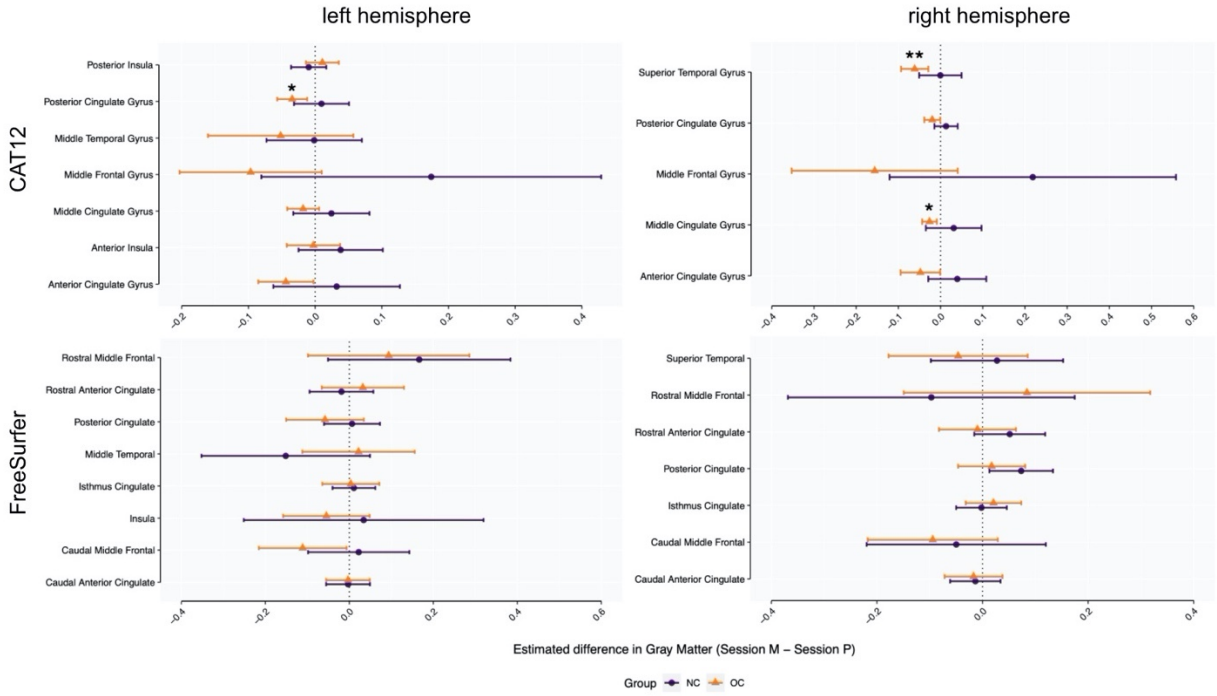


Figure 14. *Estimated differences in gray matter between sessions.*

Figures in the upper half of Figure 14 display data obtained by the CAT12 stream, figures on the lower half present data obtained by the Freesurfer stream. Figures on the left present data within the left hemisphere, figures on the right, data of the right hemisphere. Results for women with a NC are plotted in purple, these of women taking OCs in orange. Results are presented as the difference of values obtained between session M (menstruation/pill-withdrawal) and session P (midluteal phase/pill-intake); thus, positive values on the right side of the dotted line represent areas that present larger GM values during session M (menstruation/pill-withdrawal) compared to session P (midluteal phase/pill-intake) and vice versa.

* = p < 0.05 fdr corrected; ** = p < 0.01 fdr corrected; NC = natural cycle; OC = oral contraceptives

However, neither the left posterior insula nor the left middle cingulate gyrus survived after controlling for multiple testing (p_{FDR} *0.1367 and* p_{FDR} *0.1367*, respectively).

Descriptively, ROIs obtained via the CAT12 stream showed overall larger GM values in women with a NC compared to women taking OCs.

Focusing on the Freesurfer stream results, a comparable pattern could be observed: Across hemispheres, women with a NC showed descriptively larger GM values except the left rostral anterior and the left caudal anterior cingulate gyrus, as well as the right isthmus cingulate gyrus. The only significant GM volume difference was yielded within the left insula with larger values in women with a NC than women taking OCs ($p_{unc.}$ *0.0238, estimate 0.3776,* t_{ratio} *2.3734)*; however, this effect did not survive controlling for multiple testing (*FDR corrected 0.1903)*. The results are summarized in Figure 15 and Table 9 and 10 in the appendix).

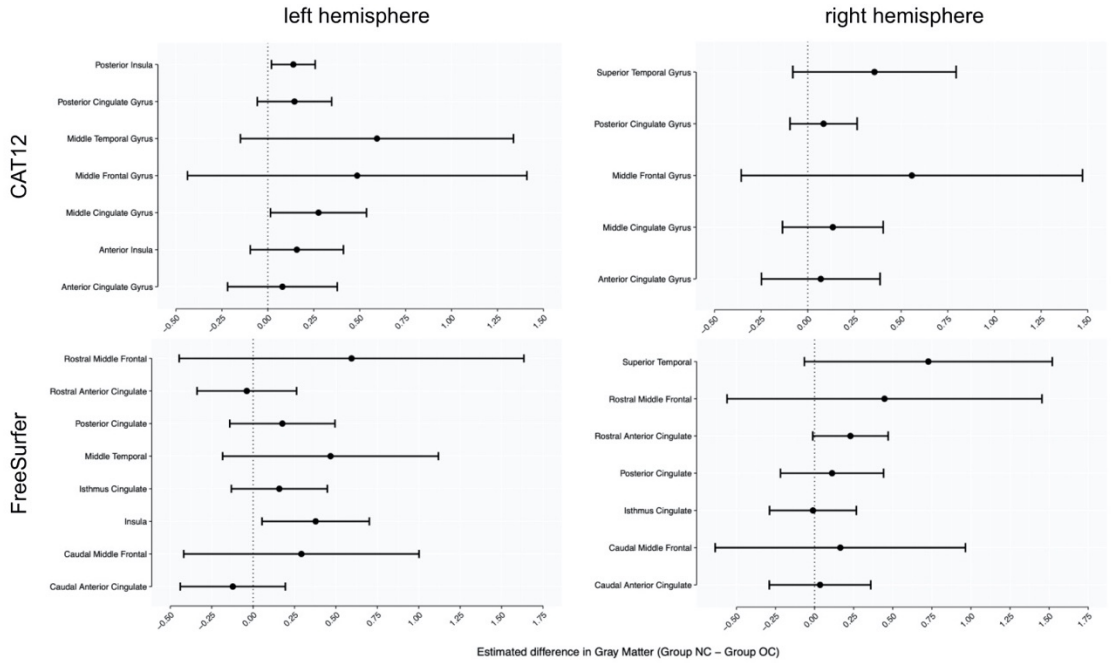


Figure 15. *Estimated differences in gray matter between women with a NC and women taking OCs.*

Figures in the upper half of Figure 15 display data obtained by the CAT12 stream, figures on the lower half present data obtained by the Freesurfer stream. Figures on the left present data within the left hemisphere, figures on the right, data of the right hemisphere. Results are presented as the difference of values obtained within the NC group and the OC group; thus, positive values on the right side of the dotted line represent areas that present larger GM values in women with a NC compared to women taking OCs and vice versa.

NC = natural cycle; OC = oral contraceptives; M = menstruation and pill-withdrawal;

P = post-ovulation (midluteal phase) and pill-intake

4.4 Discussion

Studying sex differences of the human brain and highlighting differences between men and women is an acclaimed topic in popular-scientific literature. Still, it also has practical motivations, aiming to uncover sex-specific mechanisms of various neurological and psychiatric disorders, which are more common in one characteristic sex. For example, eating disorders and migraines are more common in women (Ferrante et al., 2012; Striegel-Moore & Bulik, 2007), whereas men are more likely to suffer from autism spectrum disorder and Parkinson' disease (Baldereschi et al., 2000). Yet, it is not entirely understood what exactly causes these variable incidence rates in neurological and psychiatric disorders between sexes. One prime suspect might be the different concentration of sex hormones, the major control elements of female and male characteristics.

However, little is known about their operating principles during a human's lifespan and which brain regions and structures are mainly affected. Whereas sex hormones are relatively stable in men, women experience immense fluctuations not only across their lifespan (e.g., pregnancy, menopause) but also within an approximately monthly demeanor, the menstrual cycle. Thus, in order to shed first light into the sex hormones' mode of action within the brain, and therefore to identify distinct targeted regions, women during their childbearing years have become a major study tool to investigate hormonal effects on the brain. Even more, women within this age group are also likely to prevent pregnancies by using hormonal contraceptives, with oral contraceptives being one of the most commonly used prescription drugs in the world (Brynhildsen, 2014) used by over 100 million women worldwide (Petitti, 2003). In turn, hormonal contraceptives also influence the occurrence of various neurological and psychiatric diseases, e.g., (Skovlund et al., 2016), amplifying the suspicion of the importance of the influence of sex hormones on the brain function and, therefore, the underlying brain structure.

Presently only a handful of studies are published that characteristically investigated the influence of different hormonal states on brain structure during the menstrual cycle or in women taking OCs. In this study we therefore, aimed to investigate the effects of physiological and exogenously applied changes within the hormonal system on brain structure, particularly on gray matter volume in a strictly controlled dataset. Accordingly, women with a natural cycle, and women taking a specified oral contraceptive were investigated and underwent MRI measurements twice. One appointment was arranged during menstruation or pill-withdrawal, respectively; the other examination was performed during midluteal phase or pill-intake, respectively.

The discussion is divided into the following parts:

The first part summarizes and discusses the obtained GM volume results. The second part focuses on the application of the different processing pipelines for structural data, namely CAT12 and Freesurfer. The third part reviews the conceptualized study design. The discussion is then closed with this study's limitations and outlook, followed by an overall conclusion.

Effects of menstrual cycle phases and OC intake on regional GM volumes

With the aim to shed light on the effect of the menstrual cycle phase and the intake of OCs, three research questions were addressed, namely, '*Are there any potential brain volume differences between two different cycle phases, notably the early follicular (menstruation) and the midluteal phase in women with a natural cycle?';* '*Are there potential brain volume differences between the active pill-intake and the pill-withdrawal in women taking a defined androgenic OC?';* and '*Are there any potential brain volume differences across both groups, notably NC women and women taking OCs?'*

Regarding the *first two questions,* only within-subject differences were considered to investigate potential GM volume differences between the two opposed menstrual cycle phases, namely early follicular-/menstruation and midluteal phase (in the NC group) and between pill-intake and pill-withdrawal (in the OC group), respectively.

No differences in ROI GM volume were observed between the two examined menstrual cycle phases, early follicular phase (menstruation) and midluteal phase in the NC group. So far, literature comparing the effect of these opposed cycle phases on brain structure is very sparse. De Bondt et al. (2016) measured women with a NC across three cycle phases: follicular, ovulatory, and luteal phase. Their result didn't point to any significant differences between the follicular and luteal phases, in line with the here described study. Pletzer and colleagues also compared follicular and midluteal phase and reported larger GM volumes in the right fusiform gyrus during the follicular phase, however, only when applying a liberal uncorrected threshold (Pletzer, et al. 2010). Pletzer and colleagues conducted a follow-up study, investigating three menstrual-cycle phases, namely early follicular, pre-ovulatory, and midluteal phase. They detected larger GM volumes in the left middle frontal gyrus during early follicular- compared to midluteal phase; however, this effect did not survive FDR correction (Pletzer et al., 2018). Comparable to results obtained within our study, most reported regional GM values were not significantly but only descriptively larger during the follicular compared to the midluteal phase.

One reason this study did not detect any significant GM volume changes within NC women could be the overall low included number of subjects (n=18).

Additionally, inter-individual differences in hormone concentrations (see Figure 12) varied strongly, as did yielded ROI GM values (see Figure 14, results in NC women). For example, CAT12 results of the left and right middle frontal gyrus, including also contradictory results, as well as CAT12 results for the left middle temporal gyrus. This region showed different markedness of the obtained effects with either larger or smaller GM volumes comparing the same contrast. It remains an open question if, for example, individual differences of hormonal concentration might reflect individual differences of ROI volumes shaping. Within the examined cycle phases, the fluctuation of progesterone was the most pronounced, as the early follicular phase (low progesterone levels) and the midluteal phase (high progesterone levels) were compared. As this project is part of a larger study, these cycle phases were chosen due to the highest discrepancies of progesterone levels. Fluctuations of progesterone have been shown to alter degrees of lateralization of various cognitive functions (e.g., language) and were summarized within the progesterone-dependent-interhemispheric decoupling hypothesis (Hausmann & Güntürkün, 2000). However, these cycle phases might not be accompanied by large structural changes, in line with previous results (De Bondt et al., 2016; Pletzer et al., 2015a, 2018). In contrast, studies investigating additional cycle phases that are mostly driven by high estradiol concentrations could demonstrate significant changes of GM volumes (Barth et al., 2016; Lisofsky, et al., 2015; Pletzer et al., 2010; Protopopescu et al., 2008).

Protopopescu et al. (2008) conducted one of the first studies investigating the effect of menstrual cycle phases on female brain structure. They compared GM volumes between the late follicular phase/ovulation (characterized by high estradiol concentrations) and the late luteal phase (with high progesterone concentrations). Their results showed increased GM volumes in the right anterior hippocampus and decreased volumes in the right dorsal basal ganglia during late follicular phase/ovulation. Ossewaarde et al. (2011) compared the same cycle phases and found increased GM volume in the dorsal part of the left amygdala during the late luteal phase. De Bondt et al. (2016) could not reproduce the yielded results by Protopopescu, but they detected larger regional GM volumes in the left insula during ovulation. Hagemann et al. (2011) examined whole-brain volume changes and identified a GM volume peak along with a reduction of CSF during ovulation, which is characterized by high estradiol and LH concentration. To detangle the potential effect of estradiol from any other disturbance variable, animal studies are a reliable start to consider the sole effect of this hormone on brain structure. Woolley (1998) identified increasing spine density in the female rat hippocampus accompanied by increasing levels of estradiol, Khan et al. (2013) extended the results by adding the prefrontal cortex as a targeted brain region. These results obtained from animals are

based on the rat's 4-5 days estrous cycle and suggest that the hippocampal volume is positively correlated to estradiol increase from early to late follicular phase, remaining high until the luteal phase. Afterward, it starts to decrease due to rising progesterone levels (Lisofsky et al., 2015). Focusing back to human studies, these hypotheses could be confirmed (Lisofsky, et al., 2015). They identified increased GM volumes of both hippocampi from the early to the late follicular phase. Comparing these two cycle phases, characterized by overall low gonadal hormone concentrations during the early follicular phase and exclusively high estradiol levels during the late follicular phase, they detangled the yielded differences due to estradiol effects. This effect of estradiol on hippocampal volume could be further underlined by the effect of estrogen replacement therapy during menopause that resulted in larger hippocampal volume in treated women compared to men and women without this therapy (Lord et al., 2008).

Focusing on women taking the defined androgenic OC, significantly larger ROI GM volumes were found during pill-intake as compared to pill-withdrawal. This included the right superior temporal- and middle cingulate gyrus, as well as the left posterior cingulate gyrus. Uncorrected data revealed differences additionally in the left anterior cingulate gyrus as well as in the right anterior- and posterior cingulate gyrus. Focusing on the right superior temporal gyrus (STG), this region is involved in various processes, e.g., it acts as a site of integration between dorsal and ventral visual streams contributing to the processing of object- and space-related information as lesions within the right STG result in visual neglect (Karnath, 2001) but also in social perception as larger right STG volumes in autistic individuals and abnormal activity was detected that are related to its key role in social perception (Jou et al., 2010). Additionally, this study identified larger GM volumes within the left and right cingulate gyrus that is part of the limbic system. Thus it is involved in emotion, attention and social behavior (Drevets et al., 2008; Hadland et al., 2003). Particularly, the anterior cingulate gyrus (ACC) is known to be an important brain region for cognitive and affective control (Bush et al., 2000). Mood changes, increased rates of depression, anxiety, fatigue, neurotic symptoms, compulsion and anger are identified side-effects of OC users (Kulkarni, 2007; Pletzer & Kerschbaum, 2014; Robinson et al., 2004), whose vulnerability might be related to the obtained changes within these regions.

Similar to the first research question, existing literature investigating the effect of OC use on brain structure is scarce. It is even more limited when considering androgenicity characteristics and explicit concentration of the hormone components. Nevertheless, it has already been shown that different OCs affect brain structure contrastingly, this fact has often been neglected in previous studies, e.g. (De Bondt et al., 2013a; Pletzer et al.,

2010). Until now, only two studies considered the distinct effects of OCs with either androgenic- or anti-androgenic actions (De Bondt et al., 2016; Pletzer et al., 2015a). With relevance to our study results, De Bondt and colleagues could not identify any structural changes between pill-intake and pill-withdrawal caused by the use of androgenic OCs. However, they merged users of second- and third-generation OCs. Thus, OCs with different progestin types (i.e., second-generation OCs include levonorgestrel and norgestrel, third-generation OCs include desogestrel, gestodene, and norgestodene) (Shahnazi et al., 2014) were included and the authors did not control the exact combinations of hormones. This limitation also holds true for the study of Pletzer et al. (2015a), as they included women using second and third-generation OCs containing either strong (levonorgestrel, desogestrel, norgestrel) or weak (gestodene) androgenic properties and did not report the exact combinations of the agents. They detected several ROIs displaying larger GM volumes during pill-withdrawal compared to pill-intake. However, the comparability of their results to our opposing results is restricted with regard to the described limitations of the study design. No literature considering the effect of androgenic OCs using the exact hormonal combination currently exists.

Results obtained within this study clearly demonstrate larger GM volumes within the examined ROIs during pill-intake, thus the exogenous supplementation of estradiol and progesterone derivates, notably ethinylestradiol and levonorgestrel, compared to pill-withdrawal. Focusing on the mode of action of estradiol, animal, as well as human studies confirmed larger GM volumes in distinct ROIs (Lisofsky, et al., 2015; Lord et al., 2008; Pletzer et al., 2010; Woolley, 1998).

The question arises if a comparable volume-increase effect is additionally conceivable due to pill-intake. Following this approach, the biological effectiveness of the supplemented ethinylestradiol in comparison to endogenously fluctuation estradiol needs to be considered. It is already known that upon oral administration, 90% of ethinylestradiol is absorbed within the first two hours (Endrikat et al., 2002). After absorption by the upper digestive tract, it is then metabolized by the liver and enters the systemic circulation (Goldzieher, 1993). The half-life is approximately 20 hours in women using an OC containing 100 µg of the progesterone derivate levonorgestrel and 20 µg ethinylestradiol (Sitruk-Ware & Nath, 2013). However, the repeated treatment within a typical 28 days pill-cycle (21 days of pill intake followed by seven days of pill-withdrawal) results in an increase in ethinylestradiol concentration complying with an approximately two-fold accumulation already within the first treatment cycle. After application of the last tablet, ethinylestradiol concentrations declined rapidly within 24h (Endrikat et al., 2002) (Figure 16).

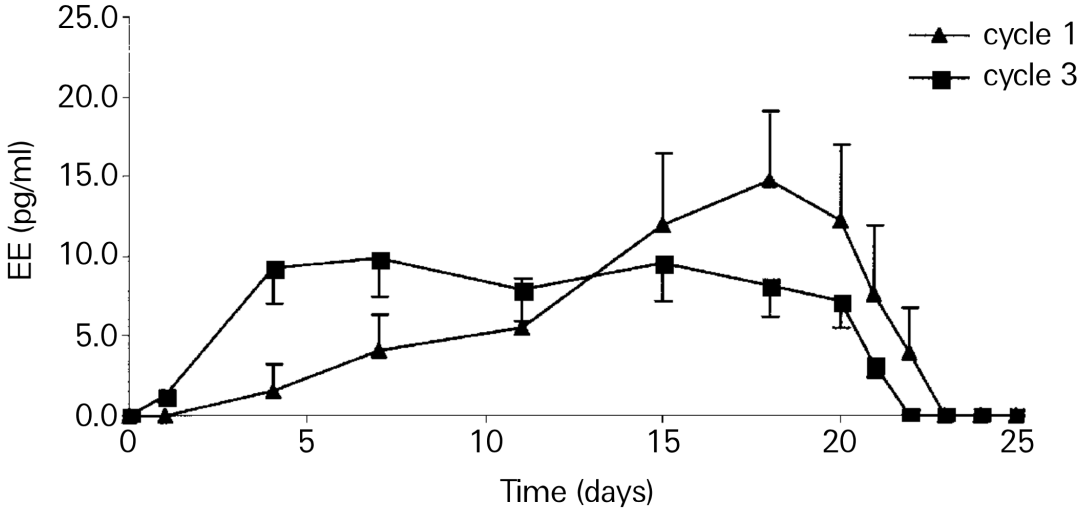


Figure 16. *Mean serum trough levels of ethinylestradiol after daily administration of Miranova®. This figure is reprinted with permission fromTaylor & Francis to reuse for a thesis or dissertation.*

Due to the described high accumulation of ethinylestradiol during pill-intake an increase biological efficiency of this estradiol derivate compared to endogenous estradiol is very likely and would be in line with previous results describing increased GM volumes caused by higher concentrations of estradiol (Lisofsky, et al., 2015; Lord et al., 2008; Woolley, 1998). Additionally, the concentration of the sex hormone-binding globulin increased about 1.5 times during the intake of the investigated OC. Consequently, concentrations of free levonorgestrel decrease, whereas the SHBG-bound fraction increases, diminishing the direct effects of this progestin derivative (Endrikat et al., 2002). This supports the assumption that the obtained GM volume differences between pill-intake and pill-withdrawal, in line with previous results, are attributed mainly to changes in ethinylestradiol concentrations.

The *third research question* addressed potential GM volume differences between both groups, notably between women with a NC and women taking OCs. Focusing on the TIV results, both analysis pipelines revealed comparable results with brain volumes ranging from 1100 cm^3 to 1700 cm^3 (≙ 1.1 l – 1.7 l).

No significant differences between groups and sessions could be detected; however, NC women descriptively showed overall larger TIVs than women taking OCs. This result is in line with results from Pletzer et al. (2015a), who also detected smaller, but not

significantly different TIV values in women taking androgenic OCs. Furthermore, they found even smaller values in women taking anti-androgenic OCs. However, within NC-women these TIV values included data from two cycle phases with overall lower estradiol concentrations. For example, considering the late-follicular phase, characterized by the highest fluctuation estradiol values, the yielded differences of TIV values are likely to be amplified. The TIV results with descriptively larger volumes in NC women are reflected in the obtained data within the investigated ROIs, however, they present only a subset of brain regions. Thus, possible changes within other not considered brain regions might also have an impact on total GM, and ultimately TIV. Descriptively, the majority of the here investigated ROIs displayed larger GM volumes for NC women compared to the OC group. Prior to FDR-correction, the uncorrected data emphasized larger GM volumes within the left insula and the left middle cingulate gyrus. However, no group differences were obtained when controlling the data for multiple comparisons. These results are in agreement with the results of previous studies. Pletzer et al. (2015a) reported larger volumes in NC (follicular phase) in left middle and superior frontal regions compared to women taking androgenic active OCs. De Bondt et al. (2016) and colleagues detected larger volumes within the left cingulate gyrus. The left insula has been detected and discussed in numerous previous studies investigating its proneness to menstrual cycle changes or OC intake: Pletzer et al. (2010) detected larger GM values within the left insula, however in women taking OCs, compared to women with a NC during the early follicular phase. De Bondt et al. (2013a) further detangled the effect of the menstrual cycle and OC-use within a highly comprehensive study. With regard to the left insula, they found short term morphometric differences, notably increased GM values during the follicular phase, associated with a peak in estradiol concentrations, compared to the luteal phase (results could be reproduced in De Bondt et al. (2016)) and during the pill-intake compared to pill withdrawal. In line with our study, however, investigating cortical thickness, instead of GM volume, Petersen et al. (2015) also detected larger values in NC women compared to women using OCs.

Focusing on the function of the insula, its vulnerability to different hormonal states is likely, as this region is known to be a hub for integrating social-emotional, sensorimotor, olfactory-gustatory, self-recognition, and cognitive networks of the brain (Craig 2009; De Bondt et al., 2016; Kurth et al., 2010; Lamm and Singer 2010; Pletzer et al., 2015a). It has been associated with emotional feelings (Damasio et al., 2013) and processing (Kl et al., 2002). Narrowing its function with regard to potential menstrual cycle-related actions, it is hypothesized to play an important role during ovulation (De Bondt et al., 2016), as during this time point, the insula is associated with feelings of romantic love, sexual arousal, empathy and orgasm quality (Craig, 2010; Ortigue et al., 2007; Vilares

et al., 2012). Thus, women may handle socio-emotional stimuli differently when chances of conception are higher (Macrae et al., 2002).
Additionally, the insula has been linked to subjective emotional feelings and emotional processing (Uddin et al., 2017). Thus, not only feelings of empathy vary as a function of cycle phase (Derntl et al., 2013) but also emotional memory formation (Ertman et al., 2011) and processing of emotional faces (Rupp et al., 2009). The reported impact of OC-use on behavioral outcomes, e.g., changes in mood and emotions is rather inhomogeneous. However, overall, women using OCs reported increased levels of anxiety, irritability, or mood swings (Lewis et al., 2019), which might be explained by changes in insula functions as OC users had lower emotion-induced reactivity in the left insula (Gingnell et al., 2013). This study's results, reporting GM volume differences within the left insula between NC women and women taking OCs could be interpreted as a structural basis for these correlates in brain function.

