Oral Contraceptive Androgenicity and Cognitive Performance Among Women

A Thesis

Submitted to the Faculty of

Drexel University

by

Katherine Alvarez

in partial fulfillment of the

requirements for the degree

of

Master of Science in Psychology

March 2014

DEDICATIONS

For my parents Dagoberto and Elsi.

For crossing an ocean and abandoning a life,
so that I may build a better one.

Consider this a blueprint.

ACKNOWLEDGEMENTS

The completion of this projection could not have been possible without the support and assistance of several individuals. I would like to express my gratitude to my advisor, Mary V. Spiers, Ph.D., for her guidance throughout this project. This study would not have been possible without the expertise of my committee members, Monica E. Calkins, Ph.D., and Pamela A. Geller Ph.D. I would also like to thank Ruben C. Gur Ph.D. and the Brain Behavior Laboratory at the University of Pennsylvania for allowing me the opportunity to complete this project, in particular Dr. Calkins, Kosha Ruparel and Chad Jackson for guiding me through the CNB database and its beautiful complexities.

I would like to extend my deepest gratitude to Kelly E. Jones, Ph.D. for her support with data organization and analysis and for her contagious enthusiasm for statistics.

I would also like to thank the individuals who have provided personal support throughout this project. First and foremost, I am eternally grateful to my dearest friend James for his unwavering support and his selfless investment in my personal and professional growth. To my friends Jennifer, Sabrina and Charles for always providing much-needed levity, encouragement and above all, a listening ear. Finally, I wish to express my sincerest thanks to my partner Peter for his invaluable emotional support throughout this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Oral Contraceptive Androgenicity and Cognitive Performance Among Women

Katherine Alvarez, B.S.

Mary Spiers, Ph.D.

Oral contraceptives (OCs) may lower endogenous sex hormones while introducing synthetic progestins with varying degrees of biological androgenicity or masculinizing effects. No study has examined the relationship between OC androgenicity and female performance on visuospatial (i.e., line orientation, matrix reasoning), verbal (i.e., word memory, analogical reasoning) and facial expression processing (i.e., emotion recognition, emotion intensity differentiation). We hypothesized a positive relationship between androgen status and line orientation, matrix reasoning and analogical reasoning performance and a negative relationship between androgen status and word memory, emotion recognition and emotion intensity differentiation. One hundred seventy nine females (82.1% Caucasian) ages 15-21 from the Philadelphia Neurodevelopmental Cohort were assigned to groups based on the androgenicity of their OC progestins: "highly androgenic" (*Levonorgestrel*), "androgenic" (*Norethindrone*), "antiandrogenic" (*Drospirenone*) and controls. A multivariate analysis of variance revealed that composite emotion recognition performance was significantly different among groups $F(6,348) = 2.23$, $p < .05$; *Wilk's* $\Lambda = 0.927$, *partial* $\eta 2 = .037$. Univariate one-way analyses of variance revealed marginally significant emotion recognition ($p = .107$, *partial* $\eta 2 = .034$) and emotion intensity differentiation ($p = .064$, *partial* $\eta 2 = .040$) among groups. Tukey's post hoc tests revealed that "highly androgenic" OC users ($M = 81.85$, $SD = 8.21$) outperformed controls ($M = 77.05$, $SD = 8.25$) on emotion intensity differentiation. There were no differences in line orientation, matrix reasoning, word memory, analogical

reasoning, anger recognition or anger intensity differentiation. Findings suggest that some aspects of facial expression processing (i.e., emotion intensity) may be more sensitive to sex hormone changes than others (i.e., emotion recognition). Emotion intensity differentiation patterns suggest that OCs may not exert clinically meaningful androgenic effects on cognition.

CHAPTER 1: BACKGROUND AND LITERATURE REVIEW

Oral contraceptives (OCs) or "birth control pills" present an avenue to explore female cognitive performance because they alter sex hormones that may be implicated in cognition (Erlanger, Kutner, & Jacobs, 1999). The following literature review examines female cognitive performance changes in the visuospatial, verbal, and social facial expression processing domains with natural and pharmacologically induced hormonal fluctuations. The hormone-altering effects of OCs are discussed in relation to the pharmacology and the biological androgenicity parameters of the three progestins of interest to the current study – *Levonorgestrel*, *Norethindrone* and *Drospirenone.* Lastly, existing literature on the androgenizing "masculinizing" and antiandrogenizing "anti-masculinizing" cognitive effects of different OC progestins is summarized across visuospatial, verbal, and facial expression processing domains.