Differences of yielded results by the two processing pipelines

Within this section, the results achieved by the two different processing pipelines for structural data, namely CAT12 and Freesurfer are descriptively compared. Concerning research question one, no significant results in neither the CAT12 nor the Freesurfer analysis stream were obtained. Focusing on the effect within the OC group, all above-described results were received only by the CAT12 stream. With regard to the Freesurfer stream, no significant differences were yielded.
Addressing general differences between the NC and the OC group, investigated within the third research question, both pipelines, although using different calculations of TIV values, result in comparable values. Whereas within CAT12, CSF, GM, and WM are summed up to the obtained TIV, Freesurfer uses an atlas normalization tool, including an atlas scaling factor and providing an estimated TIV (eTIV) (Buckner et al., 2004). ROI data obtained via CAT12 uniformly showed descriptively larger GM volumes for women with an NC compared to women taking OCs. A majority of yielded data through Freesurfer displayed similar results, with only the left posterior and caudal anterior cingulate gyrus and the right isthmus cingulate gyrus demonstrating larger GM volumes in the OC group. Significantly larger GM volumes were detected within the left middle cingulate gyrus (by CAT12) and in the left insula by both pipelines. Data processing of both packages displays disparities during computing steps (see Methods 4.2.4.1 and 4.2.4.2). Among this, the most obvious difference might be the usage of different parcellations. Whereas CAT12 per default operates with the Neuromorphometric atlas, the Desikan atlas applied in FreeSurfer only calculates 35 ROIs per hemisphere. In

contrast, the Neuromorphometrics atlas provides a more detailed parcellation with 71 examined ROIs per hemisphere, additionally including ventricles and cerebellar regions. Concerning the obtained differences between both pipelines, it is important to address the question, which pipeline is more reliable. As MRI-based morphometry evolved into an important tool in research and diagnosis, this question gains increasing relevance. Qualitative and quantitative comparison of both pipelines focusing on cortical thickness measurements revealed high correspondence in the yielded results (Guo et al., 2019; Righart et al., 2017; Seiger et al., 2018). Seiger et al. (2018) identified a strong correlation of ROI results and overall excellent reliability values of both software. However, they observed larger cortical thickness values in the achieved CAT12 results, with the highest discrepancies between both pipelines in the insula, indicating systemic differences between the methods. Fillmer et al. (2018) examined the accuracy of CAT12 and Freesurfer, focusing on volumetric measurements. They compared the results obtained from both pipelines to a ground truth from simulated data sets that included the exact volume values of the different investigated subcortical structures. They detected CAT12 results being closer to the ground truth and more robust against different levels of SNR. Freesurfer provided close to the ground truth values, in particular for small subcortical volumes, however, these results were more prone to SNR changes. Focusing on cortical thickness, Freesurfer is known as the "gold-standard" as this software has been intensively validated by different post-mortem data sets (Fischl & Dale, 2000; Popescu et al., 2016; Rosas et al., 2002). Here, CAT12, however, occurs less sensitive for detecting differences (Righart et al., 2017). Thus, the answer to the question, which processing pipeline is more reliable, may depend on the research question and the applied methodology.

The impact of the study design

Generally, comparability between these previously described and our study is limited due to the following two points: Firstly, the study designs, including participant numbers and hormonal states, varied tremendously:
Petersen et al. (2015) examined 46 NC women, from which 21 were scanned during the follicular and 25 during the luteal phase, and 44 OC users, from which 22 were scanned during the pill-withdrawal and 22 during pill intake. Although the investigated cycle- and pill phases were in line with our study, they applied a cross-sectional design, restricting the results' comparability to our study's longitudinal design. Comparable to our study design, Pletzer et al. (2010) investigated 14 NC women twice, once during the early follicular phase and once during the midluteal phase. However, when defining the cycle

phases, they relied on verbal reports. Additionally, they included 14 women taking OCs, but they were only examined during the pill-intake. Furthermore, the combination of OC was not recorded. These issues unequivocally limited the study's validity. Pletzer et al. (2015a) measured 20 NC women, once during the early follicular phase, 22 users of anti-androgenic OCs, from whom 15 could be additionally examined during pill withdrawal, and 18 users of androgenic OCs, but with different types of progestins, from whom 12 were retrievable for a second measurement. They observed contrary results due to the pill androgenicity, with overall larger GM volumes in users of anti-androgenic medication compared to GM volumes of users of androgenic OCs. The studies by De Bondt et al. (2013a, 2016) were mostly in line and comparable to our study characteristic, as 15 NC women were examined twice, during the follicular- and luteal-phase, and 15 women taking OCs were also measured twice, during pill-intake and pill-withdrawal. However, with regard to the contraception type, different formulations were allowed. In the follow-up study, they considered this shortcoming and investigated 23 women taking an androgenic OC and ten users of anti-androgenic OCs twice, during pill-withdrawal and pill-intake. Furthermore, 24 NC women were scanned three times, during the follicular, the ovulatory and the luteal phase.

To overcome these obstacles, an elaborated study design has been developed within the here presented study. To verify the correct measurement times, blood samples were collected, and progesterone and estradiol concentrations were analyzed. Here, the expected levels were detected, with overall higher values in women with a NC. We achieved the predicted distribution with regard to the different phases within the physiological menstrual cycle (Mihm et al., 2010) and the suppression of these hormones during pill-intake (Fleischman et al., 2010). As a side-note, these results nicely demonstrate the high effectiveness of the correct use of OCs, resulting in the desired effect of suppressing ovulation. Overall, hormone concentrations were significantly higher in NC women. During the menstruation or pill-withdrawal, we detected significantly smaller hormone concentrations compared to midluteal phase or pill-intake respectively. The obtained results implicate that hormonal differences alone highly distinguish between NC women and women taking OCs and the two different sessions, thus they might already explain a vast amount of the obtained variance concerning GM values. Therefore, the factor hormone concentration was not included as a predictor variable in the here conducted subsequent analysis. Furthermore, this study concentrated only on the effect of one precisely defined oral contraceptive including the same amount of ethinylestradiol and levonorgestrel within the OC treatment of each subject within this group.

Constituting a general confounding factor in neuroimaging studies, secondly, the applied analysis pipelines, and statistical analyses vary. Focusing on the application of different analysis pipelines: Pletzer et al. (2010) and Pletzer, et al. (2015a) used the VBM5 toolbox in SPM5 in both cited studies, De Bondt et al. (2013a) applied the Jacobian-modulated VBM preprocessing in SPM8, whereas Petersen et al.(2015) used Freesurfer. Resulting statistics might to some extent, depend on the choice of the software used, and the reproducibility of findings might be restricted (Kiar et al., 2020). De Bondt et al. (2016) initially investigated the effects of different pipelines, particularly software versions for VBM (SPM8 and SPM12), to assess the reproducibility of hormone-driven regional GM volume differences in women. They could only reproduce a relatively small portion of statistically significant clusters.
In this study, we went one step further and applied two different software packages, e.g., CAT12 in SPM12 and Freesurfer, using an ROI-based approach in order to enable a direct comparison. Comparable ROIs from two parcellation atlases were included in the analysis. Besides the detected significant GM volume differences within the left insula, no further similar results could be seen. However, both packages yielded comparable tendencies towards the obtained results, with NC women showing larger GM values than women taking OCs.

In summary, an elaborated study design was established to investigate the potential influences of menstrual cycle phases and OC-treatment. Therefore, the included two cycle phases, namely early follicular/menstruation and mid luteal phase, were validated with sex hormone concentration in collected blood samples and the impact of one distinct androgenic effective OC was examined. Furthermore, the effect of the application of different software has been tested. However, the more factors are controlled and standardized, the greater limitations arise: The explanatory power of the general effects of the menstrual cycle is restricted to only two cycle phases, both with overall lower estradiol concentrations. Additionally, only a subset of ROIs was investigated, thus providing only limited insights on the effect of menstrual cycle phases and OC-treatment on GM volumes.

Limitations and Outlook of this study

This study unfortunately, contains important limitations that need to be addressed. The number of included study participants is generally low.
Nevertheless, pertaining to the investigated study sample, the drop-out rate was 5%. This can be considered low, as comparable MRI study commonly experience a drop-out

rate of 20% (e.g., De Bondt et al., 2016), particularly in concern with our strict requirements regarding the small-time frame and the high importance of capturing the correct hormonal state. All participants were required to take part in two MRI sessions, requiring high compliance of the tested subjects. An additional constraint is that this study did not account for each subject's hormonal background. This means, it was not controlled for prior pregnancies and prior OC intake in women with a NC beyond the required six months before the beginning of the study. This could potentially be a confounding factor (Pletzer et al., 2015a, 2019a). In line with Pletzer et al. (2019a), this study did not take potential other confounding variables related to previous OC use into account, e.g., relationship status. Notably, the motivation to start OC treatment is oftentimes related to entering a long-lasting relationship, while the decision to discontinue the OC use could be related to a break-up or planned pregnancy (Pletzer et al., 2019a). These relationships dynamically impact women's psychological well-being and are challenging to disentangle from the effect of OC intake. Furthermore, potential hormonal effects were investigated only within GM volume, neglecting potential impact on other structural features, e.g., cortical thickness, surface, curvature. During the last years, researchers interested in brain structure and function have been blessed with a variety of available brain atlases. However, this most apparent benefit, the use of different parcellations across studies, limits the reproducibility of the obtained results (e.g., comparing across parcellations with various organizations and numbers of ROIs) (Lawrence et al., 2020), besides prominent instabilities in numerical computations (Kiar et al., 2020). Nonetheless, without well-established principles, scientists are confronted with limited information about the used atlas, thus connecting neuroscientific findings to the organization of the atlas is strongly restricted.

Summarizing the discussed results with a focus on future research regarding the effects of female sex hormone on the brain, the consideration of interindividual characteristics concerning each subjects' hormonal state and history, e.g., taking into account previous pregnancies, the prior intake of varying hormonal contraceptives and the duration of OC intake, is highly recommended.

Future studies are advised to include additional cycle phases within an elaborated study design to further investigate the effect of estradiol on GM volumes. Additionally, the effect of menopause and pregnancy, as well as reversible effects after OC-treatment need to be more enlightened (Hoekzema et al., 2017; Lord et al., 2008; Pletzer et al., 2019a).

Although hormones are already well known to be potent neuromodulators for decades, their influence has been almost constantly ignored in neuroscience. Most likely, the exact molecular mechanisms of hormonal effects on brain plasticity are not fully understood and elucidated (De Bondt et al., 2016). Some potential mechanisms might be an

influence of sex hormones on synaptic plasticity (Leranth & Shanabrough, 2001), glial remodeling (Naftolin et al., 2007), and neurogenesis (Galea et al., 2013). However, the neglect of individual hormonal states and hormonal history could introduce wide heterogeneity in the obtained data and consequently reduce the sensitivity and accuracy of the yielded findings. Considering these points, this should be extended to become a fundamental prerequisite to future MRI studies, including women.

Conclusion

In summary, this study revealed that ROI-based morphometry results are prone, to various extents, to the influence of hormonal environments in females and the applied ROI-based morphometry processing approaches.

In general, descriptively larger GM volumes within the investigated ROIs were observed in NC women compared to women taking OCs and both software packages revealed differences within the left insula. Within-subject differences of GM volume of the examined ROIs were only detected between pill-intake and pill-withdrawal, with larger GM values during pill-intake. However, these results were yielded by only one pipeline (CAT12). A strictly controlled study design was established, including two opposed menstrual cycle phases, namely the early follicular phase/menstruation and the midluteal phase, accompanied by large changes in progesterone concentrations. Concerning the effects of OC-treatment, only one distinct androgenic effective medication was investigated. Overall low differences between the NC- and the OC group, as well as within-subject designs, were observed. However, this obvious detriment can also be interpreted as a strength of this design, as potential confounding factors might be limited. This approach is emphasized, as overall low differences in GM volumes within and between groups were revealed by both independent pipelines. Thus, reliable, reproducible analysis pipelines are an essential prerequisite. When applying different atlases, providing metadata describing similarities and differences between them in detail is also recommended to compare the obtained results feasibly. Thus, well-established standards are required to connect neuroscientific findings across different atlas organizations successfully. Taking the hormonal status into account, together with the usage of reliable data processing pipelines, is necessary to increase the reliability and validity of future research and to lay the foundations on understanding the underlying mechanism of sex-specific diseases in the brain.

5

General Discussion

In the last chapter of this doctoral thesis, first main findings of both projects are summarized, integrated and evaluated with regard to the general aim of the here presented work. Subsequently, important limitations of the conducted research are outlined and possible ways to address them in the future are proposed. Finally, an outlook on the future of sex hormones in neuroscience is given, concluding with a summary of the thesis' s accomplishments.

5 GENERAL DISCUSSION

5.1 Summary of the main findings

It is more than 100 years ago since the groundbreaking discovery of the first hormone (Starling & Royal College of Physicians of London, 1905; Tata, 2005), around 90 years ago since the revelation of progesterone as a key player in the menstrual cycle (Bickenbach 1944.; Ludwig 2011), more than 60 years ago since "the pill's" revolutionized breakthrough (Djerassi, 1992, 2006), more than 50 years ago, since scientists published their first assumption of the discovered sex hormones' actions on the nervous system (Young et al., 1964). It is around 45 years ago, since the first human MRI measurement was acquired (Hinshaw et al., 1977), around 40 years since MRI was introduced into clinical usage (Cammoun et al., 1985), and around 30 years since the birth of fMRI (Belliveau et al., 1991). It is less than 20 years ago since the researchers Lauterbur and Sir Peter Mansfield were awarded with the 2003 Nobel Prize in Physiology or Medicine for their discoveries concerning MRI (Edelman, 2014). This illustrative timeline emphasizes the expeditious development and accumulation of knowledge, expertise, and techniques within the last decades. Thanks to this accelerated growth of knowledge, we already know a lot about the function of the human brain and how it is influenced by certain factors. However, an extensive amount is still missing and might never be revealed completely.

It is known since the earliest records that the menstrual cycle affects mood and behavior in women and thus the central operation system, the brain. Plato suspected the mourning womb, grieving over not carrying a child might cause the monthly experienced symptoms (Tasca et al., 2012). Even today, women's behavior is often attributed to her menstrual cycle, e.g., 'if a woman is moody, she is asked if it is 'that time of the month. If she's feeling sexual, she's told she might be ovulating'[2]. It has emerged that these popular expressions are not unsubstantiated. Scientific studies proved that some women do experience an increase of anxiety around their menstruation (Welz et al., 2016) and that the menstrual cycle but also the intake of 'the pill' effects the brain with regard to structure and function (for a general overview, please see the introduction, more detailed information can be found here, e.g. (Rehbein et al., 2020; Sundström Poromaa and Gingnell 2014). However, precise theories of how menstrual cycle phases and the intake

[2] https://www.bbc.com/future/article/20180806-how-the-menstrual-cycle-changes-womens-brains-every-month (accessed: 24.09.2020, 14:30)

of hormonal contraceptives influence and shape the human brain is yet limited. Within this thesis, therefore, the aim was to provide necessary prerequisites and to increase previous knowledge supporting future research on this fascinating topic. Therefore, before acquiring any data, a sophisticated study design was developed, including the elaboration of ethics approval and grant application, as well as the programming of the fMRI paradigms and analyses scripts.

The aim of the major study summarized in Figure 5 was to investigate the effect of female sex hormones on distinct lateralized brain functions, notably language functions as measured applying a word generation and a semantic decision task and face processing by using a face perception task. Furthermore, resting state- as well as DTI and structural data were acquired. Whereas reliable and robust paradigms to investigate left-hemispheric functions already exist, these are lacking concerning right-hemispheric dominant processes. Thus, firstly, the establishment and examination of three fMRI paradigms in order to develop a suitable measuring device to investigate right-hemispheric functions were performed. This has been implemented in Project One. This first project was divided into two substudies: In substudy one, the utility of the applied paradigms to induce right-hemispheric lateralization in the spatial processing network was analyzed using a rather liberal “robustness” criteria, as visuospatial attention, comparable to any other lateralized cognitive function, shows high interindividual variability with regard to the distribution of the activation patterns. Therefore, three paradigms, in particular the mental rotation task, “dots-in-space” task, and Landmark task were tested with regard to the obtained data’s robustness and reliability. The most suitable paradigm was proposed to be applied in a follow-up study, investigating the effects of female sex hormones on intra- and interhemispheric brain connectivity. Under these criteria, the Landmark task was identified as the most robust but also the most suitable paradigm, as it was the shortest of the tested fMRI tasks. The landmark task was further examined in substudy two with regard to the test-retest reliability. It was therefore optimized into two versions (A and B), using a more sophisticated control condition compared to the version in substudy one with both versions in substudy two slightly differing in their difficulty level, respectively. Here more reliable data applying the easier version B, indicated by higher ICCs values. In summary, the first project demonstrated the utility of the Landmark task in general for mapping right hemispheric lateralization while simultaneously underlining the current limitation of the paradigm and its susceptibility to marginal changes of the, e.g., control condition or scanner parameters. Within this project, the first comparison of the robustness of widely applied and currently established paradigms for assessing right-hemispheric dominance during visuospatial processing and its principle and systematic investigation was performed.

Optimizing paradigms, especially these to investigate hemispheric specialization, a mechanism with a strong heterogeneous nature with regard to inter-individual brain activations, is an essential prerequisite in order to investigate which factors drive hemispheric specialization, how lateralization processes of different cognitive processes interact with each other, how the brain integrated processes that are lateralized to opposite hemispheres and how physiological factors (e.g., oxygen saturation, blood pressure, hormonal state) influence these functions. This first project of this dissertation strived to fulfill these demands of identifying such paradigms with regard to the right-hemispheric lateralization of visuospatial processing.

Secondly, a further goal of this work was to establish a sophisticated study design that enables the examination of the effects of female sex hormones, as they occur during the menstrual cycle or during the intake of OCs. Optimizing study designs, especially with regard to the consideration of strong-interindividual differences evoked by fluctuations of sex hormones as well as validating the detected results using alternative processing approaches, is an essential prerequisite in order to further understand and investigate how different hormonal states affect brain structure in general. Within the second project of this thesis, the effects of menstrual cycle phases and the intake of a distinct OC on GM volume within ROIs and how these results are influenced by the application of two different commonly used analysis pipelines and brain parcellation methods were investigated. Women were examined twice within a longitudinal study design, and blood samples were collected in order to determine hormone levels of progesterone and estradiol and to confirm the correct cycle phases. NC women were examined during menstruation, indicated by overall low concentrations of female sex hormones, and midluteal phase, specified by overall high sex hormones values. Women taking an identical androgenic effective OC were investigated during pill-intake and pill-withdrawal, both states revealing overall low-hormonal values. Using a ROI-based approach within two state-of-the-arts pipelines, notably CAT12 and Freesurfer, they provided a variety of different brain regions being descriptively affected by different hormonal states, with only one brain region, in particular, the left insula being descriptively identified with larger GM values in NC women compared to OC women across both applied processing streams. Both pipelines revealed descriptively greater GM volumes in women with a NC compared for women taking OCs in a majority of tested ROIs. Significant within-group differences could only be detected in women taking OCs. These revealed larger GM volumes during pill-intake in the left posterior cingulate gyrus as well as the right superior temporal- and middle cingulate gyrus compared to pill-withdrawal. However, these results were only yielded by the CAT12 processing stream. The second project of this dissertation emphasized the effect of menstrual cycle phase and OC intake on certain brain regions,

notably GM volumes within predefined regions, and identified overall larger GM volumes within NC women. Furthermore, the effect of the applied analysis pipeline is highlighted and described, as different processing streams and parcellation methods resulted in largely variable brain regions being affected by different hormonal states.

In this doctoral thesis, the crucial prerequisite to thoroughly study the effects of female sex hormones have been established: It is an indispensable prerequisite to apply robust and reliable paradigms in addition to reproducible processing pipelines to address these questions. With an increase in awareness regarding a troubling lack of reproducibility of yielded data of different analytical software tools, the validity in scientific derivatives and, thus, their downstream results have become questionable (Kiar et al., 2020). Further improvements of the developed and examined paradigms, as well as the use of more sophisticated scanner hardware and/or acquisition sequences, should constitute major endeavors in order to boost the sensitivity and stability of single-subject measures. Furthermore, it could be shown that one the one hand, menstrual cycle phases, and OC intake exert an effect on GM volume on a subset of investigated ROIs, with significant within-subject differences in OC women and descriptively overall larger GM volumes in NC women; and on the other hand that these results were prone to the applied processing pipeline.

5.2 Discussion of the obtained results

Within this paragraph, the obtained results are recapitulated. Firstly, brain regions that were affected by the menstrual cycle or OC intake changes are considered, followed by a general discussion on the conducted study design. Subsequently, these thesis's results are reviewed, especially with regard to reliability and reproducibility.

Within the second project, namely, "*Shedding light on regional gray matter differences in women - how do menstrual cycle phases and the intake of 'the pill' shape the female brain,*" several brain regions were observed that were prone to hormonal changes. Focusing on NC women, no significant differences in GM volume between the two examined cycle phases, notably the early follicular phase/menstruation and the midluteal phase could be observed in any investigated ROI. Descriptively the majority of ROIs showed overall larger values during the early follicular compared to the midluteal phase. These results are in line with previous literature, as GM volume differences between these distinct cycle phases are overall rare (Barth et al., 2016; De Bondt et al., 2016; De Bondt et al., 2013a; Pletzer et al., 2010; Pletzer et al., 2018, 2015a). A possible

explanation of the only meagerly detected GM volume differences could be that levels of the sex-hormone estradiol across these two investigated cycle phase were overall low. Estradiol, however, is known to possess neuromodulatory properties (Behl & Manthey, 2000; Krentzel et al., 2019; Srivastava & Penzes, 2011) and has been attributed to increasing GM volumes, not only in animals (Hao et al., 2006; McEwen & Woolley, 1994; Sheppard et al., 2019; Woolley, 1998) but also in the female human brain in specific brain regions, e.g., the hippocampus and the frontal cortex (Barth et al., 2016; De Bondt et al., 2016; Lisofsky et al., 2015; Pletzer et al., 2010; Protopopescu et al., 2008). Interestingly, these brain areas are rich in sex hormone receptors and thus show high susceptibility to hormonal changes (Barth et al., 2015). On the other side, the sex hormone progesterone seems to exert lower effects on brain structure differences between these two cycle phases. As this project is part of a larger study, these cycle phases, namely the early follicular and the midluteal phase, were chosen due to the highest discrepancies of progesterone levels. Fluctuations of progesterone have been shown to alter degrees of lateralization of various cognitive functions (e.g., language) and were summarized within the progesterone-dependent-interhemispheric decoupling hypothesis (Hausmann & Güntürkün, 2000). However, these cycle phases were not accompanied by large structural changes, in line with previous results (De Bondt et, al. 2013a; Pletzer et al., 2010; Pletzer et al., 2015a).

In the OC group, significant differences in GM volumes were observed during pill-intake in the right STG, the right middle cingulate gyrus, and the left posterior cingulate gyrus compared to pill-withdrawal. Descriptively, also the left and right anterior cingulate gyrus and the right posterior cingulate gyrus showed larger GM volumes during pill-intake. Previous research has shown that these regions are involved in social perception, emotion and attention regulation, in social behavior as well as cognitive and affective control (Bush et al., 2000; Drevets et al., 2008; Hadland et al., 2003; Jou et al., 2010) and local changes within these might attribute to the identified side-effects some OC users experienced (Kulkarni, 2007; Pletzer & Kerschbaum, 2014; Robinson et al., 2004). One assumption for the observed volume differences within women taking OCs might be due to the strong exogenous of ethinylestradiol, one of the major components within hormonal contraceptives, between pill-intake and pill-withdrawal. Ethinylestradiol is a synthetic derivative of estradiol and may exert comparable effects concerning GM structure (Endrikat et al., 2002) Thus, it is likely to contribute to the observed GM volume changes within a subset of ROIs.

Focusing on comparing GM volumes between the NC- and the OC group, no significant differences were detected within the investigated ROIs. However, the total intracranial volume and the majority of investigated ROIs descriptively showed larger volumes in the

NC group. With regard to previous literature, these results are in line with a study performed by (Pletzer et al., 2015a), who also observed smaller TIVs in women taking androgenic OCs compared to NC women. TIVs include values not only from GM, but also WM and CSF. Thus, descriptively greater TIVs could also be caused by increasing values of the two latter structures as they are prone to sex hormone fluctuations (De Bondt et al. 2013b). These non-significant differences might also be caused by the included menstrual cycle phases within the NC group: early follicular phase as well as midluteal phase are characterized by low concentrations of the female sex hormones estradiol, which in turn is known to contribute to higher GM volumes in distinct brain regions in previous animal but also human studies (Barth et al. 2016; De Bondt et al., 2016; Hao et al., 2006; Lisofsky, et al., 2015; McEwen and Woolley 1994; Pletzer et al., 2010; Protopopescu et al., 2008; Sheppard et al., 2019; Woolley 1998). Interestingly, overall larger GM volumes were detected in NC women, although the menstrual cycle phases were not chosen ideally due to overall low estradiol levels. Thus, potential higher biological effects of the synthetic ethinylestradiol compared to the endogenous estradiol (Endrikat et al., 2002) might not cause larger global GM volumes, but might be involved in the opposite effect. Furthermore, ethinylestradiol was identified to have a lower, notably 38%, affinity for the ERβ receptor, which is widely distributed in the female brain, involved in abundant functions, including neurogenesis but also in the modulation of signaling processes in emotional behavior (Kudwa et al., 2014; Sugiyama et al., 2010; Suzuki et al., 2007; Vargas et al., 2016). Possibly, the effect of both endogenous and exogenously applied female sex-hormones on GM volumes is not determined by the concentrations of these hormones in a straight-forward way, but rather is the result of their interaction and potential other factors (De Bondt et al., 2016; Österlund & Hurd, 2001). Extending this possible explanation to the effect of oral contraceptives, this could also explain observed differences: Comparing the effects of the second-generation OCs with androgenic actions with these of third-generation OCs having anti-androgenic actions, opposing effects concerning GM volumes were observed (De Bondt et al. 2016; Pletzer et al. 2015a; Sitruk-Ware 2006). Previous studies reported larger GM volumes in various ROIs in women taking anti-androgenic OCs compared to women using second generation OCs (De Bondt et al., 2016; Pletzer et al,. 2015a). Whereas progesterone analogs differ between these generations, notably second-generation OCs contain progestins related to testosterone (e.g., levonorgestrel), and third-generation OCs contain progestins more closely resembling progesterone (e.g., drospirenone), the included estradiol derivative is the same, namely ethinylestradiol, in the majority of combined OCs (Gogos et al., 2014; Guerra et al., 2013; Gurvich et al., 2020). In summary, the effect of the endogenous sex hormones estradiol and progesterone, as

well as the synthetic analogs, notably ethinylestradiol and various progestins, are difficult to detangle.

For investigating such research questions, a strictly controlled study design was established. Women with a NC were monitored with regard to their menstrual cycle three to four months before the MRI measurement. This surveillance included tracking of the basal body temperature as well as confirming their ovulation using LH-sensitive ovulation testing kids. These prerequisites were performed in order to plan the MRI measurement within the exact cycle phase of interest, namely the early follicular phase/ menstruation and the midluteal phase. Furthermore, blood samples were collected prior to the MRI measurements to determine the sex hormones estradiol and progesterone concentrations but also to validate the measurement to the striven time point. Results of the hormonal analysis revealed that all women, despite one, were examined during the striven menstrual cycle phases. Women included in the OC-group were also strictly controlled: Only women using a defined second-generation OC containing a combination of ethinylestradiol and levonorgestrel were included. Furthermore, they need to have started the OC treatment with this distinct preparation for more than six months prior to the first MRI measurement. All women were nulliparous and were investigated in a counterbalanced design. Notably half of the women were measured firstly during the early follicular phase or pill-withdrawal, the other half during the midluteal phase or pill-intake, respectively. Furthermore, individual appointments were arranged in the same daytime to diminish possible circadian effects. Comparing the conducted study designs with these of previous research, it shows very stringent regulations. Earlier studies often relied on verbal reports to determine menstrual cycle phases (Pletzer et al., 2011; Rumberg et al., 2010) or ovulation tests only (Pletzer & Kerschbaum, 2014; Pletzer et al., 2010; Protopopescu et al., 2008). Thus, the determination of the examined cycle phase was rather imprecise. Focusing on women taking OCs, the different generations were often not considered or merged (De Bondt et al., 2013a; Lisofsky et al., 2015; Pletzer et al., 2010), however, more recent studies have described a contrary effect of these concerning GM volumes (De Bondt et al., 2016; Pletzer et al., 2019a, 2015a). The study design established within this thesis has considered these issues, however, simultaneously presents resulting problems. The more factors are controlled and standardized, the greater limitations arise: Only two menstrual cycle phases were investigated, notably the early follicular/menstruation and the midluteal phase, both with overall lower estradiol concentrations (Mihm et al., 2010), thus restricting the explanatory power of general effects of the menstrual cycle on the brain. Focusing on the second study of this thesis, only a subset of ROIs was investigated, thus providing only limited

significance of menstrual cycle but also OC effects on GM volumes. In summary, within this thesis, a comprehensive study design was established to strictly control the effects of two menstrual cycle phases and a defined androgenic OC on the female brain.

With a boost of awareness regarding a concerning shortage of reliability and reproducibility in neuroscientific research, the degree of validity of the yielded results has become uncertain. Various forms of instability have been identified in structural and functional measurements, including across operating system versions (Glatard et al., 2015), minor noise injections (Lewis et al., 2017), and data set or implementation of theoretically equivalent algorithms (Bowring et al., 2019; Klein et al., 2009). These issues hold practical applications in order to decide which tool/implementation should be applied for an experiment (Kiar et al., 2020). Focusing on conventional fMRI, regional brain activity is estimated by measuring the BOLD signal that indicates changes in blood oxygenation associated neural activity (Logothetis et al., 2001). Commonly, researchers map brain activity evoked by specific cognitive functions by contrasting the regional BOLD signal during a control condition with the BOLD signal during a condition of interest (Elliott et al., 2020). Thanks to this approach, task-fMRI enables unique insights into the brain, ranging from basic perception to complex thought and, with a clinical focus, the opportunity to directly measure neurological and psychiatric dysfunction (Elliott et al., 2020). The original idea of task-fMRI was to examine functions of the average human brain by measuring within-subject differences in brain activation between task and control conditions and averaging them together across subjects to obtain a group effect, resulting in mostly robust brain activity. This led to the idea of using the same paradigms to study between-subject differences. Thus nowadays, fMRI is widely used for studying how the brains of individuals differ. However, the reliability of the most commonly applied paradigm is largely unknown and an object of current debate within this research field (Bennett & Miller, 2010; Herting et al., 2018; Nord et al., 2017). Recently, concerns have been raised that the conclusions drawn from some neuroimaging studies are either bogus or not generalizable. This might be caused by the high vulnerability of fMRI results to low statistical power, flexibility in data analysis, software error, and a lack of direct replication (Poldrack et al., 2017).

Having these in mind, project one of this thesis strived to establish a paradigm that robustly and reliably evokes right-hemispheric dominance of fMRI activation patterns both at the group and single-subject level. Concerning a thorough analysis of hemispheric lateralization, systemic analyses of inter-individual differences need to be considered. For this, experimental paradigms must provide robust measures of hemispheric lateralization not only on the group but also in individual subjects (Brandt et

al., 2013; Lohmann et al., 2004; Schuster et al., 2017). This refers to the test-theoretical concept of test-retest reliability and has already been studied for various lateralized cognitive functions, namely language (Gorgolewski et al., 2013; Jansen et al., 2006), face processing (Frässle et al., 2016), motor processing (Frässle, Paulus, et al., 2015) and declarative memory (Brandt et al., 2013). However, the reliability of imaging paradigms assessing a typical right-lateralized cognitive function, i.e., visuospatial processing, has received considerably less attention so far.

Furthermore, not only the fMRI methodology suffers from partly meager validity reputation but other measurements for which data collection or analysis offers a variety of opportunities (Kiar et al., 2020). Within project two analysis pipelines were applied to analyze structural brain data in order to validate the obtained results. Volume-based morphometry is an already well-established tool to examine brain structures, however, reliability and robustness of these tools have not been systematically tested by now (Fillmer et al., 2018). Applying two commonly used pipelines for structural analysis, namely CAT12 and Freesurfer, within this study, a descriptively comparable trend of higher GM ROI volumes within NC women compared to OC women was observed. However, within-subjects results differ, and significant differences were only provided by the CAT12 pipeline. Fillmer et al. (2018) compared the accuracy of CAT12 and Freesurfer, focusing on volumetric measurements. Whereas CAT 12 revealed results closer to the ground truth and more robust against different levels of SNR, Freesurfer results were more prone to SNR changes. Focusing on other structural features, e.g., cortical thickness, Freesurfer is known as the “gold-standard” as this software has been intensively validated by different post-mortem data sets (Fischl & Dale, 2000; Popescu et al., 2016; Rosas et al., 2002). Here, CAT12, however, occurs less sensitive for detecting differences (Righart et al., 2017). Thus, the answer to the question, which processing pipeline is more reliable, may depend on the research question and the applied methodology.

5.3 Limitations and outlook

With regard to the limitations of the performed projects, the overall small number of participants in both projects should firstly be noted. Examining around twenty participants might be sufficient while studying strong and robust effects and functions, such as language functions, but not investigating potential small individual differences, as they occur concerning fluctuating sex hormones.

With regard to the first project, the measured overall high inter-individual variability, particularly the high inconsistency of the yielded frontal and parietal activation patterns, impeded to devise a principled criterion to assess the robustness of the applied paradigms. Furthermore, in addition to these inherent differences, MRI acquisition parameters across the applied paradigms in study one were dissimilar, potentially resulting in different SNRs. Additionally, differences in cognitive and behavioral characteristics of the task themselves but also between individuals could have also distorted the effort to identify a suitable robust paradigm to investigate right-hemispheric lateralization during visuospatial processing. These limitations, however, highlight that future refinement of the utilized paradigms is strongly needed and is an essential prerequisite for a more thorough investigation of the right-hemispheric lateralization in visuospatial attention. This is even more important before applying these paradigms, providing rather heterogeneous activation patterns across subjects to research questions, which expect to obtain comparatively small differences. This is the case when studying the effects of sex hormones on the brain.

It must be noted that none of the examined paradigms assessing visuospatial functions has been included in this study. The justification for that was that the expected changes in brain function and activation patterns evoked by hormonal fluctuation are expected to be very small per se. Including a paradigm providing widely distributed activation patterns and markedly inter- as well as intra-subject heterogeneity would add further variability and might reduce the obtained effect sizes. However, as fMRI studies sophisticatedly examining the effects of menstrual cycle phases and OCs on visuospatial attention are still missing, the intriguing question arises of how these different hormonal environments might modulate the brain. Nowadays, it is known that spatial tasks favor men and that women during high-hormonal phases, notably late follicular- and luteal phase, score lower on mental rotation than during the low-hormonal phase (Schöning et al., 2007). Surprisingly, the effect of OCs and, particularly their androgenic activity has not been systematically investigated using fMRI, despite a behavioral study impressively demonstrated that performance in the applied mental rotation task was best in OC users on an androgenic treatment compared to users of antiandrogenic OCs and nonusers (Wharton et al., 2008). Thus, the questions arise: What are the underlying neural correlates of these results, and how do OCs alter these brain functions? However, to pursue this issue, suitable, markedly reliable and robust fMRI paradigms are a fundamental prerequisite that development needs to be addressed in the future.

Focusing on the second project, investigating the effects of menstrual cycle phases and OC intake on GM brain structure, only meager differences could be observed. The small number of participants is likely to explain the mostly no-significant effects. Additionally, with regard to NC women, the chosen cycle phases, notably the early follicular phase/menstruation and the midluteal phase are mostly driven by high differences in progesterone concentrations. At the same time estradiol levels are reduced across these distinct time points (Mihm et al., 2010). However, it has been hypothesized that high estradiol levels are responsible for increasing GM volumes in characteristic ROIS (Lisofsky et al., 2015; Pletzer et al., 2010; Pletzer et al., 2018; Protopopescu et al., 2008). It has to be noted that the established main study (summarized in figure 5) was conceptualized with the aim also to detangle the effect of female sex hormones on the lateralization of various cognitive functions. These are typically lateralized to one of the brain's hemispheres, and the degree of lateralization has been shown to be prone to mostly progesterone changes (Hausmann & Güntürkün, 2000). Therefore, these cycle phases were included, and the late follicular phase, driven by characteristically high estradiol levels, was neglected.

However, it needs to be emphasized that the study design was highly elaborated, as each of the subjects with a natural menstrual cycle was closely supervised and monitored throughout the entire length of the study, reflecting a crucial overall prerequisite. In this study, the intensive support and collaboration of all included study subjects lead to a very low drop-out rate. Establishing such a laborious and well-controlled study design is of high importance as menstrual cycle research is costly and time-consuming per se. The high difficulty of obtaining data from women with a natural cycle on the corresponding days of the menstrual cycle to investigate a distinct cycle phase is a major challenge. Concerning this point, different questions concerning the effects of menstrual cycle phases and OC intake need to consider distinct cycle phases: Whereas studying potential effects of sex hormones on different brain functions, particularly their degree of lateralization, might be most suitable comparing high- and low progesterone driven cycle phases (with regard to the progesterone dependent hemispheric decoupling hypothesis (Hausmann & Güntürkün, 2000); structural differences might be more prominent during high- and low estradiol differing phases (De Bondt et al., 2013a; De Bondt et al., 2016; Lisofsky et al., 2015; Pletzer et al., 2010; Protopopescu et al., 2008).

Furthermore, reliable information about the current hormonal state by collecting blood samples to precisely examine the exact hormone concentration and thus validating the cycle phase of interest is highly recommended.

Additionally, each subject's "hormonal history", in particular, previous pregnancies, previous intake of hormonal contraceptives, hormonal treatments should be taken into account, as these have shown an impact on the human brain (Hoekzema et al., 2017; Pletzer et al., 2019a).

Furthermore, one has to be careful with regard to the obtained results' interpretation. As menstrual cycle phases are governed largely by the concentration and fluctuation of the captured female sex hormones estradiol and progesterone, these might not be the only potential important influence factors affecting gray matter volumes. As already described earlier, the natural menstrual cycle and oral contraceptive both affect various metabolic processes (e.g., basal body temperature, heart rate and breathing patterns), which might affect the yielded MRI signal. Therefore, the results cannot be interpreted solely as an effect of fluctuating sex hormones but could also be evoked by further physiological parameter changes, the study did not control for. Additionally, the decision to start or end treatment of OCs might also be accompanied by changes in personal circumstances that, in turn, may affect overall psychological well-being.

Thus, it is practically impossible to detangle and identify the sole effect of the hormonal key players of the menstrual cycle on the female human brain.

In summary, future studies are, therefore advised to consider the following recommendation concerning the study design to control as many factors as possible:

Investigating NC women, favorably more cycle phases should be included. This enables one to narrow down the obtained results to a specific sex hormone, e.g., either estradiol or progesterone. Additionally, cycle phases should be validated by evaluating the exact hormone concentration in, e.g., blood samples. When investigating women under an OC-treatment, it is highly recommended to include only one explicit OC-type, containing the exact amount of estradiol and progestin derivatives. For both NC and OC women, a sophisticated anamnestic interview is advised, including each individual's hormonal history, e.g., previous pregnancies, use of other OC types, or hormonal replacement therapies. Regarding data collection and analysis, it is recommended to include robust and reliable paradigms to increase the obtained data's validity.

Applied processing pipelines should be stable and controlled concerning their reproducibility. Substantively, future research is advised to include more menstrual cycle phases as well as defined OC from third- and fourth generations to further detangle the effects of distinct endogenous and synthetic sex hormones on brain structure and function. Additionally, fMRI paradigms, investigating emotion and empathy processing, and social interactions should be included to investigate sex hormones' effects on these essential interpersonal functions. Including various neuropsychological tests and

questionnaires could further increase knowledge about the hormonal effects on cognition and psychological well-being.

The future of sex and gender in neuroscience

To discuss the future of sex and gender in neuroscience, it is necessary to look back to a time where sex differences have been disregarded as a confound in neuroscience. Nevertheless, still today, the prevailing state of research is by no means consistent, which might be due to small sample sizes, low power, and a considerable variation in methodologies (Kiar et al., 2020). This is most particularly true for research on how the menstrual cycle or hormonal contraceptives affect the human brain.

Why are we interested in sex differences in the brain at all? Many scientists aim to identify sex differences in the human brain and thus their link to sex differences in behavior, to increase the understanding of why men and women differ behaviorally and why many psychiatric diagnoses show unequal sex ratios (Hines, 2020). The gained knowledge may help to improve therapeutic targets. We already know that there are sex differences in brain structure. However, it is still unknown what is exactly responsible for this. The existing sex differences in behavior presuppose sex differences in the brain. Additionally, the factors that influence sex differences in behavioral development are liable also to affect sex differences in the brain. In summary, genes on the sex chromosomes, the concentration of fluctuating hormones over different lifespans, beginning already prenatally over puberty, the socialization by family, peers and others, but also self-socialization which is based on the cognitive understanding of gender, are all likely to contribute to sex differences in both, behavior and the brain, as these two are inseparability intertwined (Hines, 2020). Understanding at least the effect of sex hormones on the brain might shed the first light on this complex topic.

'Because, in the end, sex may be one important factor influencing our behavior, but not only men and women differ. Every individual is different and should be accepted as such' (Pletzer, 2015b).

A recently published study showed that even the identification to a gender role affected grey matter volume (Pletzer, 2019b): The author corroborated findings of sex hormones on brain structure and demonstrated testosterone driven effects in women to more male-like brain morphologies. Furthermore, estradiol led to more female-like brain morphologies. The author described a positive association between a more feminine gender role and a more female-like brain morphology in men, notably concerning the left

middle frontal gyrus. Additionally, differences in gender roles and gray matter volumes between OC-users and NC women were described (Pletzer, 2019b). Focusing on the left middle frontal gyrus, interestingly, this brain region is typically larger in women and has already been addressed in an earlier study where researchers reported larger cortical thickness in untreated male-to-female transsexuals compared to men (Luders et al., 2012). These results are in line with prior results. A prior study showed that androgen treatment increases the female brain's volume towards male proportions and anti-androgen and estrogen treatment reduced the size of the male brains towards a female morphology (Pol et al., 2006). The findings imply the plasticity of the adult human brain structure towards the opposite sex under the influence of cross-sex hormones (Guillamon et al., 2016; Pol et al., 2006).

We need to increase the knowledge about the effect of sex hormones on the brain, and thus on all resulting behavior and a majority of neurological conditions; but also, on how they affect the whole body's physiology. This is essential to understand how they affect sex specific disease mechanisms. It is already known that a vast amount of clinical manifestation of conditions and diseases are either more prevalent and men or in women (Fig.17).

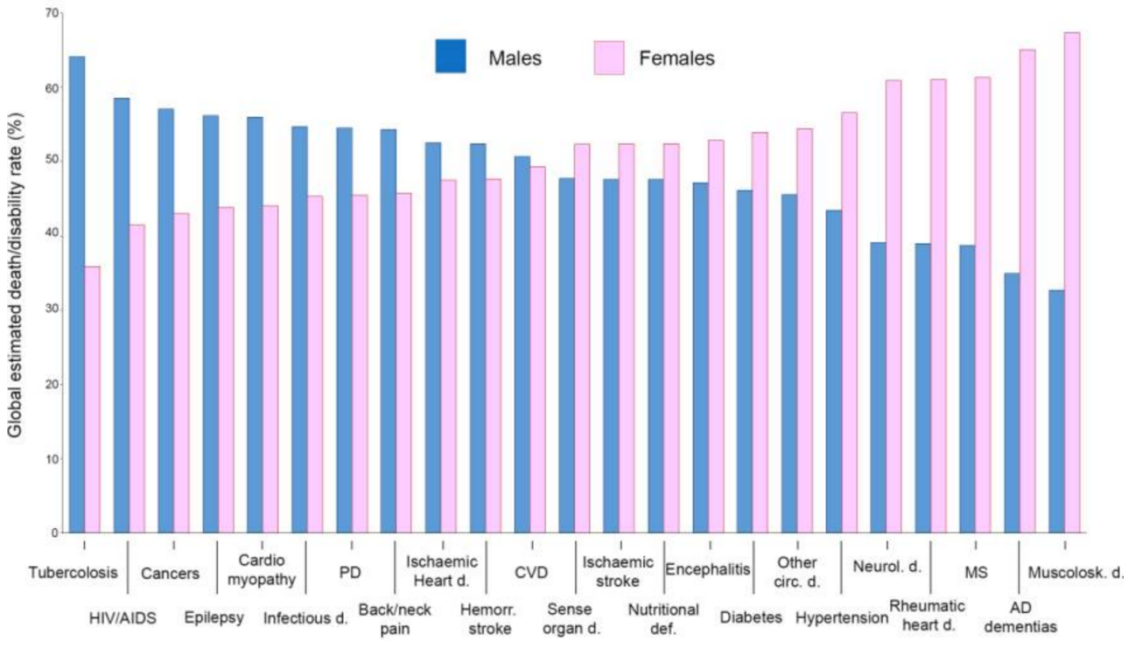


Figure 17. *Global estimated death / disability rate in males and females.*

Disease; PD: Parkinson's disease; Haemorr.: Haemorrhagic; CVD: Cardiovascular Disease; def.: deficiency; circ.: circulatory; neurol.: neurological; MS: Multiple Sclerosis; AD: Alzheimer's Disease; Musculosk.: musculoskeletal.

This figure is reprinted from (Gemmati et al., 2020)*, Creative Commons BY 4.0.*

For example, women are more prone to suffer from Autoimmune diseases, like Multiple Sclerosis (Ysrraelit & Correale, 2019), whereas men are on greater risk to develop Parkinson's disease (Miller & Cronin-Golomb, 2010).
But also, symptoms are experienced differently between sexes: Whereas chest pain is a cardinal clinical feature among patients suffering from acute myocardial infarction and a commonly known "classic" hallmark symptom, this does not always account for women, who are more likely to experience "atypical" symptoms, resulting in fatal consequences (Canto et al., 2000, 2012).
As the correct diagnosis is crucial in order to receive a timely therapy, women often experience larger infarcts and worse prognosis (Canto et al., 2007). An obvious disadvantage! This might be due to the fact that our medical knowledge is mostly based on studies exclusively including male subjects or including them to a vast excess:

'Historically, the default model for medical research has been male.'[3]

An illustrative example is the Physician's Health study[4], within which the effect of aspirin on cardiovascular disease in over 22 000 participants was investigated, none of them was female. Even more, in 1977 women in their childbearing years were officially banned by the FDA to participate in phase 1 clinical trials for nearly 20 years (Merkatz et al., 1993). One explanation leading to the exclusion of women in clinical studies was, that they would cause avoidable complexity resulting from their hormonal fluctuations and it was rather assumed, that women would react in a way comparable to men (Liu & Mager, 2016). However, the differences between the circulating levels of endogenous hormones, such as testosterone and estradiol can indeed affect pharmacokinetic or pharmacodynamic (Beierle et al., 1999; Liu & Mager, 2016; Miller, 2001) and thus the desired – but also possible side effects.
But general disadvantage in medicine, ranging from the accurate diagnosis to the correct treatment is not only a "female" problem, as also men are lately or wrongly diagnosed in rather typical female diseases, for example depression and osteoporosis (Kautzky-Willer, 2012). To overcome these problems, the impact of sex but also gender must be taken into account in research and health care practices, although it requires time, commitment, institutional and political support (Regitz-Zagrosek, 2012). In summary, I want to conclude this paragraph with a quote taken from "A Report of the Mary Horrigan

[3] Taken from http://sitn.hms.harvard.edu/flash/2018/treating-men-and-women-differently-sex-differences-in-the-basis-of-disease/) (2020-08-04; 12:10)

[4] Description of the study can be retrieved from clinical trials.gov (https://clinicaltrials.gov/ct2/show/NCT00000500)

Connors Center for Women's Health & Gender Biology at Brigham and Women's Hospital" (Johnson et al., 2016):

'It is important to note that the study of sex/gender differences benefits men as much as it benefits women. Therefore, when we fail to routinely consider the impact of sex/gender in research, we are leaving everyone's health to chance'.

5.4 Final Conclusion

The findings of this thesis might not complete the earlier described puzzle. However, they provide additional missing puzzle pieces to further understand the effect of sex hormones on the female brain. They might let us search for them in a narrower environment and with the knowledge and awareness of the importance of well-established study designs, including robust and reliable paradigms and also the significance of suitable processing and analysis pipelines. Additionally, with the future analysis of the further data yielded within this study design, including functional, e.g., cognitive functions (word-generation and semantic decision), and resting-state; as well as structural features (DTI), further light will be shed on the effects of different hormonal states on the brain.

This thesis provided suitable methodological, indispensable prerequisites to allow further examination of the effect of hormonal variation on the brain and showed the influence of two opposed menstrual cycle phases and the intake of a defined hormonal contraceptive on brain gray matter volumes within predefined ROIS.

A well-designed study was established, enabling the investigation of the effect of different menstrual cycle phases and OCs various brain functions and brain structure. Additionally, a pipeline to thoroughly investigate the robustness and reliability of a paradigm to assess right-lateralized brain activation by visuospatial processing was developed. Even as none of the investigated paradigms has proven its usefulness, this project is of rather high importance for future research endeavors in order to boost sensitivity and stability of single subject measures. Especially if the overall goal is to investigate rather small differences, as it is the case, when examining the effect of sex hormones on the brain, it is of high importance to exclude any possible other interfering sources. One has to ensure that the only yielded difference is due to the factor of interest, notably distinct values of fluctuating sex hormones.

Therefore, study design, data collection, but also, data analysis need to be addressed. Concerning study design and included subjects, it is absolutely necessary to obtain any physiological data which might impact the measured MRI signal, e.g., body temperature, heart- and breathing, oxygen saturation, additionally, as these measurements are feasible and inexpensive. Furthermore, the collection of more anamnestic information is recommended, with special regard to physiological and pathophysiological metabolic processes, e.g., previous intake or implantation of hormonal contraceptives, pregnancies, and further hormonal treatment but also disease affecting the measured MRI signal, e.g., circulatory disturbances or diabetes. Focusing on MRI data collection, meager computational reproducibility has become increasingly apparent over the last years, scrutinizing the validity of scientific findings (Kiar et al., 2020). Thus, it is highly recommended to apply stable, robust and reliable paradigms to receive good qualitative results and a stable and quality assured hardware.

Regarding data analysis, robust and reliable analysis pipelines, considering the study-design, e.g. longitudinal data, should be favored as well as the validation of the obtained results using at least one further available analysis stream.

Furthermore, different parcellation methods have to be considered.

Additionally, sensitive statistical methods, allowing complex models and the inclusion of fixed and random factors, are an important requirement.

In summary, regardless of the applied neuroimaging pipeline, e.g., for the analysis of functional or structural data, reproducibility should be investigated and reported. These results need to be considered when interpreting the obtained data.

Having thoroughly considered these effects within the previous pages, the overall research question of how sex hormones in general shape the human brain does not stop at his point but furthermore needs to overcome the simplified classification of binary definitions of sexes. It’s time to not only consider but to appreciate and embrace interindividual differences and diversity, as nature and thus science itself are diverse.

"There never were in the world two opinions alike,
no more than two hairs or two grains;
the most universal quality is diversity."