1.1 Activating Effects of Sex Hormones on Cognition

Sex hormones (estrogen, progesterone, testosterone) may differentially "organize" the prenatal brains of males and females, with higher testosterone levels in males resulting in "masculinized" cognitive patterns and the absence or near absence of testosterone in females resulting in "feminized" cognitive patterns. Sex hormone fluctuations throughout the life span may "activate" or facilitate processing in previously organized brain structures and neural networks leading to cognitive performance differences between males and females (Kelso, Nicholls, & Warne, 1999; Kimura, 1996). The organizing effects of sex hormones are believed to be permanent whereas the activating effects are thought of as temporary and reversible. Throughout the present

study “female-typical” (FT) and “male-typical” (MT) will used in reference to cognitive performance patterns that are typical of that sex in comparison to the opposite sex.

1.1.1 Activating Effects of the Menstrual Cycle

The “activating” role of sex hormones is supported by studies of cognitive performance variations in females across the menstrual cycle. The menstrual cycle is divided into three phases (menstrual, follicular, and luteal) that correspond with patterns of hormonal fluctuations. Estrogen and progesterone are lower during the menstrual phase, gradually increase in the follicular phase and peak at mid-cycle during the luteal phase (Hampson & Young, 2007). Although the literature mainly reports cycle-related variations in estrogen and progesterone, testosterone levels also fluctuate throughout the menstrual cycle. Studies suggest that in women, testosterone levels are at their lowest during the late luteal, menstrual and early follicular phases, progressively increase throughout the follicular phase and peak between the late follicular and early luteal phases of the menstrual cycle (Guay, 2002; Massafra, De Felice, Agnusdei, Gioia, & Bagnoli, 1999; Rothman et al., 2011; Salonia et al., 2008).

Although not all findings are consistent, the higher estrogen and progesterone (luteal) phase of the menstrual cycle has been associated with more accurate (FT) verbal fluency performance (Maki, Rich, & Rosenbaum, 2002; Rosenberg & Park, 2002; Solís-Ortiz & Corsi-Cabrera, 2008). Conversely, studies show that during menstruation and the early follicular phase, when estrogen and progesterone are at their lowest, there is a slight shift in women’s cognitive performance towards the MT pattern of increased accuracy in visuospatial performance (Hampson, 1990; Hausmann, Slabbekoorn, Van

Goozen, Cohen-Kettenis, & Gunturkun, 2000; Maki et al., 2002; Silverman & Phillips, 1993) (See Figure 1).

Most studies exploring sex hormone fluctuations and female social cognition performance via facial expression processing (FEP) report a negative association between FEP performance and estrogen and progesterone levels. Contrary to the cycle-related performance patterns observed with visuospatial and verbal fluency measures, more accurate (FT) FEP performance is observed during the early follicular cycle phase when estrogen and progesterone levels are lower and less accurate (MT) FEP performance is observed during the luteal phase when estrogen and progesterone are higher (See Figure 1). Although these studies do not report cycle-related testosterone variations, it is important to note that in women, testosterone levels are also lower during the early follicular phase and peak between the late follicular and early luteal phases (Guay, 2002; Massafra et al., 1999; Rothman et al., 2011; Salonia et al., 2008). A series of studies by Derntl and colleagues (2008a, 2008b, 2013) consistently report more accurate FEP performance in the early follicular phase than in the luteal phase. Similarly, a 2009 study by Guapo and colleagues reported more accurate recognition of angry faces and a nearly significant trend of more accurate recognition of fearful faces in the early follicular phase. Only one study to date has reported incongruous results of more accurate fear recognition during the luteal than the follicular phase (Pearson & Lewis, 2005). Unlike the aforementioned studies, Pearson and Lewis (2005) did not confirm estrogen and progesterone levels via hormonal assays and did not estimate cycle phase on testing day. Cycle phase was estimated at a later date based on participant-initiated email reports of the onset of menses following study participation, potentially introduced additional error

to the study's cycle estimates and confounding results. For example, reporting a menses date that is off by a couple of days is enough to erroneously place a late follicular/early luteal (higher estrogen and progesterone) participant in the "low hormone" follicular group or to place a late luteal - approaching menses (lower estrogen and progesterone) participant in the "high hormone" luteal group.