Michel de Montaigne

Summary

The stereotypic and oversimplified relationship between female sex hormones and undesirable behavior dates back to the earliest days of human society, as already the ancient Greek word for the uterus “hystera” indicated an aversive connection. Remaining and evolving throughout the centuries, transcending across cultures and various aspects of everyday life, its perception was only recently reframed. Contemporarily, the complex interaction of hormonal phases (i.e., the menstrual cycle), hormonal medication (i.e., oral contraceptives), women’s psychological well-being, and behavior is the subject of multifaceted and more reflected discussions. A driving force of this ongoing paradigm shift was the introduction of this highly interesting and important topic into the realm of scientific research. In particular, this refers to neuroscientific research as it enables a multimodal approach combining aspects of physiology, medicine, and psychology. Here a growing body of literature pointed towards significant alterations of both brain function, such as lateralization of cognitive functions, and structure, such as gray matter concentrations, due to fluctuations and changes in hormonal levels. This especially concerns female sex hormones. However, the more research is conducted within this field, the less reliable these observations and derived insights appear to be. Among other reasons, this is grounded in two particular factors: measurement inconsistencies and diverse hormonal phases accompanied by interindividual differences. The first factor refers to the prominent unreliability of one of the primarily utilized neuroscientific research instruments: functional magnetic resonance imaging (fMRI). This unreliability is seemingly present in paradigms and analyses, as well as their interplay and additionally affected by the second factor. In more detail, hormonal phases and levels apparently further influence neuroscientific results obtained through fMRI as outcomes vary drastically across different cycle phases and medication. This resulting vast uncertainty thus tremendously hinders the further advancement of our understanding of how female sex hormones might alter brain structure and function and, ultimately, behavior. Therefore, precisely controlled study designs need to be assembled and thoroughly validated as a prerequisite for the sufficient and robust research on this exceedingly important and fascinating part of biology. This important endeavor was at the very core of the here presented thesis and aimed to address the outlined concerns through two projects. The first one focused on the unreliability of a multitude of fMRI results, including those studies investigating the lateralization of specific cognitive functions. Additional variability is included by investigating women, as prior research described an influence of sex hormones on the degree of lateralization.

The research of the lateralization of various cognitive functions is generally impeded due to a quality discrepancy between paradigms addressing either left- or right-hemispheric functions. Whereas left-dominant functions can be examined with robust and reliable paradigms, equivalents are missing to investigate functions that are lateralized to the right hemisphere, especially visuospatial attention. Accordingly, the thesis's first project evaluated the robustness of three paradigms for assessing right-hemispheric dominance during visuospatial processing within a repeated measurement design. To this end, reliability and lateralization indices were assessed for each paradigm. The yielded results demonstrated the general utility of the examined paradigms while simultaneously underlining their current limitations concerning reliability and susceptibility to marginal changes of, e.g., measurement parameters. Thus, none of the examined tasks were incorporated into the second project of this thesis, which strived the conceptualization of a comprehensive study design to systematically investigate the effects of female sex hormones on gray and white matter, typically lateralized cognitive functions, and resting state. Women were investigated within a longitudinal study design, and blood samples were collected to determine blood hormone levels. Women with a natural cycle were examined during menstruation, indicated by overall low concentrations of female sex hormones, and midluteal phase, specified by overall high sex hormones levels. Women taking an androgenic effective hormonal contraceptive were investigated during pill-intake and pill-withdrawal, revealing overall low-hormonal values. As structural data is generally less prone to analytical perturbations, and prior studies reported high reliability for volume measures, within this thesis's extent, the focus was on elucidating the effects on gray matter volume within predefined regions and how these results are influenced by the application of two different analysis pipelines. Overall, descriptively larger gray matter volumes were found within naturally cycling women across pipelines. Distinct brain regions were additionally affected by pill-intake vs. pill-withdrawal within results obtained by only one processing stream. In sum, this thesis's outcomes highlight the importance of reliable paradigms, comprehensive study designs, and the application of validated analysis pipelines, as indicated by the yielded results differences examining the influence of menstrual cycle phases and oral contraceptive treatment on brain structure. The analysis of the additional acquired functional and structural data obtained within the study will further elucidate the effect of female sex hormones on the brain. Ultimately, this thesis outlines the essential requirements to further investigate and understand the female brain's underlying physiological and anatomical features that may have motivated ancient greek anatomists to designate the uterus after an outmoded psychiatric condition.

Zusammenfassung

Die Stereotype und stark vereinfachte Beziehung zwischen weiblichen Sexualhormonen und ‘typisch’ weiblichen Verhaltens geht auf die frühesten Tage der menschlichen Gesellschaft zurück, da bereits das altgriechische Wort für die Gebärmutter „Hystera“ auf eine solche Verbindung hinweist. Gegenwärtig ist das komplexe Zusammenspiel von Phasen des Menstruationszyklus und hormoneller Medikation mittels oraler Kontrazeptiva auf das psychische Wohlbefinden und Verhalten von Frauen, Gegenstand vielfältiger und stärker reflektierter Diskussionen. Eine treibende Kraft hierfür war die Einführung dieses hochinteressanten Themas in den Bereich der neurowissenschaftlichen Forschung. Diese ermöglicht einen multimodalen Ansatz, der Aspekte der Physiologie, Medizin und Psychologie kombiniert. Hier weist eine wachsende Anzahl von Literatur auf signifikante Veränderungen sowohl der Gehirnfunktion, z.B. der Lateralisierung kognitiver Funktionen, als auch der Hirnstruktur, z.B. der Konzentration der grauen Substanz, aufgrund von Veränderungen des Hormonspiegels hin. Dies trifft insbesondere auf die Fluktuation weiblicher Sexualhormone zu. Je mehr dieser Forschungszweig jedoch beleuchtet wird, desto weniger zuverlässig scheinen die Beobachtungen und abgeleiteten Erkenntnisse zu sein. Dies beruht unter anderem auf zwei Faktoren: Messinkonsistenzen und verschiedene Hormonkonzentrationen, begleitet von interindividuellen Unterschieden. Der erste Faktor bezieht sich auf die unzureichende Reliabilität eines der hauptsächlich verwendeten neurowissenschaftlichen Forschungsinstrumente: der funktionellen Magnetesonanztomographie (fMRT). Die mangelnde Reliabilität spiegelt sich in Paradigmen und Analysen sowie deren Zusammenspiel wieder und wird durch den zweiten Faktor zusätzlich verstärkt. Sexualhormone, die über verschiedene Zyklusphasen und Medikamente hinweg drastisch variieren, beeinflussen offenbar die durch fMRT erzielten Ergebnisse weiter. Die daraus resultierende enorme Instabilität behindert den Wissensgewinn des Verständnisses, wie weibliche Sexualhormone die Struktur und Funktion des Gehirns und letztendlich das Verhalten verändern. Ziel ist daher die Etablierung und gründliche Validierung präzise kontrollierte Studiendesigns, um eine reliable und robuste Forschung zu diesem faszinierenden Thema zu ermöglichen. Dieses wichtige Vorhaben stand im Mittelpunkt der hier vorgestellten Arbeit und zielte darauf ab, die skizzierten Probleme durch zwei Projekte anzugehen. Das erste Projekt adressierte die Instabilität einer Vielzahl von fMRT-Ergebnissen, einschließlich derer von Studien, die die Lateralisierung spezifischer kognitiver Funktionen untersuchten. Frühere Studien belegen, dass diese ebenfalls von Sexualhormonen beeinflusst werden. Die Untersuchung der Lateralisierung verschiedener kognitiver

Funktionen wird weiter durch eine Diskrepanz zwischen der Qualität der Paradigmen eingeschränkt: Während linksdominante Funktionen mit reliablen Paradigmen untersucht werden können, fehlen diese zur Untersuchung der rechten Hemisphäre (insbesondere von visuell-räumlichen Funktionen). Daher konzentrierte sich das erste Projekt auf die Analyse der Robustheit von drei Paradigmen zur Untersuchung rechtsdominanter visuell-räumlicher Funktionen. Zusätzlich wurden Reliabilitäts- und Lateralisierungsindizes bewertet. Die Ergebnisse zeigten die allgemeine Verwendbarkeit dieser Paradigmen, jedoch auch ihre gegenwärtigen Einschränkungen hinsichtlich der Zuverlässigkeit und Anfälligkeit für marginale Änderungen von beispielsweise Scannerparametern. Somit wurde keine der untersuchten Aufgaben in das zweite Projekt dieser Arbeit inkorporiert. Dieses Projekt strebte die Konzeptualisierung eines umfassenden Studiendesigns zur systematischen Untersuchung der Auswirkungen weiblicher Sexualhormone auf graue und weiße Substanz, lateralisierte kognitive Funktionen und Ruhezustandsfunktionen an. Frauen wurden im Rahmen eines Längsschnittstudie untersucht und Bluthormonspiegel wurden bestimmt. Frauen mit einem natürlichen Zyklus wurden während der Menstruation, mit insgesamt niedrigen Konzentrationen der Sexualhormone, und in der Midluteal Phase, charakterisiert durch hohe Sexualhormonwerte, untersucht. Frauen, unter Medikation eines androgen wirksamen Kontrazeptivums wurden während der Einnahme und des Entzugs der Pille untersucht. Da strukturelle Daten generell weniger anfällig für analytische Störgrößen sind und frühere Studien insbesondere eine hohe Reliabilität für Volumenmessungen berichteten, lag der Schwerpunkt dieser Arbeit auf der Analyse des Volumens der grauen Substanz. Zusätzlich wurde der Einfluss zweier verschiedener Analysesoftware auf die Ergebnisse untersucht. Natürlich menstruierende Frauen wiesen deskriptiv größere Volumina an grauer Substanz aus. Frauen unter hormoneller Medikation zeigten während der Einnahme im Vergleich zum Entzug der Pille größere Volumina, die jedoch nur von einer Software berichtet wurden. Zusammenfassend unterstreichen die Ergebnisse dieser Arbeit die Bedeutung zuverlässiger Paradigmen, umfassender Studiendesigns, sowie der Anwendung validierter Analysepipelines. Letzteres zeigt sich in unterschiedlichen Ergebnissen des Einflusses des Menstruationszyklus und der Einnahme oraler Kontrazeptiva auf die Gehirnstruktur. Die Analyse der zusätzlichen im Rahmen der Studie erhobenen funktionellen und strukturellen Daten, wird die Wirkung weiblicher Sexualhormone auf das Gehirn weiter vervollständigen. Diese Arbeit etablierte die wesentlichen Anforderungen, um die zugrunde liegenden physiologischen und anatomischen Merkmale des weiblichen Gehirns weiter zu untersuchen und zu verstehen, was antike griechische Anatomen dazu motiviert haben könnten, die Gebärmutter nach einer ehemals psychiatrischen Erkrankung zu benennen.

REFERENCES

Abler, B., Kumpfmüller, D., Grön, G., Walter, M., Stingl, J., & Seeringer, A. (2013). Neural Correlates of Erotic Stimulation under Different Levels of Female Sexual Hormones. *PLoS ONE*, *8*(2), e54447–e54447. https://doi.org/10.1371/journal.pone.0054447

Adcock, J. E., Wise, R. G., Oxbury, J. M., Oxbury, S. M., & Matthews, P. M. (2003). Quantitative fMRI assessment of the differences in lateralization of language-related brain activation in patients with temporal lobe epilepsy. *NeuroImage*, *18*(2), 423–438. https://doi.org/10.1016/S1053-8119(02)00013-7

Allen, J. S., Damasio, H., Grabowski, T. J., Bruss, J., & Zhang, W. (2003). Sexual dimorphism and asymmetries in the gray-white composition of the human cerebrum. *NeuroImage*, *18*(4), 880–894.

Alonso, L. C., & Rosenfield, R. L. (2002). Oestrogens and puberty. *Best Practice & Research Clinical Endocrinology & Metabolism*, *16*(1), 13–30. https://doi.org/10.1053/beem.2002.0177

Alvergne, A., & Lummaa, V. (2010). Does the contraceptive pill alter mate choice in humans? *Trends in Ecology & Evolution*, *25*(3), 171–179. https://doi.org/10.1016/j.tree.2009.08.003

Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H. B. M., & Zilles, K. (1999). Broca's region revisited: Cytoarchitecture and intersubject variability. *The Journal of Comparative Neurology*, *412*(2), 319–341. https://doi.org/10.1002/(SICI)1096-9861(19990920)412:2<319::AID-CNE10>3.0.CO;2-7

Andreasen, N. C., Flaum, M., Swayze, V., O'Leary, D. S., Alliger, R., Cohen, G., Ehrhardt, J., & Yuh, W. T. (1993). Intelligence and brain structure in normal individuals. *The American Journal of Psychiatry*, *150*(1), 130–134. https://doi.org/10.1176/ajp.150.1.130

Arélin, K., Mueller, K., Barth, C., Rekkas, P. V., Kratzsch, J., Burmann, I., Villringer, A., & Sacher, J. (2015). Progesterone mediates brain functional connectivity changes during the menstrual cycle-a pilot resting state MRI study. *Frontiers in Neuroscience*, *9*, 44–44. https://doi.org/10.3389/fnins.2015.00044

Arnoni-Bauer, Y., Bick, A., Raz, N., Imbar, T., Amos, S., Agmon, O., Marko, L., Levin, N., & Weiss, R. (2017). Is It Me or My Hormones? Neuroendocrine Activation Profiles to Visual Food Stimuli Across the Menstrual Cycle. *The Journal of Clinical Endocrinology & Metabolism*, *102*(9), 3406–3414. https://doi.org/10.1210/jc.2016-3921

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, *5*(3), 119–126.

Axmacher, N., Bialleck, K. A., Weber, B., Helmstaedter, C., Elger, C. E., & Fell, J. (2009). Working memory representation in atypical language dominance. *Human Brain Mapping*, *30*(7), 2032–2043. https://doi.org/10.1002/hbm.20645

Baldereschi, M., Di Carlo, A., Rocca, W. A., Vanni, P., Maggi, S., Perissinotto, E., Grigoletto, F., Amaducci, L., & Inzitari, D. (2000). Parkinson's disease and parkinsonism in a longitudinal study: Two-fold higher incidence in men. ILSA Working Group. Italian Longitudinal Study on Aging. *Neurology*, *55*(9), 1358–1363. https://doi.org/10.1212/wnl.55.9.1358

Bale, T. L., & Epperson, C. N. (2015). Sex differences and stress across the lifespan. *Nature Neuroscience*, *18*(10), 1413–1420. https://doi.org/10.1038/nn.4112

Bali, N., Arimoto, J. M., Iwata, N., Lin, S. W., Zhao, L., Brinton, R. D., Morgan, T. E., & Finch, C. E. (2012). Differential Responses of Progesterone Receptor Membrane Component-1 (Pgrmc1) and the Classical Progesterone Receptor (Pgr) to 17β-Estradiol and Progesterone in Hippocampal Subregions that Support Synaptic Remodeling and Neurogenesis. *Endocrinology*, *153*(2), 759–769. https://doi.org/10.1210/en.2011-1699

Baratloo, A., Safari, S., Rouhipour, A., Hashemi, B., Rahmati, F., Motamedi, M., Forouzanfar, M., & Haroutunian, P. (2014). The Risk of Venous Thromboembolism with Different Generation of Oral Contraceptives; a Systematic Review and Meta-Analysis. *Emergency*, *2*(1), 1–11.

Barbieri, R. L. (2014). The Endocrinology of the Menstrual Cycle. In Z. Rosenwaks & P. M. Wassarman (Eds.), *Human Fertility: Methods and Protocols* (pp. 145–169). Springer New York. https://doi.org/10.1007/978-1-4939-0659-8_7

Barnes, J. N. (2017). Sex-specific factors regulating pressure and flow. *Experimental Physiology*, *102*(11), 1385–1392. https://doi.org/10.1113/EP086531

Barth, C., Steele, C. J., Mueller, K., Rekkas, V. P., Arélin, K., Pampel, A., Burmann, I., Kratzsch, J., Villringer, A., & Sacher, J. (2016). In-vivo Dynamics of the Human Hippocampus across the Menstrual Cycle. *Scientific Reports*, *6*(1), 32833–32833. https://doi.org/10.1038/srep32833

Barth, C., Villringer, A., & Sacher, J. (2015). Sex hormones affect neurotransmitters and shape the adult female brain during hormonal transition periods. *Frontiers in Neuroscience*, *9*. https://doi.org/10.3389/fnins.2015.00037

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Baudry, M., Bi, X., & Aguirre, C. (2013). Progesterone–estrogen interactions in synaptic plasticity and neuroprotection. *Neuroscience*, *239*, 280–294. https://doi.org/10.1016/j.neuroscience.2012.10.051

Bayer, J., Schultz, H., Gamer, M., & Sommer, T. (2014). Menstrual-cycle dependent fluctuations in ovarian hormones affect emotional memory. *Neurobiology of Learning and Memory*, *110*, 55–63. https://doi.org/10.1016/j.nlm.2014.01.017

Beaumont, J. G. (1997). Future research directions in laterality. *Neuropsychology Review*, *7*(3), 107–126.

Behl, C., & Manthey, D. (2000). Neuroprotective activities of estrogen: An update. *Journal of Neurocytology*, *29*(5), 351–358. https://doi.org/10.1023/A:1007109222673

Beierle, I., Meibohm, B., & Derendorf, H. (1999). Gender differences in pharmacokinetics and pharmacodynamics. *International Journal of Clinical Pharmacology and Therapeutics*, *37*(11), 529–547.

Belliveau, J. W., Kennedy, D. N., McKinstry, R. C., Buchbinder, B. R., Weisskoff, R. M., Cohen, M. S., Vevea, J. M., Brady, T. J., & Rosen, B. R. (1991). Functional mapping of the human visual cortex by magnetic resonance imaging. *Science*, *254*(5032), 716–719. https://doi.org/10.1126/science.1948051

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300. JSTOR.

Bennett, C. M., & Miller, M. B. (2010). How reliable are the results from functional magnetic resonance imaging? *Annals of the New York Academy of Sciences*, *1191*, 133–155. https://doi.org/10.1111/j.1749-6632.2010.05446.x

Benton, A. L., University), A. B. (Assistant P. of P. and P. S., Assistant Professor of Psychiatry and Psychology Rush, Sivan, A. B., Hamsher, K. deS, Varney, N. R., & Spreen, O. (1994). *Contributions to Neuropsychological Assessment: A Clinical Manual*. Oxford University Press.

Bickenbach: Hemmung der Follikelreifung durch Progesteron... - Google Scholar. Retrieved September 18, 2019, from https://scholar.google.com/scholar_lookup?author=W.%20Bickenbach&author=E.%20Paulikovics&journal=Zbl%20Gynak&volume=68&pages=153-157&publication_year=1944

Biswal, B. B. (2012). Resting state fMRI: A personal history. *NeuroImage*, *62*(2), 938–944. https://doi.org/10.1016/j.neuroimage.2012.01.090

Bitzer, J., & Simon, J. A. (2011). Current issues and available options in combined hormonal contraception. *Contraception*, *84*(4), 342–356. https://doi.org/10.1016/j.contraception.2011.02.013

Bleton, H., Perera, S., & Sejdić, E. (2016). Cognitive tasks and cerebral blood flow through anterior cerebral arteries: A study via functional transcranial Doppler ultrasound recordings. *BMC Medical Imaging*, *16*, 22. https://doi.org/10.1186/s12880-016-0125-0

Bonenberger, M., Groschwitz, R. C., Kumpfmueller, D., Groen, G., Plener, P. L., & Abler, B. (2013). It's all about money: Oral contraception alters neural reward processing. *Neuroreport*, *24*(17), 951–955. https://doi.org/10.1097/WNR.0000000000000024

Bowring, A., Maumet, C., & Nichols, T. E. (2019). Exploring the impact of analysis software on task fMRI results. *Human Brain Mapping*, *40*(11), 3362–3384. https://doi.org/10.1002/hbm.24603

Brandt, D. J., Sommer, J., Krach, S., Bedenbender, J., Kircher, T., Paulus, F. M., & Jansen, A. (2013). Test-Retest Reliability of fMRI Brain Activity during Memory Encoding. *Frontiers in Psychiatry*, *4*, 163. https://doi.org/10.3389/fpsyt.2013.00163

Brinton, R. D., Thompson, R. F., Foy, M. R., Baudry, M., Wang, J., Finch, C. E., Morgan, T. E., Pike, C. J., Mack, W. J., Stanczyk, F. Z., & Nilsen, J. (2008). Progesterone receptors:

Form and function in brain. *Frontiers in Neuroendocrinology*, *29*(2), 313–339. https://doi.org/10.1016/j.yfrne.2008.02.001

Broca, P. (1861a). Nouvelle observation d'aphémie produite par une lésion de la troisième circonvolution frontale. *Bulletins de La Société d'anatomie*, *2e serie*(6), 398–407.

Broca, P. (1861b). Perte de la parole: Ramollissement chronique et destruction partielle du lobe anterieur gauche du cerveau. *Bulletins de La Societe d'anthropologie*, *1re serie*(2), 253–258.

Broca, P. (1861c). Remarques sur le siège de la facultè du langage articulé, suivies d'une observation d'aphémie. *Bulletins de La Société d'anatomie*, *2e serie*(6), 330–357.

Broca, Paul. (1865). Sur le siège de la faculté du langage articulé. *Bulletins et Mémoires de la Société d'Anthropologie de Paris*, *6*(1), 377–393. https://doi.org/10.3406/bmsap.1865.9495

Brynhildsen, J. (2014). Combined hormonal contraceptives: Prescribing patterns, compliance, and benefits versus risks. *Therapeutic Advances in Drug Safety*, *5*(5), 201–213. https://doi.org/10.1177/2042098614548857

Buckner, R. L., Head, D., Parker, J., Fotenos, A. F., Marcus, D., Morris, J. C., & Snyder, A. Z. (2004). A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: Reliability and validation against manual measurement of total intracranial volume. *NeuroImage*, *23*(2), 724–738. https://doi.org/10.1016/j.neuroimage.2004.06.018

Craig. (2009). How do you feel — now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*(1), 59–70. https://doi.org/10.1038/nrn2555

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*(6), 215–222. https://doi.org/10.1016/S1364-6613(00)01483-2

Caceres, A., Hall, D. L., Zelaya, F. O., Williams, S. C. R., & Mehta, M. A. (2009). Measuring fMRI reliability with the intra-class correlation coefficient. *NeuroImage*, *45*(3), 758–768. https://doi.org/10.1016/j.neuroimage.2008.12.035

Cai, Q., Van der Haegen, L., & Brysbaert, M. (2013). Complementary hemispheric specialization for language production and visuospatial attention. *Proceedings of the National Academy*

of Sciences of the United States of America, *110*(4), E322-30. https://doi.org/10.1073/pnas.1212956110

Cammoun, D., Davis, K. A., & Hendee, W. R. (1985). Clinical Applications of Magnetic Resonance Imaging—Current Status. *Western Journal of Medicine*, *143*(6), 793–803.

Canto, J. G., Goldberg, R. J., Hand, M. M., Bonow, R. O., Sopko, G., Pepine, C. J., & Long, T. (2007). Symptom Presentation of Women With Acute Coronary Syndromes: Myth vs Reality. *Archives of Internal Medicine*, *167*(22), 2405–2413. https://doi.org/10.1001/archinte.167.22.2405

Canto, J. G., Rogers, W. J., Goldberg, R. J., Peterson, E. D., Wenger, N. K., Vaccarino, V., Kiefe, C. I., Frederick, P. D., Sopko, G., Zheng, Z.-J., & Investigators, for the N. (2012). Association of Age and Sex With Myocardial Infarction Symptom Presentation and In-Hospital Mortality. *JAMA*, *307*(8), 813–822. https://doi.org/10.1001/jama.2012.199

Canto, J. G., Shlipak, M. G., Rogers, W. J., Malmgren, J. A., Frederick, P. D., Lambrew, C. T., Ornato, J. P., Barron, H. V., & Kiefe, C. I. (2000). Prevalence, Clinical Characteristics, and Mortality Among Patients With Myocardial Infarction Presenting Without Chest Pain. *JAMA*, *283*(24), 3223–3229. https://doi.org/10.1001/jama.283.24.3223

Chiarello, C., & Maxfield, L. (1996). Varieties of interhemispheric inhibition, or how to keep a good hemisphere down. *Brain and Cognition*, *30*(1), 81–108. https://doi.org/10.1006/brcg.1996.0006

Cicek, M., Deouell, L. Y., & Knight, R. T. (2009a). Brain activity during landmark and line bisection tasks. *Frontiers in Human Neuroscience*, *3*, 7. https://doi.org/10.3389/neuro.09.007.2009

Confer, J. C., Easton, J. A., Fleischman, D. S., Goetz, C. D., Lewis, D. M. G., Perilloux, C., & Buss, D. M. (2010). Evolutionary psychology. Controversies, questions, prospects, and limitations. *The American Psychologist*, *65*(2), 110–126. https://doi.org/10.1037/a0018413

Cosgrove, K. P., Mazure, C. M., & Staley, J. K. (2007). Evolving Knowledge of Sex Differences in Brain Structure, Function, and Chemistry. *Biological Psychiatry*, *62*(8), 847–855. https://doi.org/10.1016/j.biopsych.2007.03.001

Craig, A. D. (Bud). (2010). The sentient self. *Brain Structure and Function*, *214*(5), 563–577. https://doi.org/10.1007/s00429-010-0248-y

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

Dalton, M. E. (1981). Sex hormone-binding globulin concentrations in women with severe premenstrual syndrome. *Postgraduate Medical Journal*, *57*(671), 560–561.

Damasio, A., Damasio, H., & Tranel, D. (2013). Persistence of Feelings and Sentience after Bilateral Damage of the Insula. *Cerebral Cortex (New York, NY)*, *23*(4), 833–846. https://doi.org/10.1093/cercor/bhs077

Davison Ankney, C. (1992). Sex differences in relative brain size: The mismeasure of woman, too? *Intelligence*, *16*(3–4), 329–336. https://doi.org/10.1016/0160-2896(92)90013-H

De Bondt, T, Jacquemyn, Y., Van Hecke, W., Sijbers, J., Sunaert, S., & Parizel, P. M. (2013a). Regional gray matter volume differences and sex-hormone correlations as a function of menstrual cycle phase and hormonal contraceptives use. *Brain Research*, *1530*, 22–31. https://doi.org/10.1016/j.brainres.2013.07.034

De Bondt, T., Pullens, P., Van Hecke, W., Jacquemyn, Y., & Parizel, P. M. (2016). Reproducibility of hormone-driven regional grey matter volume changes in women using SPM8 and SPM12. *Brain Structure and Function*, *221*(9), 4631–4641. https://doi.org/10.1007/s00429-016-1193-1

De Bondt, T, Van Hecke, W., Veraart, J., Leemans, A., Sijbers, J., Sunaert, S., Jacquemyn, Y., & Parizel, P. M. (2013b). Does the use of hormonal contraceptives cause microstructural changes in cerebral white matter? Preliminary results of a DTI and tractography study. *European Radiology*, *23*(1), 57–64. https://doi.org/10.1007/s00330-012-2572-5

Deppe, M., Knecht, S., Papke, K., Lohmann, H., Fleischer, H., Heindel, W., Ringelstein, E. B., & Henningsen, H. (2000). Assessment of hemispheric language lateralization: A comparison between fMRI and fTCD. *Journal of Cerebral Blood Flow and Metabolism : Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *20*(2), 263–268. https://doi.org/10.1097/00004647-200002000-00006

Derntl, B., Hack, R. L., Kryspin-Exner, I., & Habel, U. (2013). Association of menstrual cycle phase with the core components of empathy. *Hormones and Behavior*, *63*(1), 97–104. https://doi.org/10.1016/j.yhbeh.2012.10.009

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Djerassi, C. (1992). Steroid research at Syntex: "The Pill" and cortisone. *Steroids*, *57*(12), 631–641. https://doi.org/10.1016/0039-128X(92)90016-3

Djerassi, C. (2006). Chemical birth of the pill. *American Journal of Obstetrics and Gynecology*, *194*(1), 290–298. https://doi.org/10.1016/j.ajog.2005.06.010

Dorst, J., Haag, A., Knake, S., Oertel, W. H., Hamer, H. M., & Rosenow, F. (2008). Functional transcranial Doppler sonography and a spatial orientation paradigm identify the non-dominant hemisphere. *Brain and Cognition*, *68*(1), 53–58. https://doi.org/10.1016/j.bandc.2008.02.123

Drevets, W. C., Savitz, J., & Trimble, M. (2008). The subgenual anterior cingulate cortex in mood disorders. *CNS Spectrums*, *13*(8), 663–681. https://doi.org/10.1017/s1092852900013754

Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy assessment methods. *The Quarterly Journal of Experimental Psychology*, *60*(9), 1227–1245. https://doi.org/10.1080/17470210600926075

Edelman, R. R. (2014). The History of MR Imaging as Seen through the Pages of Radiology. *Radiology*, *273*(2S), S181–S200. https://doi.org/10.1148/radiol.14140706

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, *25*(4), 1325–1335. https://doi.org/10.1016/j.neuroimage.2004.12.034

Elliott, M. L., Knodt, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., Sison, M. L., Moffitt, T. E., Caspi, A., & Hariri, A. R. (2020). What is the test-retest reliability of common task-fMRI measures? New empirical evidence and a meta-analysis. *BioRxiv*, 681700. https://doi.org/10.1101/681700

Endrikat, J., Blode, H., Gerlinger, C., Rosenbaum, P., & Kuhnz, W. (2002). A pharmacokinetic study with a low-dose oral contraceptive containing 20 µg ethinylestradiol plus 100 µg

levonorgestrel. *The European Journal of Contraception & Reproductive Health Care*, *7*(2), 79–90. https://doi.org/10.1080/ejc.7.2.79.90

Ertman, N., Andreano, J. M., & Cahill, L. (2011). Progesterone at encoding predicts subsequent emotional memory. *Learning & Memory*, *18*(12), 759–763. https://doi.org/10.1101/lm.023267.111

Fernández, G, Specht, K., Weis, S., Tendolkar, I., Reuber, M., Fell, J., Klaver, P., Ruhlmann, J., Reul, J., & Elger, C. E. (2003). Intrasubject reproducibility of presurgical language lateralization and mapping using fMRI. *Neurology*, *60*(6), 969–975. https://doi.org/10.1212/01.WNL.0000049934.34209.2E

Fernández, de Greiff, A., von Oertzen, J., Reuber, M., Lun, S., Klaver, P., Ruhlmann, J., Reul, J., & Elger, C. E. (2001). Language Mapping in Less Than 15 Minutes: Real-Time Functional MRI during Routine Clinical Investigation. *NeuroImage*, *14*(3), 585–594. https://doi.org/10.1006/nimg.2001.0854

Ferrante, T., Castellini, P., Abrignani, G., Latte, L., Russo, M., Camarda, C., Veronesi, L., Pasquarella, C., Manzoni, G. C., & Torelli, P. (2012). The PACE study: Past-year prevalence of migraine in Parma's adult general population. *Cephalalgia: An International Journal of Headache*, *32*(5), 358–365. https://doi.org/10.1177/0333102411434811

Filipek, P. A., Richelme, C., Kennedy, D. N., & Caviness, V. S. (1994). The Young Adult Human Brain: An MRI-based Morphometric Analysis. *Cerebral Cortex*, *4*(4), 344–360. https://doi.org/10.1093/cercor/4.4.344

Fillmer, A., Kuehne, A., Goeschel, L., Köbe, T., Flöel, A., & Ittermann, B. (2018, June 16). Brain Tissue Segmentation and Subcortical Parcellation: How Reliable Are Different Tools? *SMRT27th Annual Meeting*. Joint Annual Meeting ISMRM-ESMRMB, Paris, France. http://archive.ismrm.org/2018/0698.html

Fink, G. R., Marshall, J. C., Shah, N. J., Weiss, P. H., Halligan, P. W., Grosse-Ruyken, M., Ziemons, K., Zilles, K., & Freund, H. J. (2000). Line bisection judgments implicate right parietal cortex and cerebellum as assessed by fMRI. *Neurology*, *54*(6), 1324–31.

Fischl, B., Liu, A., & Dale, A. M. (2001). Automated manifold surgery: Constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Medical Imaging*, *20*(1), 70–80.

Fischl, & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(20), 11050–11055.

Fischl, Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., & Dale, A. M. (2002). Whole Brain Segmentation. *Neuron*, *33*(3), 341–355. https://doi.org/10.1016/S0896-6273(02)00569-X

Fischl, Salat, D. H., Kouwe, A. J. W. van der, Makris, N., Ségonne, F., Quinn, B. T., & Dale, A. M. (2004). Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, *23*(Supplement 1), S69–S84. https://doi.org/DOI: 10.1016/j.neuroimage.2004.07.016

Fischl, Sereno, M. I., & Dale, A. M. (1999). Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System. *NeuroImage*, *9*(2), 195–207. https://doi.org/10.1006/nimg.1998.0396

Fischl , van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., & Dale, A. M. (2004). Automatically Parcellating the Human Cerebral Cortex. *Cerebral Cortex*, *14*(1), 11–22. https://doi.org/10.1093/cercor/bhg087

Fleischman, D. S., Navarrete, C. D., & Fessler, D. M. T. (2010). Oral Contraceptives Suppress Ovarian Hormone Production. *Psychological Science*, *21*(5), 750–752. https://doi.org/10.1177/0956797610368062

Fliessbach, K., Rohe, T., Linder, N. S., Trautner, P., Elger, C. E., & Weber, B. (2010). Retest reliability of reward-related BOLD signals. *NeuroImage*, *50*(3), 1168–1176. https://doi.org/10.1016/j.neuroimage.2010.01.036

Flöel, A., Jansen, A., Deppe, M., Kanowski, M., Konrad, C., Sommer, J., & Knecht, S. (2005). Atypical hemispheric dominance for attention: Functional MRI topography. *Journal of Cerebral Blood Flow & Metabolism*, *25*(9), 1197–1208. https://doi.org/10.1038/sj.jcbfm.9600114

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression.: Vol. Third edition*. Sage, Thousand Oaks CA. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Frank, R. T. (1931). THE HORMONAL CAUSES OF PREMENSTRUAL TENSION. *Archives of Neurology & Psychiatry*, *26*(5), 1053–1057. https://doi.org/10.1001/archneurpsyc.1931.02230110151009

Frässle, S., Paulus, F. M., Krach, S., & Jansen, A. (2016). Test-retest reliability of effective connectivity in the face perception network. *Human Brain Mapping*, *37*(2), 730–744. https://doi.org/10.1002/hbm.23061

Frässle, S., Paulus, F. M., Krach, S., Schweinberger, S. R., Stephan, K. E., & Jansen, A. (2015). Mechanisms of hemispheric lateralization: Asymmetric interhemispheric recruitment in the face perception network. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2015.09.055

Frässle, S., Stephan, K. E., Friston, K. J., Steup, M., Krach, S., Paulus, F. M., & Jansen, A. (2015). Test-retest reliability of dynamic causal modeling for fMRI. *NeuroImage*, *117*, 56–66. https://doi.org/10.1016/j.neuroimage.2015.05.040

Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S., & Turner, R. (1995). Analysis of fMRI time-series revisited. *NeuroImage*, *2*(1), 45–53. https://doi.org/10.1006/nimg.1995.1007

Frost, J. A. (1999). Language processing is strongly left lateralized in both sexes: Evidence from functional MRI. *Brain*, *122*(2), 199–208. https://doi.org/10.1093/brain/122.2.199

Galea, L. a. M., Wainwright, S. R., Roes, M. M., Duarte-Guterman, P., Chow, C., & Hamson, D. K. (2013). Sex, Hormones and Neurogenesis in the Hippocampus: Hormonal Modulation of Neurogenesis and Potential Functional Implications. *Journal of Neuroendocrinology*, *25*(11), 1039–1061. https://doi.org/10.1111/jne.12070

Gemmati, D., Varani, K., Bramanti, B., Piva, R., Bonaccorsi, G., Trentini, A., Manfrinato, M. C., Tisato, V., Carè, A., & Bellini, T. (2020). “Bridging the Gap” Everything that Could Have Been Avoided If We Had Applied Gender Medicine, Pharmacogenetics and Personalized Medicine in the Gender-Omics and Sex-Omics Era. *International Journal of Molecular Sciences*, *21*(1), 296. https://doi.org/10.3390/ijms21010296

Giatti, S., Boraso, M., Melcangi, R. C., & Viviani, B. (2012). Neuroactive steroids, their metabolites, and neuroinflammation. *Journal of Molecular Endocrinology*, *49*(3), R125–R134. https://doi.org/10.1530/JME-12-0127

Gingnell, M., Engman, J., Frick, A., Moby, L., Wikström, J., Fredrikson, M., & Sundström-Poromaa, I. (2013). Oral contraceptive use changes brain activity and mood in women with previous negative affect on the pill—A double-blinded, placebo-controlled randomized trial of a levonorgestrel-containing combined oral contraceptive. *Psychoneuroendocrinology*, *38*(7), 1133–1144. https://doi.org/10.1016/j.psyneuen.2012.11.006

Glatard, T., Lewis, L. B., Ferreira da Silva, R., Adalat, R., Beck, N., Lepage, C., Rioux, P., Rousseau, M.-E., Sherif, T., Deelman, E., Khalili-Mahani, N., & Evans, A. C. (2015). Reproducibility of neuroimaging analyses across operating systems. *Frontiers in Neuroinformatics*, *9*. https://doi.org/10.3389/fninf.2015.00012

Gogos, A., Wu, Y. C., Williams, A. S., & Byrne, L. K. (2014). The Effects of Ethinylestradiol and Progestins ("the pill") on Cognitive Function in Pre-menopausal Women. *Neurochemical Research*, *39*(12), 2288–2300. https://doi.org/10.1007/s11064-014-1444-6

Goldstein, J. M., Seidman, L. J., Horton, N. J., Makris, N., Kennedy, D. N., Caviness, V. S., Faraone, S. V., & Tsuang, M. T. (2001a). Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cerebral Cortex (New York, N.Y. : 1991)*, *11*(6), 490–497.

Goldzieher, J. W. (1993). Pharmacology of Contraceptive Steroids. In D. Shoupe & F. P. Haseltine (Eds.), *Contraception* (pp. 17–24). Springer. https://doi.org/10.1007/978-1-4612-2730-4_2

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Friston, K. J., & Frackowiak, R. S. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*, *14*(1 Pt 1), 21–36. https://doi.org/10.1006/nimg.2001.0786

Gorgolewski, K. J., Storkey, A. J., Bastin, M. E., Whittle, I., & Pernet, C. (2013). Single subject fMRI test-retest reliability metrics and confounding factors. *NeuroImage*, *69*, 231–243. https://doi.org/10.1016/j.neuroimage.2012.10.085

Guerra, J. A., López-Muñoz, F., & Álamo, C. (2013). Progestins in Combined Contraceptives. *Journal of Experimental & Clinical Medicine*, *5*(2), 51–55. https://doi.org/10.1016/j.jecm.2013.02.007

Guillamon, A., Junque, C., & Gómez-Gil, E. (2016). A Review of the Status of Brain Structure Research in Transsexualism. *Archives of Sexual Behavior*, *45*(7), 1615–1648. https://doi.org/10.1007/s10508-016-0768-5

Guo, C., Ferreira, D., Fink, K., Westman, E., & Granberg, T. (2019). Repeatability and reproducibility of FreeSurfer, FSL-SIENAX and SPM brain volumetric measurements and the effect of lesion filling in multiple sclerosis. *European Radiology*, *29*(3), 1355–1364. https://doi.org/10.1007/s00330-018-5710-x

Gur, R. C., Turetsky, B. I., Matsui, M., Yan, M., Bilker, W., Hughett, P., & Gur, R. E. (1999). Sex differences in brain gray and white matter in healthy young adults: Correlations with cognitive performance. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(10), 4065–4072.

Gurvich, C., Warren, A. M., Worsley, R., Hudaib, A.-R., Thomas, N., & Kulkarni, J. (2020). Effects of Oral Contraceptive Androgenicity on Visuospatial and Social-Emotional Cognition: A Prospective Observational Trial. *Brain Sciences*, *10*(4), 194. https://doi.org/10.3390/brainsci10040194

Hadland, K. A., Rushworth, M. F. S., Gaffan, D., & Passingham, R. E. (2003). The effect of cingulate lesions on social behaviour and emotion. *Neuropsychologia*, *41*(8), 919–931. https://doi.org/10.1016/S0028-3932(02)00325-1

Hagemann, G., Ugur, T., Schleussner, E., Mentzel, H.-J., Fitzek, C., Witte, O. W., & Gaser, C. (2011). Changes in brain size during the menstrual cycle. *PloS One*, *6*(2), e14655–e14655. https://doi.org/10.1371/journal.pone.0014655

Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., Busa, E., Pacheco, J., Albert, M., Killiany, R., Maguire, P., Rosas, D., Makris, N., Dale, A., Dickerson, B., & Fischl, B. (2006). Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *NeuroImage*, *32*(1), 180–194.

Handelsman, D. J., Hirschberg, A. L., & Bermon, S. (2018). Circulating Testosterone as the Hormonal Basis of Sex Differences in Athletic Performance. *Endocrine Reviews*, *39*(5), 803–829. https://doi.org/10.1210/er.2018-00020

Hao, J., Rapp, P. R., Leffler, A. E., Leffler, S. R., Janssen, W. G. M., Lou, W., McKay, H., Roberts, J. A., Wearne, S. L., Hof, P. R., & Morrison, J. H. (2006). Estrogen Alters Spine Number

and Morphology in Prefrontal Cortex of Aged Female Rhesus Monkeys. *Journal of Neuroscience*, *26*(9), 2571–2578. https://doi.org/10.1523/JNEUROSCI.3440-05.2006

Harrington, G. S., Tomaszewski Farias, S., Buonocore, M. H., & Yonelinas, A. P. (2006). The intersubject and intrasubject reproducibility of FMRI activation during three encoding tasks: Implications for clinical applications. *Neuroradiology*, *48*(7), 495–505. https://doi.org/10.1007/s00234-006-0083-2

Hattemer, K., Plate, A., Heverhagen, J. T., Haag, A., Keil, B., Klein, K. M., Hermsen, A., Oertel, W. H., Hamer, H. M., Rosenow, F., & Knake, S. (2011a). Determination of hemispheric dominance with mental rotation using functional transcranial Doppler sonography and FMRI. *Journal of Neuroimaging : Official Journal of the American Society of Neuroimaging*, *21*(1), 16–23. https://doi.org/10.1111/j.1552-6569.2009.00402.x

Hausmann, M, & Güntürkün, O. (1999). Sex differences in functional cerebral asymmetries in a repeated measures design. *Brain and Cognition*, *41*(3), 263–275. https://doi.org/10.1006/brcg.1999.1126

Hausmann, M. (2005). Hemispheric asymmetry in spatial attention across the menstrual cycle. *Neuropsychologia*, *43*(11), 1559–1567. https://doi.org/10.1016/j.neuropsychologia.2005.01.017

Hausmann, M.. (2017). Why sex hormones matter for neuroscience: A very short review on sex, sex hormones, and functional brain asymmetries. *Journal of Neuroscience Research*, *95*(1–2), 40–49. https://doi.org/10.1002/jnr.23857

Hausmann, M., & Güntürkün, O. (2000). Steroid fluctuations modify functional cerebral asymmetries: The hypothesis of progesterone-mediated interhemispheric decoupling. *Neuropsychologia*, *38*(10), 1362–1374. https://doi.org/10.1016/S0028-3932(00)00045-2

Herting, M. M., Gautam, P., Chen, Z., Mezher, A., & Vetter, N. C. (2018). Test-retest reliability of longitudinal task-based fMRI: Implications for developmental studies. *Developmental Cognitive Neuroscience*, *33*, 17–26. https://doi.org/10.1016/j.dcn.2017.07.001

Heßling, A., & Bundeszentrale für Gesundheitliche Aufklärung (Eds.). (2011). *Verhütungsverhalten Erwachsener 2011: Aktuelle repräsentative Studie im Rahmen einer telefonischen Mehrthemenbefragung* (1. Aufl.). BZgA.

Hines, M. (2020). Neuroscience and Sex/Gender: Looking Back and Forward. *The Journal of Neuroscience*, *40*(1), 37–43. https://doi.org/10.1523/JNEUROSCI.0750-19.2019

Hinshaw, W. S., Bottomley, P. A., & Holland, G. N. (1977). Radiographic thin-section image of the human wrist by nuclear magnetic resonance. *Nature*, *270*(5639), 722–723. https://doi.org/10.1038/270722a0

Hoekzema, E., Barba-Müller, E., Pozzobon, C., Picado, M., Lucco, F., García-García, D., Soliva, J. C., Tobeña, A., Desco, M., Crone, E. A., Ballesteros, A., Carmona, S., & Vilarroya, O. (2017). Pregnancy leads to long-lasting changes in human brain structure. *Nature Neuroscience*, *20*(2), 287–296. https://doi.org/10.1038/nn.4458

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*, *6*(6), 435. https://doi.org/10.1038/nrn1684

Hughes, C. M., & Peters, A. (1992). Symmetric synapses formed by callosal afferents in rat visual cortex. *Brain Research*, *583*(1), 271–278. https://doi.org/10.1016/S0006-8993(10)80033-2

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, *65*, 373–398. https://doi.org/10.1146/annurev-psych-010213-115057

Jansen, A, Menke, R., Sommer, J., Förster, A. F., Bruchmann, S., Hempleman, J., Weber, B., & Knecht, S. (2006). The assessment of hemispheric lateralization in functional MRI--robustness and reproducibility. *NeuroImage*, *33*(1), 204–217.

Jansen, A., Flöel, A., Deppe, M., van Randenborgh, J., Dräger, B., Kanowski, M., & Knecht, S. (2004). Determining the hemispheric dominance of spatial attention: A comparison between fTCD and fMRI. *Human Brain Mapping*, *23*(3), 168–180. https://doi.org/10.1002/hbm.20055

Jansen, A., Flöel, A., Menke, R., Kanowski, M., & Knecht, S. (2005a). Dominance for language and spatial processing: Limited capacity of a single hemisphere. *Neuroreport*, *16*(9), 1017–1021.

Jansen, A, Flöel, A., Van Randenborgh, J., Konrad, C., Rotte, M., Förster, A.-F., Deppe, M., & Knecht, S. (2005b). Crossed cerebro-cerebellar language dominance. *Human Brain Mapping*, *24*(3), 165–172. https://doi.org/10.1002/hbm.20077

Jansen, A, Lohmann, H., Scharfe, S., Sehlmeyer, C., Deppe, M., & Knecht, S. (2007). The association between scalp hair-whorl direction, handedness and hemispheric language

dominance: Is there a common genetic basis of lateralization? *NeuroImage*, *35*(2), 853–861. https://doi.org/10.1016/j.neuroimage.2006.12.025

Johnson, P. A., Fitzgerald, Therese, Glynn, Amy, Salganicoff, Alina, Wood, Susan F., & Goldstein, Jill M. (2016). *Precision Medicine—How Sex and Gender Drive Innovation*. https://www.tbf.org/~/media/TBFOrg/Files/Reports/ConnorsCenterBrochure-2016.pdf2016

Jou, R. J., Minshew, N. J., Keshavan, M. S., Vitale, M. P., & Hardan, A. Y. (2010). Enlarged right superior temporal gyrus in children and adolescents with autism. *Brain Research*, *1360*, 205–212. https://doi.org/10.1016/j.brainres.2010.09.005

Jovicich, J., Czanner, S., Greve, D., Haley, E., Kouwe, A. van der, Gollub, R., Kennedy, D., Schmitt, F., Brown, G., MacFall, J., Fischl, B., & Dale, A. (2006). Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *NeuroImage*, *30*(2), 436–443. https://doi.org/DOI: 10.1016/j.neuroimage.2005.09.046

Juch, H., Zimine, I., Seghier, M. L., Lazeyras, F., & Fasel, J. H. D. (2005). Anatomical variability of the lateral frontal lobe surface: Implication for intersubject variability in language neuroimaging. *NeuroImage*, *24*(2), 504–514. https://doi.org/10.1016/j.neuroimage.2004.08.037

Kammoun, I., Ben Saâda, W., Sifaou, A., Haouat, E., Kandara, H., Ben Salem, L., & Ben Slama, C. (2017). Change in women's eating habits during the menstrual cycle. *Annales d'Endocrinologie*, *78*(1), 33–37. https://doi.org/10.1016/j.ando.2016.07.001

Karnath, H.-O. (2001). New insights into the functions of the superior temporal cortex. *Nature Reviews Neuroscience*, *2*(8), 568–576. https://doi.org/10.1038/35086057

Kautzky-Willer. (2012). Sex and gender differences in endocrinology. In *Sex and Gender Aspects in Clinical Medicine* (pp. 125–149). London, UK: Springer.

Kawaguchi, Y. (1992). Receptor subtypes involved in callosally-induced postsynaptic potentials in rat frontal agranular cortex in vitro. *Experimental Brain Research*, *88*(1), 33–40. https://doi.org/10.1007/BF02259126

Khan, M. M., Dhandapani, K. M., Zhang, Q.-G., & Brann, D. W. (2013). Estrogen regulation of spine density and excitatory synapses in rat prefrontal and somatosensory cerebral cortex. *Steroids*, *78*(6), 614–623. https://doi.org/10.1016/j.steroids.2012.12.005

Kiar, G., de Oliveira Castro, P., Rioux, P., Petit, E., Brown, S. T., Evans, A. C., & Glatard, T. (2020). Comparing perturbation models for evaluating stability of neuroimaging pipelines. *The International Journal of High Performance Computing Applications*, *34*(5), 491–501. https://doi.org/10.1177/1094342020926237

Kinsbourne, M., & Bruce, R. (1987a). Shift in visual laterality within blocks of trials. *Acta Psychologica*, *66*(2), 139–55.

Phan, K.L., Wager, T., Taylor, S.T., Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*, *16*(2), 331–348. https://doi.org/10.1006/nimg.2002.1087

Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M.-C., Christensen, G. E., Collins, D. L., Gee, J., Hellier, P., Song, J. H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R. P., Mann, J. J., & Parsey, R. V. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*, *46*(3), 786–802. https://doi.org/10.1016/j.neuroimage.2008.12.037

Knecht, S., Deppe, M., Dräger, B., Bobe, L., Lohmann, H., Ringelstein, E., & Henningsen, H. (2000). Language lateralization in healthy right-handers. *Brain : A Journal of Neurology*, *123 (Pt 1*, 74–81. https://doi.org/10.1093/brain/123.1.74

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., Ringelstein, E. B., & Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. In *Brain: A journal of neurology: Vol. 123 Pt 12*. https://doi.org/10.1093/brain/123.12.2512

Krentzel, A. A., Barrett, L. R., & Meitzen, J. (2019). Estradiol rapidly modulates excitatory synapse properties in a sex- and region-specific manner in rat nucleus accumbens core and caudate-putamen. *Journal of Neurophysiology*, *122*(3), 1213–1225. https://doi.org/10.1152/jn.00264.2019

Kristo, G., Rutten, G.-J., Raemaekers, M., de Gelder, B., Rombouts, S. A. R. B., & Ramsey, N. F. (2014). Task and task-free FMRI reproducibility comparison for motor network identification. *Human Brain Mapping*, *35*(1), 340–352. https://doi.org/10.1002/hbm.22180

Kudwa, A. E., McGivern, R. F., & Handa, R. J. (2014). Estrogen Receptor β and Oxytocin Interact to Modulate Anxiety-like Behavior and Neuroendocrine Stress Reactivity in Adult Male

and Female Rats. *Physiology & Behavior*, *129*, 287–296. https://doi.org/10.1016/j.physbeh.2014.03.004

Kulkarni, J. (2007). Depression as a side effect of the contraceptive pill. *Expert Opinion on Drug Safety*, *6*(4), 371–374. https://doi.org/10.1517/14740338.6.4.371

Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., & Eickhoff, S. B. (2010). A link between the systems: Functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Structure and Function*, *214*(5), 519–534. https://doi.org/10.1007/s00429-010-0255-z

Lamm, C., & Singer, T. (2010). The role of anterior insular cortex in social emotions. *Brain Structure & Function*, *214*(5–6), 579–591. https://doi.org/10.1007/s00429-010-0251-3

Lawrence, R. M., Bridgeford, E. W., Myers, P. E., Arvapalli, G. C., Ramachandran, S. C., Pisner, D. A., Frank, P. F., Lemmer, A. D., Nikolaidis, A., & Vogelstein, J. T. (2020). Standardizing Human Brain Parcellations. *BioRxiv*, 845065. https://doi.org/10.1101/845065

Lenth, R., Buerkner, P., Herve, M., Love, J., Riebl, H., & Singmann, H. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (1.5.1) [Computer software]. https://CRAN.R-project.org/package=emmeans

Lentini, E., Kasahara, M., Arver, S., & Savic, I. (2012). Sex Differences in the Human Brain and the Impact of Sex Chromosomes and Sex Hormones. *Cerebral Cortex*, *23*(10), 2322–2336. https://doi.org/10.1093/cercor/bhs222

Leonard, C. M., Towler, S., Welcome, S., Halderman, L. K., Otto, R., Eckert, M. A., & Chiarello, C. (2008). Size matters: Cerebral volume influences sex differences in neuroanatomy. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(12), 2920–2931. https://doi.org/10.1093/cercor/bhn052

Leranth, C., & Shanabrough, M. (2001). Supramammillary Area Mediates Subcortical Estrogenic Action on Hippocampal Synaptic Plasticity. *Experimental Neurology*, *167*(2), 445–450. https://doi.org/10.1006/exnr.2000.7585

Lewis, C. A., Kimmig, A.-C. S., Zsido, R. G., Jank, A., Derntl, B., & Sacher, J. (2019). Effects of Hormonal Contraceptives on Mood: A Focus on Emotion Recognition and Reactivity, Reward Processing, and Stress Response. *Current Psychiatry Reports*, *21*(11), 115. https://doi.org/10.1007/s11920-019-1095-z

Lewis, L., Lepage, C., Khalili-Mahani, N., & et al. (2017). *Robust- ness and reliability of cortical surface reconstruction in CIVET and FreeSurfer.* Annual Meeting of the Organization for Human Brain Mapping, Vancouver. http://www.bic.mni.mcgill.ca/users/llewis/CIVET_vs_FS_HBM2017_final.pdf

Liao, P. V., & Dollin, J. (2012). Half a century of the oral contraceptive pill. *Canadian Family Physician*, *58*(12), e757–e760.

Lidzba, K., Staudt, M., Wilke, M., Grodd, W., & Krägeloh-Mann, I. (2006). Lesion-induced right-hemispheric language and organization of nonverbal functions. *Neuroreport*, *17*(9), 929–933. https://doi.org/10.1097/01.wnr.0000221841.12632.d6

Lisofsky, N., Mårtensson, J., Eckert, A., Lindenberger, U., Gallinat, J., & Kühn, S. (2015). Hippocampal volume and functional connectivity changes during the female menstrual cycle. *NeuroImage*, *118*, 154–162. https://doi.org/10.1016/j.neuroimage.2015.06.012

Liu, K. A., & Mager, N. A. D. (2016). Women's involvement in clinical trials: Historical perspective and future implications. *Pharmacy Practice*, *14*(1). https://doi.org/10.18549/PharmPract.2016.01.708

Lj, E., Ke, M., Rd, W., Bf, Q., & O, S. (2008). An R2 statistic for fixed effects in the linear mixed model. *Statistics in Medicine*, *27*(29), 6137–6157. https://doi.org/10.1002/sim.3429

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–157. https://doi.org/10.1038/35084005

Lohmann, H., Deppe, M., Jansen, A., Schwindt, W., & Knecht, S. (2004). Task repetition can affect functional magnetic resonance imaging-based measures of language lateralization and lead to pseudoincreases in bilaterality. *Journal of Cerebral Blood Flow and Metabolism : Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *24*(2), 179–187. https://doi.org/10.1097/01.WCB.0000100066.36077.91

Lord, C., Buss, C., Lupien, S. J., & Pruessner, J. C. (2008). Hippocampal volumes are larger in postmenopausal women using estrogen therapy compared to past users, never users and men: A possible window of opportunity effect. *Neurobiology of Aging*, *29*(1), 95–101. https://doi.org/10.1016/j.neurobiolaging.2006.09.001

Luders, E., Gaser, C., Narr, K. L., & Toga, A. W. (2009). Why sex matters: Brain size independent differences in gray matter distributions between men and women. *The Journal of*

Neuroscience : The Official Journal of the Society for Neuroscience, *29*(45), 14265–14270. https://doi.org/10.1523/JNEUROSCI.2261-09.2009

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. *Journal of Behavioral and Brain Science*, *2*(3), 357–362. https://doi.org/10.4236/jbbs.2012.23040

Ludwig, H. (2011). Lokale Spuren eines Jahrhunderts. *Der Gynäkologe*, *44*(1), 64–70. https://doi.org/10.1007/s00129-010-2667-5

MacLaughlin, D. T., & Donahoe, P. K. (2004). Sex Determination and Differentiation. *New England Journal of Medicine*, *350*(4), 367–378. https://doi.org/10.1056/NEJMra022784

Macrae, C. N., Alnwick, K. A., Milne, A. B., & Schloerscheidt, A. M. (2002). Person perception across the menstrual cycle: Hormonal influences on social-cognitive functioning. *Psychological Science*, *13*(6), 532–536. https://doi.org/10.1111/1467-9280.00493

Maki, P. M., Rich, J. B., & Shayna Rosenbaum, R. (2002). Implicit memory varies across the menstrual cycle: Estrogen effects in young women. *Neuropsychologia*, *40*(5), 518–529. https://doi.org/10.1016/S0028-3932(01)00126-9

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*(3), 1233–1239. https://doi.org/10.1016/S1053-8119(03)00169-1

Marecková, K., Perrin, J. S., Nawaz Khan, I., Lawrence, C., Dickie, E., McQuiggan, D. A., & Paus, T. (2014). Hormonal contraceptives, menstrual cycle and brain response to faces. *Social Cognitive and Affective Neuroscience*, *9*(2), 191–200. https://doi.org/10.1093/scan/nss128

McEwen, B. S., & Woolley, C. S. (1994). Estradiol and progesterone regulate neuronal structure and synaptic connectivity in adult as well as developing brain. *Experimental Gerontology*, *29*(3), 431–436. https://doi.org/10.1016/0531-5565(94)90022-1

Merkatz, R. B., Temple, R., Sobel, S., Feiden, K., & Kessler, D. A. (1993). Women in Clinical Trials of New Drugs – A Change in Food and Drug Administration Policy. *New England Journal of Medicine*, *329*(4), 292–296. https://doi.org/10.1056/NEJM199307223290429

Mesulam, M. M. (1981). A cortical network for directed attention and unilateral neglect. *Annals of Neurology*, *10*(4), 309–325. https://doi.org/10.1002/ana.410100402

Mihm, M, Gangooly, S., & Muttukrishna, S. (2010). The normal menstrual cycle in women. *Animal Reproduction Science*, *124*, 229–236. https://doi.org/10.1016/j.anireprosci.2010.08.030

Miller, I. N., & Cronin-Golomb, A. (2010). Gender differences in Parkinson's disease: Clinical characteristics and cognition: Gender Differences in Parkinson's disease. *Movement Disorders*, *25*(16), 2695–2703. https://doi.org/10.1002/mds.23388

Miller, M. A. (2001). Gender-based differences in the toxicity of pharmaceuticals—The Food and Drug Administration's perspective. *International Journal of Toxicology*, *20*(3), 149–152. https://doi.org/10.1080/109158101317097728

Naftolin, F., Garcia-Segura, L. M., Horvath, T. L., Zsarnovszky, A., Demir, N., Fadiel, A., Leranth, C., Vondracek-Klepper, S., Lewis, C., Chang, A., & Parducz, A. (2007). Estrogen-Induced Hypothalamic Synaptic Plasticity and Pituitary Sensitization in the Control of the Estrogen-Induced Gonadotrophin Surge. *Reproductive Sciences*, *14*(2), 101–116. https://doi.org/10.1177/1933719107301059

Nakamura, K., Fox, R., & Fisher, E. (2011). CLADA: Cortical longitudinal atrophy detection algorithm. *NeuroImage*, *54*(1), 278–289. https://doi.org/10.1016/j.neuroimage.2010.07.052

Ngun, T. C., Ghahramani, N., Sánchez, F. J., Bocklandt, S., & Vilain, E. (2011). The Genetics of Sex Differences in Brain and Behavior. *Frontiers in Neuroendocrinology*, *32*(2), 227–246. https://doi.org/10.1016/j.yfrne.2010.10.001

Nopoulos, P., Flaum, M., O'Leary, D., & Andreasen, N. C. (2000). Sexual dimorphism in the human brain: Evaluation of tissue volume, tissue composition and surface anatomy using magnetic resonance imaging. *Psychiatry Research*, *98*(1), 1–13. https://doi.org/10.1016/s0925-4927(99)00044-x

Nord, C. L., Gray, A., Charpentier, C. J., Robinson, O. J., & Roiser, J. P. (2017). Unreliability of putative fMRI biomarkers during emotional face processing. *NeuroImage*, *156*, 119–127. https://doi.org/10.1016/j.neuroimage.2017.05.024

Ober, C., Loisel, D. A., & Gilad, Y. (2008). Sex-specific genetic architecture of human disease. *Nature Reviews Genetics*, *9*(12), 911–922. https://doi.org/10.1038/nrg2415

Ortigue, S., Grafton, S. T., & Bianchi-Demicheli, F. (2007). Correlation between insula activation and self-reported quality of orgasm in women. *NeuroImage*, *37*(2), 551–560. https://doi.org/10.1016/j.neuroimage.2007.05.026

Ossewaarde, L., van Wingen, G. A., Rijpkema, M., Bäckström, T., Hermans, E. J., & Fernández, G. (2011). Menstrual cycle-related changes in amygdala morphology are associated with changes in stress sensitivity. *Human Brain Mapping*, *34*(5), 1187–1193. https://doi.org/10.1002/hbm.21502

Österlund, M. K., & Hurd, Y. L. (2001). Estrogen receptors in the human forebrain and the relation to neuropsychiatric disorders. *Progress in Neurobiology*, *64*(3), 251–267. https://doi.org/10.1016/S0301-0082(00)00059-9

Paus, T. (2010). Chapter 2 - Sex differences in the human brain: A developmental perspective. In I. Savic (Ed.), *Progress in Brain Research* (Vol. 186, pp. 13–28). Elsevier. https://doi.org/10.1016/B978-0-444-53630-3.00002-6

Petersen, N., Kilpatrick, L. A., Goharzad, A., & Cahill, L. (2014). Oral contraceptive pill use and menstrual cycle phase are associated with altered resting state functional connectivity. *NeuroImage*, *90*, 24–32. https://doi.org/10.1016/j.neuroimage.2013.12.016

Petersen, N., Touroutoglou, A., Andreano, J. M., & Cahill, L. (2015). Oral contraceptive pill use is associated with localized decreases in cortical thickness. *Human Brain Mapping*, *36*(7), 2644–2654. https://doi.org/10.1002/hbm.22797

Petitti, D. B. (2003). Combination Estrogen–Progestin Oral Contraceptives. *New England Journal of Medicine*, *349*(15), 1443–1450. https://doi.org/10.1056/NEJMcp030751

Pletzer, B, (2015b). Editorial: From sex differences in neuroscience to a neuroscience of sex differences: New directions and perspectives. *Frontiers in Neuroscience*, *9*, 330. https://doi.org/10.3389/fnins.2015.00330

Pletzer, B. (2019b). Sex Hormones and Gender Role Relate to Gray Matter Volumes in Sexually Dimorphic Brain Areas. *Frontiers in Neuroscience*, *13*. https://doi.org/10.3389/fnins.2019.00592

Pletzer, B., & Kerschbaum, H. H. (2014). 50 years of hormonal contraception-time to find out, what it does to our brain. *Frontiers in Neuroscience*, *8*, 256–256. https://doi.org/10.3389/fnins.2014.00256

Pletzer, B., Harris, T. A., & Hidalgo-Lopez, E. (2018). Subcortical structural changes along the menstrual cycle: Beyond the hippocampus. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-34247-4

Pletzer, B., Harris, T., & Hidalgo-Lopez, E. (2019a). Previous contraceptive treatment relates to grey matter volumes in the hippocampus and basal ganglia. *Scientific Reports*, *9*(1), 11003. https://doi.org/10.1038/s41598-019-47446-4

Pletzer, B., Kronbichler, M., Aichhorn, M., Bergmann, J., Ladurner, G., & Kerschbaum, H. H. (2010). Menstrual cycle and hormonal contraceptive use modulate human brain structure. *Brain Research*, *1348*, 55–62. https://doi.org/10.1016/j.brainres.2010.06.019

Pletzer, B., Kronbichler, M., & Kerschbaum, H. (2015a). Differential effects of androgenic and anti-androgenic progestins on fusiform and frontal gray matter volume and face recognition performance. *Brain Research*, *1596*, 108–115. https://doi.org/10.1016/j.brainres.2014.11.025

Pletzer, B., Kronbichler, M., Ladurner, G., Nuerk, H.-C., & Kerschbaum, H. (2011). Menstrual cycle variations in the BOLD-response to a number bisection task: Implications for research on sex differences. *Brain Research*, *1420*, 37–47. https://doi.org/10.1016/j.brainres.2011.08.058

Pletzer, B., Kronbichler, M., Nuerk, H.-C., & Kerschbaum, H. (2014). Hormonal contraceptives masculinize brain activation patterns in the absence of behavioral changes in two numerical tasks. *Brain Research*, *1543*, 128–142. https://doi.org/10.1016/j.brainres.2013.11.007

Pol, H. E. H., Cohen-Kettenis, P. T., Haren, N. E. M. V., Peper, J. S., Brans, R. G. H., Cahn, W., Schnack, H. G., Gooren, L. J. G., & Kahn, R. S. (2006). Changing your sex changes your brain: Influences of testosterone and estrogen on adult human brain structure. *European Journal of Endocrinology*, *155*(suppl_1), S107–S114. https://doi.org/10.1530/eje.1.02248

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., Nichols, T. E., Poline, J.-B., Vul, E., & Yarkoni, T. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, *18*(2), 115–126. https://doi.org/10.1038/nrn.2016.167

Popescu, V., Klaver, R., Versteeg, A., Voorn, P., Twisk, J. W. R., Barkhof, F., Geurts, J. J. G., & Vrenken, H. (2016). Postmortem validation of MRI cortical volume measurements in MS. *Human Brain Mapping*, *37*(6), 2223–2233. https://doi.org/10.1002/hbm.23168

Protopopescu, X., Butler, T., Pan, H., Root, J., Altemus, M., Polanecsky, M., McEwen, B., Silbersweig, D., & Stern, E. (2008a). Hippocampal structural changes across the menstrual cycle. *Hippocampus*, *18*(10), 985–988. https://doi.org/10.1002/hipo.20468

Publishing, H. H. (n.d.). *Testosterone, aging, and the mind*. Harvard Health. Retrieved August 4, 2019, from https://www.health.harvard.edu/newsletter_article/Testosterone_aging_and_the_mind

Pujol, J., Deus, J., Losilla, J. M., & Capdevila, A. (1999). Cerebral lateralization of language in normal left-handed people studied by functional MRI. *Neurology*, *52*(5), 1038–1043. https://doi.org/10.1212/WNL.52.5.1038

Raghava, N., Das, B. C., & Ray, S. K. (2017). Neuroprotective effects of estrogen in CNS injuries: Insights from animal models. *Neuroscience and Neuroeconomics*, *6*, 15–29. https://doi.org/10.2147/NAN.S105134

Regitz-Zagrosek, V. (2012). Sex and gender differences in health: Science & Society Series on Sex and Science. *EMBO Reports*, *13*(7), 596–603. https://doi.org/10.1038/embor.2012.87

Rehbein, E., Hornung, J., Sundström Poromaa, I., & Derntl, B. (2020). Shaping of the female human brain by sex hormones – a review. *Neuroendocrinology*. https://doi.org/10.1159/000507083

Reuter, M., & Fischl, B. (2011). Avoiding asymmetry-induced bias in longitudinal image processing. *NeuroImage*, *57*(1), 19–21. https://doi.org/10.1016/j.neuroimage.2011.02.076

Reuter, M., Rosas, H. D., & Fischl, B. (2010). Highly accurate inverse consistent registration: A robust approach. *NeuroImage*, *53*(4), 1181–1196. https://doi.org/10.1016/j.neuroimage.2010.07.020

Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, *61*(4), 1402–1418. https://doi.org/10.1016/j.neuroimage.2012.02.084

Righart, R., Schmidt, P., Dahnke, R., Biberacher, V., Beer, A., Buck, D., Hemmer, B., Kirschke, J. S., Zimmer, C., Gaser, C., & Mühlau, M. (2017). Volume versus surface-based cortical thickness measurements: A comparative study with healthy controls and multiple

sclerosis patients. *PLOS ONE*, *12*(7), e0179590. https://doi.org/10.1371/journal.pone.0179590

Robinson, S. A., Dowell, M., Pedulla, D., & McCauley, L. (2004). Do the emotional side-effects of hormonal contraceptives come from pharmacologic or psychological mechanisms? *Medical Hypotheses*, *63*(2), 268–273. https://doi.org/10.1016/j.mehy.2004.02.013

Rosas, H. D., Liu, A. K., Hersch, S., Glessner, M., Ferrante, R. J., Salat, D. H., van der Kouwe, A., Jenkins, B. G., Dale, A. M., & Fischl, B. (2002). Regional and progressive thinning of the cortical ribbon in Huntington's disease. *Neurology*, *58*(5), 695–701.

Ruigrok, A. N. V., Salimi-Khorshidi, G., Lai, M.-C., Baron-Cohen, S., Lombardo, M. V., Tait, R. J., & Suckling, J. (2014). A meta-analysis of sex differences in human brain structure. *Neuroscience and Biobehavioral Reviews*, *39*(100), 34–50. https://doi.org/10.1016/j.neubiorev.2013.12.004

Rumberg, B., Baars, A., Fiebach, J., Ladd, M. E., Forsting, M., Senf, W., & Gizewski, E. R. (2010). Cycle and gender-specific cerebral activation during a verb generation task using fMRI: Comparison of women in different cycle phases, under oral contraception, and men. *Neuroscience Research*, *66*(4), 366–371. https://doi.org/10.1016/j.neures.2009.12.011

Rupp, H. A., James, T. W., Ketterson, E. D., Sengelaub, D. R., Janssen, E., & Heiman, J. R. (2009). Neural activation in the orbitofrontal cortex in response to male faces increases during the follicular phase. *Hormones and Behavior*, *56*(1), 66–72. https://doi.org/10.1016/j.yhbeh.2009.03.005

Salli, E., Korvenoja, A., Visa, A., Katila, T., & Aronen, H. J. (2001). Reproducibility of fMRI: Effect of the Use of Contextual Information. *NeuroImage*, *13*(3), 459–471. https://doi.org/10.1006/nimg.2000.0702

Schöning, S., Engelien, A., Kugel, H., Schäfer, S., Schiffbauer, H., Zwitserlood, P., Pletziger, E., Beizai, P., Kersting, A., Ohrmann, P., Greb, R. R., Lehmann, W., Heindel, W., Arolt, V., & Konrad, C. (2007). Functional anatomy of visuo-spatial working memory during mental rotation is influenced by sex, menstrual cycle, and sex steroid hormones. *Neuropsychologia*, *45*(14), 3203–14. https://doi.org/10.1016/j.neuropsychologia.2007.06.011

Schuster, V., Herholz, P., Zimmermann, K. M., Westermann, S., Frässle, S., & Jansen, A. (2017). Comparison of fMRI paradigms assessing visuospatial processing: Robustness and

reproducibility. *PLOS ONE*, *12*(10), e0186344–e0186344. https://doi.org/10.1371/journal.pone.0186344

Segonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, *22*(3), 1060–1075. https://doi.org/DOI: 10.1016/j.neuroimage.2004.03.032

Seiger, R., Ganger, S., Kranz, G. S., Hahn, A., & Lanzenberger, R. (2018). Cortical Thickness Estimations of FreeSurfer and the CAT12 Toolbox in Patients with Alzheimer's Disease and Healthy Controls. *Journal of Neuroimaging*, *28*(5), 515–523. https://doi.org/10.1111/jon.12521

Shahnazi, M., Farshbaf Khalili, A., Ranjbar Kochaksaraei, F., Asghari Jafarabadi, M., Gaza Banoi, K., Nahaee, J., & Bayati Payan, S. (2014). A Comparison of Second and Third Generations Combined Oral Contraceptive Pills' Effect on Mood. *Iranian Red Crescent Medical Journal*, *16*(8). https://doi.org/10.5812/ircmj.13628

Shaywitz, B. A., Shaywltz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., Bronen, R. A., Fletcher, J. M., Shankweiler, D. P., Katz, L., & Gore, J. C. (1995). Sex differences in the functional organization of the brain for language. *Nature*, *373*(6515), 607. https://doi.org/10.1038/373607a0

Sheppard, P. A. S., Choleris, E., & Galea, L. A. M. (2019). Structural plasticity of the hippocampus in response to estrogens in female rodents. *Molecular Brain*, *12*(1), 22. https://doi.org/10.1186/s13041-019-0442-7

Shin, Y.-W., Kim, D. J., Ha, T. H., Park, H.-J., Moon, W.-J., Chung, E. C., Lee, J. M., Kim, I. Y., Kim, S. I., & Kwon, J. S. (2005). Sex differences in the human corpus callosum: Diffusion tensor imaging study. *Neuroreport*, *16*(8), 795–798.

Sitruk-Ware, R. (2006). New progestagens for contraceptive use. *Human Reproduction Update*, *12*(2), 169–178. https://doi.org/10.1093/humupd/dmi046

Sitruk-Ware, R., & Nath, A. (2013). Characteristics and metabolic effects of estrogen and progestins contained in oral contraceptive pills. *Best Practice & Research Clinical Endocrinology & Metabolism*, *27*(1), 13–24. https://doi.org/10.1016/j.beem.2012.09.004

Skovlund, C. W., Mørch, L. S., Kessing, L. V., & Lidegaard, Ø. (2016). Association of Hormonal Contraception With Depression. *JAMA Psychiatry*, *73*(11), 1154–1162. https://doi.org/10.1001/jamapsychiatry.2016.2387

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex (New York, N.Y.: 1991)*, *6*(1), 11–20. https://doi.org/10.1093/cercor/6.1.11

Smith, S. S., Waterhouse, B. D., Chapin, J. K., & Woodward, D. J. (1987). Progesterone alters GABA and glutamate responsiveness: A possible mechanism for its anxiolytic action. *Brain Research*, *400*(2), 353–359. https://doi.org/10.1016/0006-8993(87)90634-2

Sommer, B. (1973). The Effect of Menstruation on Cognitive and Perceptual-Motor Behavior: A Review. *Psychosomatic Medicine*, *35*(6), 515–534.

Springer, J. A., Binder, J. R., Hammeke, T. A., Swanson, S. J., Frost, J. A., Bellgowan, P. S., Brewer, C. C., Perry, H. M., Morris, G. L., & Mueller, W. M. (1999). Language dominance in neurologically normal and epilepsy subjects: A functional MRI study. *Brain : A Journal of Neurology*, *122 (Pt 1*, 2033–2046.

Srivastava, D. P., & Penzes, P. (2011). Rapid Estradiol Modulation of Neuronal Connectivity and Its Implications for Disease. *Frontiers in Endocrinology*, *2*. https://doi.org/10.3389/fendo.2011.00077

Staley, J. K., Krishnan-Sarin, S., Zoghbi, S., Tamagnan, G., Fujita, M., Seibyl, J. P., Maciejewski, P. K., O'Malley, S., & Innis, R. B. (2001). Sex differences in [123I]?-CIT SPECT measures of dopamine and serotonin transporter availability in healthy smokers and nonsmokers. *Synapse*, *41*(4), 275–284. https://doi.org/10.1002/syn.1084

Starling, E. H., & Royal College of Physicians of London. (1905). *The Croonian lectures on the chemical correlation of the functions of the body: Delivered before the Royal College of Physicians of London on June 20th, 22nd, 27th & 29th, 1905*. London : Publisher not identified. http://archive.org/details/b2497626x

Staudt, M., Lidzba, K., Grodd, W., Wildgruber, D., Erb, M., & Krägeloh-Mann, I. (2002). Right-Hemispheric Organization of Language Following Early Left-Sided Brain Lesions: Functional MRI Topography. *NeuroImage*, *16*(4), 954–967. https://doi.org/10.1006/nimg.2002.1108

Stephan, K. E., Marshall, J. C., Penny, W. D., Friston, K. J., & Fink, G. R. (2007). Interhemispheric integration of visual processing during task-driven lateralization. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(13), 3512–3522. https://doi.org/10.1523/JNEUROSCI.4766-06.2007

Stewart, M., & Black, K. (2015). Choosing a combined oral contraceptive pill. *Australian Prescriber*, *38*(1), 6–11. https://doi.org/10.18773/austprescr.2015.002

Striegel-Moore, R. H., & Bulik, C. M. (2007). Risk factors for eating disorders. *The American Psychologist*, *62*(3), 181–198. https://doi.org/10.1037/0003-066X.62.3.181

Sugiyama, N., Barros, R. P. A., Warner, M., & Gustafsson, J.-Å. (2010). ERβ: Recent understanding of estrogen signaling. *Trends in Endocrinology & Metabolism*, *21*(9), 545–552. https://doi.org/10.1016/j.tem.2010.05.001

Sundström Poromaa, I., & Gingnell, M. (2014a). Menstrual cycle influence on cognitive function and emotion processing-from a reproductive perspective. *Frontiers in Neuroscience*, *8*, 380–380. https://doi.org/10.3389/fnins.2014.00380

Suzuki, S., Gerhold, L. M., Böttner, M., Rau, S. W., Dela Cruz, C., Yang, E., Zhu, H., Yu, J., Cashion, A. B., Kindy, M. S., Merchenthaler, I., Gage, F. H., & Wise, P. M. (2007). Estradiol enhances neurogenesis following ischemic stroke through estrogen receptors alpha and beta. *The Journal of Comparative Neurology*, *500*(6), 1064–1075. https://doi.org/10.1002/cne.21240

Tasca, C., Rapetti, M., Carta, M. G., & Fadda, B. (2012). Women And Hysteria In The History Of Mental Health. *Clinical Practice and Epidemiology in Mental Health : CP & EMH*, *8*, 110–119. https://doi.org/10.2174/1745017901208010110

Tata, J. R. (2005). One hundred years of hormones. *EMBO Reports*, *6*(6), 490–496. https://doi.org/10.1038/sj.embor.7400444

Tjandra, T., Brooks, J. C. W., Figueiredo, P., Wise, R., Matthews, P. M., & Tracey, I. (2005). Quantitative assessment of the reproducibility of functional activation measured with BOLD and MR perfusion imaging: Implications for clinical trial design. *NeuroImage*, *27*(2), 393–401. https://doi.org/10.1016/j.neuroimage.2005.04.021

Toffoletto, S., Lanzenberger, R., Gingnell, M., Sundström-Poromaa, I., & Comasco, E. (2014). Emotional and cognitive functional imaging of estrogen and progesterone effects in the female human brain: A systematic review. *Psychoneuroendocrinology*, *50*, 28–52. https://doi.org/10.1016/j.psyneuen.2014.07.025

Trenor, C. C., Chung, R. J., Michelson, A. D., Neufeld, E. J., Gordon, C. M., Laufer, M. R., & Emans, S. J. (2011). Hormonal Contraception and Thrombotic Risk: A Multidisciplinary Approach. *Pediatrics*, *127*(2), 347–357. https://doi.org/10.1542/peds.2010-2221

Tsai, M.-J., & O'Malley, B. W. (1994). Molecular Mechanisms of Action of Steroid/Thyroid Receptor Superfamily Members. *Annual Review of Biochemistry*, *63*(1), 451–486. https://doi.org/10.1146/annurev.bi.63.070194.002315

Uddin, L. Q., Nomi, J. S., Hebert-Seropian, B., Ghaziri, J., & Boucher, O. (2017). Structure and function of the human insula. *Journal of Clinical Neurophysiology : Official Publication of the American Electroencephalographic Society*, *34*(4), 300–306. https://doi.org/10.1097/WNP.0000000000000377

Van Reen, E., & Kiesner, J. (2016). Individual differences in self-reported difficulty sleeping across the menstrual cycle. *Archives of Women's Mental Health*, *19*(4), 599–608. https://doi.org/10.1007/s00737-016-0621-9

Vargas, K. G., Milic, J., Zaciragic, A., Wen, K., Jaspers, L., Nano, J., Dhana, K., Bramer, W. M., Kraja, B., van Beeck, E., Ikram, M. A., Muka, T., & Franco, O. H. (2016). The functions of estrogen receptor beta in the female brain: A systematic review. *Maturitas*, *93*, 41–57. https://doi.org/10.1016/j.maturitas.2016.05.014

Vilares, I., Howard, J. D., Fernandes, H. L., Gottfried, J. A., & Kording, K. P. (2012). Differential Representations of Prior and Likelihood Uncertainty in the Human Brain. *Current Biology*, *22*(18), 1641–1648. https://doi.org/10.1016/j.cub.2012.07.010

Wang, Q., Würtz, P., Auro, K., Morin-Papunen, L., Kangas, A. J., Soininen, P., Tiainen, M., Tynkkynen, T., Joensuu, A., Havulinna, A. S., Aalto, K., Salmi, M., Blankenberg, S., Zeller, T., Viikari, J., Kähönen, M., Lehtimäki, T., Salomaa, V., Jalkanen, S., … Ala-Korpela, M. (2016). Effects of hormonal contraception on systemic metabolism: Cross-sectional and longitudinal evidence. *International Journal of Epidemiology*, *45*(5), 1445–1457. https://doi.org/10.1093/ije/dyw147

Weis, S., & Hausmann, M. (2010). Sex hormones: Modulators of interhemispheric inhibition in the human brain. *The Neuroscientist : A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *16*(2), 132–138. https://doi.org/10.1177/1073858409341481

Weis, S., Hausmann, M., Stoffers, B., Vohn, R., Kellermann, T., & Sturm, W. (2008). Estradiol modulates functional brain organization during the menstrual cycle: An analysis of interhemispheric inhibition. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(50), 13401–13410. https://doi.org/10.1523/JNEUROSCI.4392-08.2008

Weiss, C., Nettekoven, C., Rehme, A. K., Neuschmelting, V., Eisenbeis, A., Goldbrunner, R., & Grefkes, C. (2013). Mapping the hand, foot and face representations in the primary motor cortex—Retest reliability of neuronavigated TMS versus functional MRI. *NeuroImage*, *66*, 531–542. https://doi.org/10.1016/j.neuroimage.2012.10.046

Welz, A., Huffziger, S., Reinhard, I., Alpers, G. W., Ebner-Priemer, U., & Kuehner, C. (2016). Anxiety and rumination moderate menstrual cycle effects on mood in daily life. *Women & Health*, *56*(5), 540–560. https://doi.org/10.1080/03630242.2015.1101739

Wharton, W., Hirshman, E., Merritt, P., Doyle, L., Paris, S., & Gleason, C. (2008). Oral contraceptives and androgenicity: Influences on visuospatial task performance in younger individuals. *Experimental and Clinical Psychopharmacology*, *16*(2), 156–164. https://doi.org/10.1037/1064-1297.16.2.156

Whitehouse, A. J. O., Badcock, N., Groen, M. a, & Bishop, D. V. M. (2009). Reliability of a novel paradigm for determining hemispheric lateralization of visuospatial function. *Journal of the International Neuropsychological Society : JINS*, *15*(6), 1028–1032. https://doi.org/10.1017/S1355617709990555

Whitehouse, A. J. O., & Bishop, D. V. M. (2009). Hemispheric division of function is the result of independent probabilistic biases. *Neuropsychologia*, *47*(8–9), 1938–1943. https://doi.org/10.1016/j.neuropsychologia.2009.03.005

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org}

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wiegratz, I., & Thaler, C. J. (2011). Hormonal contraception—What kind, when, and for whom? *Deutsches Ärzteblatt International*, *108*(28–29), 495–505; quiz 506. https://doi.org/10.3238/arztebl.2011.0495

Wilke, M., & Schmithorst, V. J. (2006). A combined bootstrap/histogram analysis approach for computing a lateralization index from neuroimaging data. *NeuroImage*, *33*(2), 522–530. https://doi.org/10.1016/j.neuroimage.2006.07.010

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*(2), 1205–1212. https://doi.org/10.1016/j.neuroimage.2009.09.046

Wood, W., & Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin*, *128*(5), 699–727.

Woolley, C. S. (1998). Estrogen-mediated structural and functional synaptic plasticity in the female rat hippocampus. *Hormones and Behavior*, *34*(2), 140–148. https://doi.org/10.1006/hbeh.1998.1466

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited—Again. *NeuroImage*, *2*(3), 173–181. https://doi.org/10.1006/nimg.1995.1023

Young, W. C., Goy, R. W., & Phoenix, C. H. (1964). HORMONES AND SEXUAL BEHAVIOR. *Science (New York, N.Y.)*, *143*(3603), 212–218.

Ysrraelit, M. C., & Correale, J. (2019). Impact of sex hormones on immune function and multiple sclerosis development. *Immunology*, *156*(1), 9–22. https://doi.org/10.1111/imm.13004

Zethraeus, N., Dreber, A., Ranehill, E., Blomberg, L., Labrie, F., von Schoultz, B., Johannesson, M., & Hirschberg, A. L. (2017). A first-choice combined oral contraceptive influences general well-being in healthy women: A double-blind, randomized, placebo-controlled trial. *Fertility and Sterility*, *107*(5), 1238–1245. https://doi.org/10.1016/j.fertnstert.2017.02.120

Appendix

S1: Supplementary Project 1

***Table 5**. Statistics of each Paradigm's Group Analysis.*

From top to bottom: "Dots-in space" task, Mental rotation task, Landmark task, Landmark task version B session 1, Landmark task version B session 2, Landmark task version C session 1, Landmark task version C session 2.

"Dots-in-space"task

Statistic: p-values adjusted for search volume

set-level		**cluster-level**				**peak-level**					
p	*c*	*p(FWE-corr)*	*p(FDR-corr)*	*kE*	*p(unc)*	*p(FWE-corr)*	*p(FDR-corr)*	*T*	*equivZ*	*p(unc)*	*x,y,z {mm}*
0,00115903	10	0	9,26E-101	23147	2,20E-102	1,25E-05	0,0014703	16,6488495	6,43058647	6,36E-11	[14;-90;14]
						4,24E-05	0,0014703	15,1928988	6,24282712	2,15E-10	[34;-84;-6]
						0,000141679	0,00221121	13,8630781	6,05125001	7,19E-10	[18;-96;6]
		0	3,81E-29	3663	1,81E-30	0,000155838	0,00221121	13,7626591	6,03588905	7,90E-10	[34;22;-8]
						0,006224343	0,0064947	10,3200779	5,40966459	3,16E-08	[12;14;2]
						0,037577826	0,01194703	8,91240692	5,07810972	1,91E-07	[14;10;10]
		0	4,32E-22	2412	3,08E-23	0,000759748	0,00343272	12,1821022	5,77473841	3,85E-09	[-6;32;28]
						0,002813646	0,00557716	10,9937248	5,55014633	1,43E-08	[6;20;54]
						0,011383457	0,00800356	9,8299284	5,30051011	5,77E-08	[-4;32;38]
		0	1,43E-21	2309	1,36E-22	0,002053164	0,00553207	11,2706423	5,60497637	1,04E-08	[-34;16;-10]
						0,004062381	0,0055905	10,6778908	5,48558277	2,06E-08	[-32;22;2]
						0,013812563	0,00883004	9,67673969	5,26508874	7,01E-08	[-42;18;-10]
		1,45E-06	3,45E-07	440	5,75E-08	0,063455955	0,01527139	8,53028488	4,97761379	3,22E-07	[-28;60;4]
						0,068795054	0,01582856	8,47241306	4,96194937	3,49E-07	[-46;50;0]
						0,324961758	0,04682087	7,14905739	4,56722153	2,47E-06	[-28;54;12]
		3,38E-11	1,12E-11	924	1,34E-12	0,096083048	0,01872936	8,23597813	4,89667568	4,87E-07	[-50;22;36]
						0,257975494	0,03779036	7,38954258	4,64463334	1,70E-06	[-42;-2;46]
						0,40599478	0,05765909	6,90694141	4,48640581	3,62E-06	[-34;4;58]
		6,86E-07	1,90E-07	470	2,72E-08	0,251449258	0,03718708	7,41578436	4,652915	1,64E-06	[24;52;0]
						0,562084126	0,07637271	6,51870728	4,35034433	6,80E-06	[26;48;-10]
						0,638033403	0,0888149	6,34670305	4,28734086	9,04E-06	[36;54;6]

		0,05155943	0,011005957	99	0,002096	0,65573243	0,09085623	6,30703497	4,27256159	9,66E-06	[-50;-32;-8]
						0,998893371	0,34923182	4,82889795	3,64477684	0,00013381	[-58;-30;-6]
		0,95122379	0,483524098	20	0,119618	0,999991652	0,53374951	4,40436077	3,43176359	0,00029984	[34;56;24]
		0,82980971	0,327266695	28	0,070129	0,999999318	0,62573469	4,25100422	3,35058539	0,0004032	[-2;-28;24]

Mental rotation task

Statistic: p-values adjusted for search volume

set-level		**cluster-level**				**peak-level**					
p	*c*	*p(FWE-corr)*	*p(FDR-corr)*	*kE*	*p(unc)*	*p(FWE-corr)*	*p(FDR-corr)*	*T*	*equivZ*	*p(unc)*	*x,y,z {mm}*
5,17E-14	14	4,60E-06	1,31E-06	501	2,44E-07	0,000332253	0,00770612	12,9427805	5,90520018	1,76E-09	[-48;28;24]
		0	8,81E-18	2323	4,10E-19	0,000682252	0,00770612	12,2424192	5,78543492	3,62E-09	[32;-52;64]
						0,024220244	0,02529955	9,20868969	5,15270429	1,28E-07	[38;-44;64]
						0,125309765	0,06300866	7,83571005	4,78124276	8,71E-07	[16;-62;60]
		0	3,98E-28	4609	9,26E-30	0,001795534	0,00790363	11,3508329	5,62055678	9,52E-09	[-38;-54;54]
						0,002540649	0,00790363	11,0443258	5,56028662	1,35E-08	[-30;-58;46]
						0,009112491	0,01552123	9,97294617	5,33300487	4,83E-08	[-20;-62;58]
		1,84E-08	1,40E-08	797	9,74E-10	0,002702686	0,00790363	10,9904432	5,5494868	1,43E-08	[-24;0;52]
						0,166101267	0,07566901	7,56179953	4,69842047	1,31E-06	[-22;6;58]
						0,999812064	0,61837233	4,41655922	3,43812118	0,00029288	[-20;6;72]
		4,42E-06	1,31E-06	503	2,34E-07	0,010026028	0,01552123	9,89620495	5,31563665	5,31E-08	[30;-2;58]
						0,873754693	0,25125013	5,53135729	3,96314912	3,70E-05	[30;10;60]
						0,892910488	0,25177468	5,47079897	3,93723272	4,12E-05	[24;12;50]
		0,01523698	0,003498131	160	0,000814	0,011093691	0,01552123	9,81541729	5,29718234	5,88E-08	[-30;-44;-34]
						0,956742038	0,3194611	5,20261097	3,8190987	6,70E-05	[-26;-42;-42]
						0,959487681	0,32063393	5,18639851	3,81177633	6,90E-05	[-38;-56;-34]
		0,00013774	3,49E-05	343	7,30E-06	0,093401223	0,05211752	8,12087917	4,86413421	5,75E-07	[-48;6;28]
						0,603875946	0,17229447	6,17724609	4,22353402	1,20E-05	[-40;2;32]
						0,995748403	0,43930327	4,77001524	3,61621807	0,00014947	[-50;6;18]
		1,65E-07	7,51E-08	674	8,74E-09	0,289556268	0,11396193	7,00976181	4,52108996	3,08E-06	[56;-60;-8]
						0,325713486	0,12229152	6,88835382	4,48007716	3,73E-06	[44;-60;-10]
						0,785204169	0,23217487	5,76757717	4,06169357	2,44E-05	[46;-56;2]
		7,38E-07	2,80E-07	594	3,91E-08	0,366345519	0,12342926	6,76428604	4,43736646	4,55E-06	[8;-80;-30]
						0,538990876	0,15487667	6,32375622	4,27880306	9,40E-06	[-6;-78;-38]
						0,842499374	0,24723789	5,62106848	4,00104575	3,15E-05	[26;-72;-50]
		9,54E-08	5,44E-08	704	5,06E-09	0,389628665	0,12342926	6,6978693	4,41416165	5,07E-06	[14;34;28]

Mental rotation task

Statistic: p-values adjusted for search volume

set-level		cluster-level				peak-level					
p	*c*	*p(FWE-corr)*	*p(FDR-corr)*	*kE*	*p(unc)*	*p(FWE-corr)*	*p(FDR-corr)*	*T*	*equivZ*	*p(unc)*	*x,y,z {mm}*
						0,577126387	0,1679479	6,23703671	4,24624884	1,09E-05	[0;24;40]
						0,599157374	0,17229447	6,18774748	4,22753967	1,18E-05	[2;14;56]
		0,06070773	0,01189046	115	0,003318	0,666635385	0,19362352	6,03840876	4,16991948	1,52E-05	[34;-58;-26]
						0,848182658	0,24723789	5,60542154	3,99447818	3,24E-05	[30;-66;-28]
		0,4512016	0,097644529	54	0,031791	0,727967618	0,21620437	5,90133381	4,11575901	1,93E-05	[-14;-54;-46]
		0,16265296	0,031110217	85	0,009405	0,844992091	0,24723789	5,61423731	3,99818064	3,19E-05	[46;34;18]
		0,02610938	0,00547953	142	0,001402	0,892940209	0,25177468	5,47070074	3,93719046	4,12E-05	[36;22;-6]

Landmark task

Statistic: p-values adjusted for search volume

set-level		**cluster-level**				**peak-level**					
p	*c*	*p(FWE-corr)*	*p(FDR-corr)*	*kE*	*p(unc)*	*p(FWE-corr)*	*p(FDR-corr)*	*T*	*equivZ*	*p(unc)*	*x,y,z {mm}*
4,69E-14	24	0,00478688	0,003249541	180	0,000223	0,036316335	0,30777445	8,91107368	5,07776767	1,91E-07	[-24;-2;60]
						0,999941319	0,78450685	4,42208624	3,44099699	0,00028979	[-26;-4;44]
		0	8,21E-18	2127	1,12E-19	0,0605167	0,30777445	8,5385313	4,97983611	3,18E-07	[48;-54;-6]
						0,143698942	0,35103553	7,82676411	4,77858849	8,83E-07	[46;-66;-12]
						0,185881284	0,35103553	7,57351446	4,70202974	1,29E-06	[34;-82;34]
		2,39E-08	2,70E-08	691	1,11E-09	0,278069696	0,41399601	7,16926193	4,57383269	2,39E-06	[-38;-84;16]
						0,423712896	0,41487814	6,72028923	4,42202169	4,89E-06	[-40;-92;-4]
						0,769208965	0,44084643	5,91516352	4,12127985	1,88E-05	[-30;-92;-4]
		0,24249921	0,078368287	67	0,012882	0,33597248	0,41487814	6,97258472	4,50861199	3,26E-06	[22;4;10]
						0,995653901	0,58020298	4,87855244	3,66862247	0,00012193	[28;-2;4]
		0,02617747	0,011227493	128	0,00123	0,380582578	0,41487814	6,83892536	4,46315946	4,04E-06	[-32;-32;56]
		2,44E-13	4,13E-13	1343	1,13E-14	0,422343285	0,41487814	6,72392464	4,42329361	4,86E-06	[28;-4;54]
						0,441959097	0,41487814	6,67257881	4,40526204	5,28E-06	[26;0;72]
						0,688324238	0,44084643	6,10074043	4,19414204	1,37E-05	[56;-20;52]
		6,92E-05	5,86E-05	332	3,21E-06	0,520812746	0,44084643	6,47886896	4,33590601	7,26E-06	[10;2;34]
						0,592038542	0,44084643	6,31543016	4,27569734	9,53E-06	[14;-10;40]
						0,62570652	0,44084643	6,24014616	4,24742408	1,08E-05	[8;-2;44]
		0,10607862	0,037970682	89	0,005201	0,576433378	0,44084643	6,35063934	4,28880225	8,98E-06	[-28;-14;72]
		0,94466129	0,426099297	21	0,13425	0,666126982	0,44084643	6,15025902	4,21320833	1,26E-05	[-26;-4;4]
		0,51760428	0,176315932	46	0,033814	0,756088245	0,44084643	5,94615746	4,13360626	1,79E-05	[36;34;38]
						0,99990655	0,76825958	4,46052122	3,46091355	0,00026917	[30;32;30]
		0,33894383	0,107811991	58	0,019199	0,782734493	0,44084643	5,88266802	4,10828722	1,99E-05	[-24;10;6]
		0,10607862	0,037970682	89	0,005201	0,812606688	0,44084643	5,80839872	4,07832379	2,27E-05	[12;-28;74]
						0,956291834	0,46905986	5,31166649	3,86781041	5,49E-05	[20;-12;78]
		0,01333626	0,007178166	148	0,000623	0,892037366	0,4493786	5,58125162	3,98429835	3,38E-05	[-16;-66;54]
						0,992862584	0,56484092	4,9533534	3,70414038	0,00010605	[-8;-68;58]

Landmark task

Statistic: p-values adjusted for search volume

set-level		cluster-level				peak-level					
p	*c*	*p(FWE-corr)*	*p(FDR-corr)*	*kE*	*p(unc)*	*p(FWE-corr)*	*p(FDR-corr)*	*T*	*equivZ*	*p(unc)*	*x,y,z {mm}*
						0,999521724	0,70162575	4,61222219	3,53814761	0,00020147	[-14;-58;50]
		0,01472975	0,007178166	145	0,000688	0,901699	0,4493786	5,54849863	3,97043559	3,59E-05	[-42;-8;58]
						0,962066331	0,48146737	5,27681684	3,85234623	5,85E-05	[-48;-18;50]
						0,999921435	0,76825958	4,44597387	3,45339208	0,00027679	[-38;-14;52]
		0,74144241	0,254452323	34	0,062742	0,935974821	0,46122174	5,41423845	3,9127792	4,56E-05	[-2;-78;-36]
		0,84939369	0,320507417	28	0,08781	0,953278272	0,46905986	5,32862616	3,87530178	5,32E-05	[28;52;-2]
		0,74144241	0,254452323	34	0,062742	0,977269421	0,50625571	5,16255713	3,80096989	7,21E-05	[64;-18;36]
						0,999999996	0,94392125	3,88034463	3,14440269	0,00083213	[56;-16;38]
		0,62668396	0,208525615	40	0,045704	0,98005755	0,50625571	5,13588476	3,78882571	7,57E-05	[-18;-102;2]
		0,13274735	0,043842071	83	0,006606	0,986679274	0,54201345	5,05899286	3,75349013	8,72E-05	[-18;10;-14]
		0,84939369	0,320507417	28	0,08781	0,996487742	0,58193619	4,84851408	3,6542229	0,00012898	[-56;0;40]
		0,94466129	0,426099297	21	0,13425	0,996514956	0,58193619	4,84743881	3,65370598	0,00012924	[10;-38;74]
		0,86556455	0,323559847	27	0,093079	0,998962704	0,6515966	4,69574547	3,57975428	0,00017196	[60;16;32]
						0,999994035	0,85080797	4,25595522	3,35324285	0,00039935	[62;12;24]
		0,95433344	0,435448834	20	0,143161	0,999646703	0,71406545	4,58180666	3,52283615	0,00021348	[44;24;28]
		0,58924978	0,200855725	42	0,041272	0,99972022	0,72736884	4,55914879	3,51137379	0,0002229	[16;12;-8]
						0,999999579	0,85597981	4,09781027	3,2671203	0,00054324	[14;12;-18]

Landmark task version B session 1

Statistic: p-values adjusted for search volume

set-level		**cluster-level**				**peak-level**					
p	*c*	*p(FWE-corr)*	*p(FDR-corr)*	*kE*	*p(unc)*	*p(FWE-corr)*	*p(FDR-corr)*	*T*	*equivZ*	*p(unc)*	*x,y,z {mm}*
0,00177442	10	0,000586381	0,00080024	274	2,96E-05	0,219853333	0,407434768	6,28363895	4,56720054	2,47E-06	[30;22;-2]
						0,927892422	0,655405154	4,8518219	3,86568194	5,54E-05	[30;30;2]
						0,998174995	0,655405154	4,31640005	3,55906665	0,00018609	[44;20;-12]
		0,143165808	0,04216007	87	0,0078074	0,373724546	0,407434768	5,89569139	4,39236755	5,61E-06	[6;22;42]
		0,266218516	0,06033046	69	0,0156412	0,402305104	0,407434768	5,8378315	4,36538914	6,34E-06	[-20;-96;-4]
		0,070220787	0,02483317	108	0,003679	0,585631734	0,523073098	5,51222658	4,20894756	1,28E-05	[-34;18;-2]
		0,001307222	0,00089231	243	6,61E-05	0,676487833	0,536025046	5,36273527	4,13439163	1,78E-05	[50;38;26]
						0,972075065	0,655405154	4,66465235	3,76154027	8,44E-05	[42;40;36]
						0,992269034	0,655405154	4,47697496	3,65386452	0,00012916	[40;40;20]
		0,454071664	0,09175235	53	0,0305841	0,899341958	0,655405154	4,93395329	3,91038733	4,61E-05	[44;2;32]
		0,044290659	0,0206017	122	0,0022891	0,923453576	0,655405154	4,86574554	3,87330286	5,37E-05	[18;-96;-4]
						0,931435512	0,655405154	4,84031868	3,85937277	5,68E-05	[20;-90;2]
						0,984576423	0,655405154	4,57092381	3,70817979	0,00010438	[12;-94;6]
		0,315696669	0,06469444	64	0,0191687	0,974149801	0,655405154	4,65163898	3,75418052	8,70E-05	[38;54;8]
		0,202119024	0,0513425	77	0,0114094	0,993354904	0,655405154	4,45812225	3,64286377	0,00013481	[-26;-70;-48]
						0,996646281	0,655405154	4,37926531	3,59647863	0,00016128	[-36;-60;-48]
		0,920453892	0,34536334	24	0,1279123	0,99991363	0,741642225	4,06899405	3,40801337	0,00032719	[48;10;24]

Landmark task version B session 2

Statistic: p-values adjusted for search volume

set-level		cluster-level				peak-level					
p	*c*	*p(FWE-corr)*	*p(FDR-corr)*	*kE*	*p(unc)*	*p(FWE-corr)*	*p(FDR-corr)*	*T*	*equivZ*	*p(unc)*	*x,y,z {mm}*
0,00611432	9	7,97E-09	1,00E-08	849	4,18E-10	0,015226747	0,049261398	8,03072739	5,24262065	7,92E-08	[38;20;-10]
						0,049579408	0,05329877	7,25692368	4,96390491	3,45E-07	[26;12;-8]
						0,168771598	0,145312458	6,44272375	4,63597728	1,78E-06	[20;6;0]
		4,18E-05	1,87E-05	398	2,19E-06	0,033202811	0,053088242	7,52224827	5,0628124	2,07E-07	[-20;14;0]
						0,273090658	0,14696748	6,10684443	4,48880789	3,58E-06	[-26;6;0]
						0,572335496	0,190780659	5,51177359	4,20872429	1,28E-05	[-28;18;10]
		4,46E-05	1,87E-05	395	2,34E-06	0,350238658	0,165286682	5,92283154	4,40493932	5,29E-06	[42;44;30]
						0,783487825	0,275114794	5,15665007	4,02865506	2,80E-05	[44;44;22]
						0,991323943	0,478206456	4,46903086	3,64923321	0,00013151	[48;26;30]
		0,006637953	0,00209422	190	0,000349	0,459613989	0,165286682	5,70665836	4,3033219	8,41E-06	[12;-74;52]
						0,992435195	0,478206456	4,45152187	3,63900432	0,00013685	[12;-66;60]
		0,025968037	0,00661869	143	0,0013789	0,800001735	0,275114794	5,12614965	4,01270724	3,00E-05	[6;14;52]
						0,999855618	0,645877835	4,08149099	3,41579181	0,00031798	[6;24;44]
						0,999989887	0,720461502	3,91360903	3,30995048	0,00046656	[10;22;34]
		0,282370894	0,05961881	69	0,0173888	0,880235081	0,299901923	4,95912361	3,92396903	4,36E-05	[28;56;-4]
						0,888489842	0,299901923	4,93911028	3,91317452	4,55E-05	[20;46;-10]
		0,646270913	0,14523405	42	0,0544628	0,923041139	0,335993293	4,84455585	3,8616981	5,63E-05	[26;-54;42]
						0,999999391	0,818212902	3,77135396	3,21795418	0,00064554	[28;-56;52]
		0,282370894	0,05961881	69	0,0173888	0,989221537	0,478206456	4,4977746	3,66596186	0,00012321	[-4;-74;-20]
		0,531596758	0,1192408	49	0,0397469	0,998053491	0,503247165	4,2998867	3,54917497	0,00019322	[50;6;30]
						0,999987748	0,720461502	3,92444682	3,3168717	0,00045516	[54;14;34]

Landmark task version C session 1

Statistic: p-values adjusted for search volume

set-level	cluster-level					peak-level					
p	*c*	*p(FWE-corr)*	*p(FDR-corr)*	*kE*	*p(unc)*	*p(FWE-corr)*	*p(FDR-corr)*	*T*	*equivZ*	*p(unc)*	*x,y,z {mm}*
0,10998362	6	7,63E-08	1,00E-07	723	4,02E-09	0,015847118	0,033055632	8,0091877	5,2352499	8,24E-08	[36;28;-6]
						0,78290575	0,522921036	5,15460968	4,02759064	2,82E-05	[44;20;6]
						0,994328084	0,680518595	4,41309071	3,61644905	0,00014934	[52;18;2]
		0,001984966	0,0008721	236	0,0001047	0,083354933	0,089390524	6,91080666	4,82920297	6,85E-07	[8;24;38]
						0,991698122	0,680518595	4,46021509	3,64408665	0,00013417	[4;24;48]
						0,999905531	0,865277289	4,04857779	3,39527128	0,0003428	[-6;14;48]
		7,55E-05	4,97E-05	373	3,98E-06	0,344024606	0,288695591	5,93327332	4,40976213	5,17E-06	[-32;22;10]
						0,666088308	0,450626208	5,35433054	4,13014714	1,81E-05	[-38;14;14]
						0,8463946	0,54973447	5,03138924	3,96265665	3,71E-05	[-42;16;2]
		0,211443428	0,07819885	78	0,0125118	0,929694125	0,627408665	4,82045031	3,84844769	5,94E-05	[38;42;22]
						0,999374503	0,865277289	4,19245386	3,48416069	0,00024684	[42;48;26]
						0,999980699	0,929096321	3,94757843	3,33160285	0,00043174	[32;32;28]
		0,888068198	0,57670097	27	0,1153402	0,950817748	0,627408665	4,74380636	3,80597029	7,06E-05	[18;4;74]
		0,952912015	0,67060691	21	0,1609457	0,999505703	0,865277289	4,17275333	3,47211339	0,00025819	[14;-64;58]

Landmark task version C session 2

Statistic: p-values adjusted for search volume

set-level		**cluster-level**				**peak-level**					
p	*c*	*p(FWE-corr)*	*p(FDR-corr)*	*kE*	*p(unc)*	*p(FWE-corr)*	*p(FDR-corr)*	*T*	*equivZ*	*p(unc)*	*x,y,z {mm}*
0,04706651	7	0,000123566	0,00014317	351	6,51E-06	0,286821247	0,329794682	6,06822205	4,4714002	3,89E-06	[32;26;0]
						0,769502562	0,329794682	5,17882156	4,04019894	2,67E-05	[36;24;8]
		0,009782151	0,00301468	177	0,0005177	0,623109819	0,329794682	5,42498255	4,16565121	1,55E-05	[46;6;32]
						0,761724104	0,329794682	5,19262028	4,04736269	2,59E-05	[54;10;24]
						0,999888684	0,769267592	4,05999422	3,40240173	0,00033398	[54;8;16]
		0,001835762	0,00106444	239	9,68E-05	0,665955898	0,329794682	5,35462379	4,13029533	1,81E-05	[42;42;20]
						0,774856751	0,329794682	5,16923714	4,03521373	2,73E-05	[42;42;30]
						0,999997367	0,858872727	3,83874273	3,2618018	0,00055353	[44;52;14]
		0,010354012	0,00301468	175	0,0005481	0,679855262	0,329794682	5,33168364	4,11868187	1,91E-05	[12;20;34]
						0,99175722	0,695213242	4,45937395	3,64359521	0,00013443	[6;16;50]
						0,998080477	0,695213242	4,29541826	3,5464937	0,0001952	[8;28;28]
		0,241357162	0,0457191	74	0,014547	0,7575476	0,329794682	5,19997215	4,05117302	2,55E-05	[-30;20;8]
		0,033730845	0,00795098	135	0,001807	0,872521886	0,414185834	4,97414398	3,93204753	4,21E-05	[10;-74;52]
						0,99843931	0,695213242	4,27517223	3,5343204	0,00020441	[18;-70;42]
		0,071615041	0,01434905	111	0,0039134	0,953848884	0,566946798	4,73117971	3,79892128	7,27E-05	[36;-80;28]
						0,997954344	0,695213242	4,30177307	3,55030629	0,00019239	[32;-76;34]
						0,998905925	0,695213242	4,24181986	3,51417797	0,00022056	[28;-72;28]

S2: Supplementary Project 2

Table 6. *Group and session effects on hormone concentration.*

Significant effects of the intercept, as well as factor group (NC or OC), hormonal phase (menstruation and pill-withdrawal or midluteal phase and pill-intake defined by session_id), and the interactions hormones by hormonal phase and group by hormonal phase. No significant effects of the interaction hormones by group and the threefold interaction were obtained.

Predictors	*F*	*Df*	*Df.res*	*Pr(>F)*	*Semi-partial R^2*
(Intercept)	515.8595	1	34.0000	0.0000	
hormone	16.4289	1	102.0000	0.0001	0.1387
session_id	20.1770	1	102.0000	0.0000	0.1651
group	46.3030	1	34.0000	0.0000	0.5766
hormone: session_id	36.7330	1	102.0000	0.0000	0.2648
hormone: group	2.2502	1	102.0000	0.1367	0.0216
session_id:group	69.4683	1	102.0000	0.0000	0.4051
hormone:group:session_id	0.1362	1	102.0000	0.7128	0.0013

Table 7 *Session effects on CAT12 GM-values for group NC and OC.*

Right hemisphere = upper figure; left hemisphere lower figure.

** = p < 0.05 fdr corrected; ** = p < 0.01 fdr corrected; NC = natural cycle; OC = oral contraceptives; M = menstruation and pill-withdrawal; P = post-ovulation (midluteal phase) and pill-intake*

Session effects on CAT 12 GM-values for group NC and OC.

roi	*hemisphere*	*group*	*contrast*	*estimate*	*SE*	*lower.CL*	*upper.CL*	*t.ratio*	*df*	*p.value*	*p.fdr*	*sig*
Anterior Cingulate Gyrus	Right	NC	M - P	0.0395	0.0326	-0.0292	0.1083	1.2146	16.9130	0.2412	0.4230	
Anterior Cingulate Gyrus	Right	OC	M - P	-0.0478	0.0222	-0.0947	-0.0010	-2.1537	16.9171	0.0460	0.0575	
Middle Cingulate Gyrus	Right	NC	M - P	0.0311	0.0312	-0.0348	0.0970	0.9960	16.9375	0.3333	0.4230	
Middle Cingulate Gyrus	Right	OC	M - P	-0.0261	0.0083	-0.0436	-0.0087	-3.1638	16.5385	0.0058	0.0146	*
Middle Frontal Gyrus	Right	NC	M - P	0.2185	0.1608	-0.1209	0.5579	1.3583	16.9718	0.1921	0.4230	
Middle Frontal Gyrus	Right	OC	M - P	-0.1560	0.0932	-0.3528	0.0407	-1.6734	16.9483	0.1126	0.1126	
Posterior Cingulate Gyrus	Right	NC	M - P	0.0129	0.0131	-0.0148	0.0406	0.9853	16.8616	0.3384	0.4230	
Posterior Cingulate Gyrus	Right	OC	M - P	-0.0196	0.0089	-0.0384	-0.0008	-2.1971	16.7987	0.0423	0.0575	
Superior Temporal Gyrus	Right	NC	M - P	-0.0005	0.0238	-0.0508	0.0498	-0.0208	16.7221	0.9837	0.9837	
Superior Temporal Gyrus	Right	OC	M - P	-0.0613	0.0153	-0.0936	-0.0290	-4.0018	16.8463	0.0009	0.0047	**

Session effects on CAT 12 GM-values for group NC and OC.

roi	*hemisphere*	*group*	*contrast*	*estimate*	*SE*	*lower.CL*	*upper.CL*	*t.ratio*	*df*	*p.value*	*p.fdr*	*sig*
Anterior Cingulate Gyrus	Left	NC	M - P	0.0323	0.0450	-0.0626	0.1272	0.7179	16.9665	0.4826	0.6756	
Anterior Cingulate Gyrus	Left	OC	M - P	-0.0438	0.0196	-0.0852	-0.0025	-2.2366	16.8651	0.0391	0.1369	
Anterior Insula	Left	NC	M - P	0.0383	0.0299	-0.0249	0.1014	1.2796	16.9421	0.2179	0.6756	
Anterior Insula	Left	OC	M - P	-0.0025	0.0190	-0.0425	0.0375	-0.1317	16.8946	0.8968	0.8968	
Middle Cingulate Gyrus	Left	NC	M - P	0.0244	0.0270	-0.0327	0.0814	0.9021	16.9218	0.3797	0.6756	
Middle Cingulate Gyrus	Left	OC	M - P	-0.0179	0.0113	-0.0419	0.0060	-1.5791	16.7554	0.1330	0.2327	
Middle Frontal Gyrus	Left	NC	M - P	0.1741	0.1208	-0.0807	0.4289	1.4418	16.9549	0.1676	0.6756	
Middle Frontal Gyrus	Left	OC	M - P	-0.0967	0.0505	-0.2032	0.0099	-1.9165	16.7993	0.0725	0.1691	
Middle Temporal Gyrus	Left	NC	M - P	-0.0014	0.0339	-0.0730	0.0701	-0.0426	16.8344	0.9665	0.9665	
Middle Temporal Gyrus	Left	OC	M - P	-0.0517	0.0516	-0.1607	0.0573	-1.0014	16.8331	0.3308	0.4239	
Posterior Cingulate Gyrus	Left	NC	M - P	0.0095	0.0195	-0.0317	0.0508	0.4877	16.8980	0.6320	0.7374	
Posterior Cingulate Gyrus	Left	OC	M - P	-0.0344	0.0106	-0.0568	-0.0120	-3.2391	16.8672	0.0049	0.0340	*
Posterior Insula	Left	NC	M - P	-0.0096	0.0125	-0.0361	0.0168	-0.7698	16.9441	0.4520	0.6756	
Posterior Insula	Left	OC	M - P	0.0109	0.0117	-0.0137	0.0355	0.9342	16.9341	0.3633	0.4239	

Table 8. *Session effects on Freesurfer GM-values for group NC and OC.*

Right hemisphere = upper figure; left hemisphere lower figure.

NC = natural cycle; OC = oral contraceptives; M = menstruation and pill-withdrawal;

P = post-ovulation (midluteal phase) and pill-intake

Session effects on Freesurfer GM-values for group NC and OC.

roi	*hemisphere*	*group*	*contrast*	*estimate*	*SE*	*lower.CL*	*upper.CL*	*t.ratio*	*df*	*p.value*	*p.fdr*	*sig*
Caudal Anterior Cingulate	Right	NC	M - P	-0.0136	0.0226	-0.0614	0.0342	-0.6011	16.9237	0.5557	0.7585	
Caudal Anterior Cingulate	Right	OC	M - P	-0.0170	0.0260	-0.0718	0.0379	-0.6529	16.9479	0.5226	0.6683	
Caudal Middle Frontal	Right	NC	M - P	-0.0498	0.0804	-0.2194	0.1199	-0.6190	16.9597	0.5442	0.7585	
Caudal Middle Frontal	Right	OC	M - P	-0.0942	0.0584	-0.2174	0.0290	-1.6139	16.9448	0.1250	0.6683	
Isthmus Cingulate	Right	NC	M - P	-0.0020	0.0227	-0.0499	0.0459	-0.0886	16.9461	0.9305	0.9305	
Isthmus Cingulate	Right	OC	M - P	0.0208	0.0249	-0.0318	0.0734	0.8332	16.9731	0.4163	0.6683	
Posterior Cingulate	Right	NC	M - P	0.0732	0.0285	0.0131	0.1334	2.5689	16.9555	0.0199	0.1396	
Posterior Cingulate	Right	OC	M - P	0.0173	0.0301	-0.0462	0.0808	0.5750	16.9587	0.5728	0.6683	
Rostral Anterior Cingulate	Right	NC	M - P	0.0516	0.0318	-0.0156	0.1187	1.6208	16.9816	0.1235	0.4322	
Rostral Anterior Cingulate	Right	OC	M - P	-0.0095	0.0345	-0.0824	0.0633	-0.2762	16.9799	0.7857	0.7857	
Rostral Middle Frontal	Right	NC	M - P	-0.0970	0.1288	-0.3687	0.1747	-0.7534	16.9709	0.4615	0.7585	
Rostral Middle Frontal	Right	OC	M - P	0.0842	0.1106	-0.1492	0.3176	0.7614	16.9830	0.4568	0.6683	
Superior Temporal	Right	NC	M - P	0.0274	0.0594	-0.0979	0.1527	0.4617	16.9356	0.6502	0.7585	
Superior Temporal	Right	OC	M - P	-0.0462	0.0624	-0.1779	0.0856	-0.7395	16.9429	0.4697	0.6683	

Session effects on Freesurfer GM-values for group NC and OC.

roi	*hemisphere*	*group*	*contrast*	*estimate*	*SE*	*lower.CL*	*upper.CL*	*t.ratio*	*df*	*p.value*	*p.fdr*	*sig*
Caudal Anterior Cingulate	Left	NC	M - P	-0.0030	0.0247	-0.0551	0.0492	-0.1213	16.9305	0.9049	0.9049	
Caudal Anterior Cingulate	Left	OC	M - P	-0.0031	0.0246	-0.0550	0.0487	-0.1277	16.9582	0.8999	0.9119	
Caudal Middle Frontal	Left	NC	M - P	0.0224	0.0571	-0.0982	0.1429	0.3916	16.9629	0.7002	0.9049	
Caudal Middle Frontal	Left	OC	M - P	-0.1109	0.0495	-0.2154	-0.0065	-2.2421	16.9022	0.0387	0.3093	
Insula	Left	NC	M - P	0.0341	0.1353	-0.2514	0.3196	0.2520	16.9983	0.8040	0.9049	
Insula	Left	OC	M - P	-0.0545	0.0488	-0.1576	0.0485	-1.1173	16.9791	0.2794	0.6339	
Isthmus Cingulate	Left	NC	M - P	0.0109	0.0242	-0.0402	0.0619	0.4488	16.9444	0.6593	0.9049	
Isthmus Cingulate	Left	OC	M - P	0.0036	0.0322	-0.0643	0.0715	0.1123	16.9868	0.9119	0.9119	
Middle Temporal	Left	NC	M - P	-0.1514	0.0951	-0.3521	0.0492	-1.5923	16.9883	0.1298	0.5191	
Middle Temporal	Left	OC	M - P	0.0220	0.0634	-0.1118	0.1559	0.3475	16.9603	0.7325	0.9119	
Posterior Cingulate	Left	NC	M - P	0.0064	0.0315	-0.0601	0.0730	0.2041	16.9665	0.8407	0.9049	
Posterior Cingulate	Left	OC	M - P	-0.0579	0.0438	-0.1504	0.0346	-1.3203	16.9859	0.2042	0.6339	
Rostral Anterior Cingulate	Left	NC	M - P	-0.0187	0.0360	-0.0947	0.0572	-0.5202	16.9699	0.6097	0.9049	
Rostral Anterior Cingulate	Left	OC	M - P	0.0323	0.0464	-0.0656	0.1302	0.6963	16.9890	0.4957	0.7931	
Rostral Middle Frontal	Left	NC	M - P	0.1665	0.1029	-0.0506	0.3836	1.6180	16.9614	0.1241	0.5191	
Rostral Middle Frontal	Left	OC	M - P	0.0938	0.0910	-0.0982	0.2859	1.0312	16.9566	0.3170	0.6339	

Table 9. *Group effects on CAT12 GM-values for group NC and OC.*

Right hemisphere = upper figure; left hemisphere lower figure.

NC = natural cycle; OC = oral contraceptives

Group effects on CAT 12 GM-values.

roi	*hemisphere*	*contrast*	*estimate*	*SE*	*lower.CL*	*upper.CL*	*t.ratio*	*df*	*p.value*	*p.fdr*	*sig*
Anterior Cingulate Gyrus	Right	NC - OC	0.0695	0.1560	-0.2483	0.3873	0.4454	32.1086	0.6590	0.6590	
Middle Cingulate Gyrus	Right	NC - OC	0.1342	0.1324	-0.1355	0.4039	1.0136	32.1196	0.3183	0.4325	
Middle Frontal Gyrus	Right	NC - OC	0.5574	0.4488	-0.3568	1.4715	1.2419	32.0425	0.2233	0.4325	
Posterior Cingulate Gyrus	Right	NC - OC	0.0844	0.0883	-0.0953	0.2642	0.9564	32.2044	0.3460	0.4325	
Superior Temporal Gyrus	Right	NC - OC	0.3573	0.2148	-0.0801	0.7947	1.6633	32.3097	0.1059	0.4325	

Group effects on CAT 12 GM-values.

roi	*hemisphere*	*contrast*	*estimate*	*SE*	*lower.CL*	*upper.CL*	*t.ratio*	*df*	*p.value*	*p.fdr*	*sig*
Anterior Cingulate Gyrus	Left	NC - OC	0.0794	0.1466	-0.2192	0.3780	0.5418	32.0689	0.5917	0.5917	
Anterior Insula	Left	NC - OC	0.1579	0.1244	-0.0955	0.4113	1.2692	32.0925	0.2135	0.2989	
Middle Cingulate Gyrus	Left	NC - OC	0.2759	0.1282	0.0147	0.5370	2.1515	32.1412	0.0391	0.1367	
Middle Frontal Gyrus	Left	NC - OC	0.4862	0.4535	-0.4374	1.4098	1.0723	32.0843	0.2916	0.3402	
Middle Temporal Gyrus	Left	NC - OC	0.5940	0.3649	-0.1490	1.3371	1.6280	32.2441	0.1133	0.2643	
Posterior Cingulate Gyrus	Left	NC - OC	0.1451	0.0993	-0.0572	0.3474	1.4608	32.1352	0.1538	0.2691	
Posterior Insula	Left	NC - OC	0.1388	0.0585	0.0198	0.2579	2.3749	32.0926	0.0237	0.1367	

***Table 10**. Group effects on Freesurfer GM-values for group NC and OC.*

Right hemisphere = upper figure; left hemisphere lower figure.

NC = natural cycle; OC = oral contraceptives

Group effects on Freesurfer GM-values.

roi	*hemisphere*	*contrast*	*estimate*	*SE*	*lower.CL*	*upper.CL*	*t.ratio*	*df*	*p.value*	*p.fdr*	*sig*
Caudal Anterior Cingulate	Right	NC - OC	0.0349	0.1593	-0.2895	0.3593	0.2193	32.1255	0.8278	0.9396	
Caudal Middle Frontal	Right	NC - OC	0.1646	0.3928	-0.6356	0.9647	0.4189	32.0895	0.6781	0.9396	
Isthmus Cingulate	Right	NC - OC	-0.0104	0.1364	-0.2881	0.2673	-0.0764	32.0976	0.9396	0.9396	
Posterior Cingulate	Right	NC - OC	0.1118	0.1618	-0.2178	0.4414	0.6907	32.0784	0.4947	0.8658	
Rostral Anterior Cingulate	Right	NC - OC	0.2293	0.1184	-0.0118	0.4704	1.9371	32.0362	0.0616	0.2476	
Rostral Middle Frontal	Right	NC - OC	0.4471	0.4943	-0.5597	1.4540	0.9045	32.0507	0.3725	0.8658	
Superior Temporal	Right	NC - OC	0.7274	0.3892	-0.0652	1.5200	1.8691	32.1201	0.0707	0.2476	

Group effects on Freesurfer GM-values.

roi	*hemisphere*	*contrast*	*estimate*	*SE*	*lower.CL*	*upper.CL*	*t.ratio*	*df*	*p.value*	*p.fdr*	*sig*
Caudal Anterior Cingulate	Left	NC - OC	-0.1226	0.1557	-0.4398	0.1946	-0.7871	32.1230	0.4370	0.4994	
Caudal Middle Frontal	Left	NC - OC	0.2914	0.3488	-0.4189	1.0018	0.8356	32.1111	0.4096	0.4994	
Insula	Left	NC - OC	0.3776	0.1591	0.0535	0.7017	2.3734	32.0077	0.0238	0.1903	
Isthmus Cingulate	Left	NC - OC	0.1581	0.1422	-0.1316	0.4478	1.1112	32.0767	0.2747	0.4396	
Middle Temporal	Left	NC - OC	0.4674	0.3198	-0.1840	1.1187	1.4614	32.0441	0.1536	0.4396	
Posterior Cingulate	Left	NC - OC	0.1766	0.1559	-0.1411	0.4942	1.1322	32.0488	0.2659	0.4396	
Rostral Anterior Cingulate	Left	NC - OC	-0.0381	0.1475	-0.3386	0.2623	-0.2586	32.0385	0.7976	0.7976	
Rostral Middle Frontal	Left	NC - OC	0.5942	0.5110	-0.4465	1.6349	1.1628	32.0764	0.2535	0.4396	

CURRICULUM VITAE

The CV was removed from the final print version.

The CV was removed from the final print version.

The CV was removed from the final print version.

The CV was removed from the final print version.

The CV was removed from the final print version.

The CV was removed from the final print version.

The CV was removed from the final print version.

The CV was removed from the final print version.

The CV was removed from the final print version.

The CV was removed from the final print version.

The CV was removed from the final print version.

The CV was removed from the final print version.

The CV was removed from the final print version.

Verzeichnis der akademischen Lehrer

Meine akademischen Lehrer waren die Damen/Herren in

Marburg, Deutschland:
Adamkiewicz
Bauer
Besedowsky
Bette
Bertoune
Braun
Brehm
Cetin
Culmsee
Daut
Decher
Del Rey
Depboylu
Dodel
Elsässer
Halaszovich
Höglinger
Jacob
Jansen
Krach
Leitner
Lill
Maisner
Milani
Mühlenhoff
Müller
Oberwinkler
Oertel
Oliver

Pagenstecher
Plant
Preisig-Müller
Schäfer
Schütz
Schratt
Schwarting
Sommer
Stiewe
Suske
Westermann
Weihe
Wrocklage

Stockholm, Schweden:
Axelsson

Lübeck, Deutschland:
Solbach

ACKNOWLEDGMENTS

The acknowledgments were removed from the final print version.

The acknowledgments were removed from the final print version.

EHRENWÖRTLICHE ERKLÄRUNG

This part was removed from the final print version.

This part was removed from the final print version.