1.1.2 Activating Effects of Testosterone

The "activating" role of sex hormones on female cognition has also been supported by studies demonstrating a shift towards male-typical performance patterns with higher endogenous (naturally occurring) testosterone levels and exposure to exogenous (synthetic) testosterone. The literature suggests more accurate performance on visuospatial tasks in women with higher salivary testosterone concentrations (Gouchie & Kimura, 1991; Moffat & Hampson, 1996; Ostatnikova, Putz, Celec, & Hodosy, 2009) as well as in women who are exposed to synthetic testosterone (Aleman, Bronk, Kessels, Koppeschaar, & Van Honk, 2004; Hirshman et al., 2004). Additionally, in a series of testosterone administration studies in women, Van Honk and colleagues (2007, 2011) reported decreased sensitivity to fear facial expressions, a significant reduction in sensitivity to facial anger, and significant impairment in the ability to infer emotions, intentions, and other mental states from the eyes region of the face after a single dose of testosterone.

Lastly, compelling evidence of the "activating" role of sex hormones on cognition has been gathered from studies of female-to-male transsexuals undergoing testosterone treatments. In their 1994 study, Van Goozen and colleagues reported male-typical

cognitive patterns of increased accuracy in visuospatial performance (measured via the Mental Rotation Task (MRT)) and declines in female-typical verbal fluency performance after three months of testosterone treatments. In a follow-up study, Slabbekoorn et al. (1999) replicated Van Goozen's findings of stronger visuospatial performance in female-to-male transsexuals and demonstrated that these effects did not quickly reverse; lasting up to five weeks after terminating testosterone treatments. No deteriorating effects on verbal fluency were identified (Slabbekoorn, Van Goozen, Megens, Gooren, & Cohen-Kettenis, 1999). There are studies however, that report no relationship between testosterone levels and spatial or verbal performance in women (Halari et al., 2006; Puts et al., 2010).

1.1.3 Sex Hormones and Brain Activation

Currently, one of the most comprehensive theoretical models to integrate sex hormones, brain activation and sex differences in visuospatial and verbal cognitive performance may be Hampson's (1990) differential activation theory. Hampson (1990) theorized that sex hormone differences in males and females may differentially activate the cerebral hemispheres, leading to sex differences in cognitive performance. Hampson suggested that higher estrogen and progesterone levels (e.g., luteal phase) may cause increased activation of the left hemisphere (where language is typically mediated) while simultaneously inhibiting right hemisphere functioning (where visuospatial processing is typically mediated). This differential activation pattern is believed to result in females outperforming males in verbal cognitive tasks as well as females with higher estrogen and progesterone levels (greater left hemisphere activation) outperforming females with

lower estrogen and progesterone levels (decreased left hemisphere activation) in verbal cognitive tasks. Conversely, lower estrogen and progesterone levels (i.e., menstrual phase) were theorized to result in decreased left hemisphere activation and enhanced right hemisphere activation, leading to increased visuospatial performance accuracy in males over females as well as in females with lower estrogen and progesterone levels over females with higher estrogen and progesterone levels (Hampson, 1990).

Although females consistently outperform males on FEP measures and do so as early as infancy, the underlying mechanisms are not well understood (Hall, Hutton, & Morgan, 2010; McClure, 2000). Currently, there is no leading model to account for sex differences in FEP performance however some researchers theorize that sex hormones may play a moderating role (Mareckova et al., 2012). For example, sex differences in exploratory eye movements tend to emerge alongside the hormonal changes of puberty (Miyahira, Morita, Yamaguchi, Nonaka, & Maeda, 2000) and eye-tracking methods have revealed that females tend to show increased fixations and dwell time to the eyes, which may account for their stronger performance on facial expression recognition (Hall et al., 2010). Interestingly, a single dose of synthetic testosterone has been shown to significantly impair women's ability to infer emotions specifically from the eyes region of the face (Van Honk et al., 2011). Additionally, poor FEP performance has been observed after testosterone administration (Van Honk & Schutter, 2007) and during the luteal cycle phase when natural testosterone levels are higher (Dernt et al., 2008a, 2008b, 2013; Guapo et al., 2009).

In conclusion, the literature suggests that sex hormones (estrogen, progesterone, testosterone) fluctuations are related to changes in female cognitive performance.

Menstrual cycle related increases in estrogen and progesterone at mid-cycle are associated with increased verbal fluency (female-typical) performance whereas decreased estrogen and progesterone during the menstrual and early follicular phases are associated with increased visuospatial (male-typical) performance. A reversed pattern of decreased FEP accuracy (male-typical) is observed at mid-cycle when estrogen and progesterone as well as testosterone are higher. The relationship between sex hormones and female cognition is supported by studies showing a shift towards male-typical cognitive performance in the verbal fluency, visuospatial and FEP domains with synthetic testosterone administration. Although the underlying mechanism is unknown, these collective findings suggest a relationship between natural and exogenous (synthetic) sex hormones and female cognitive performance. Oral contraceptives (OCs) are a popular group of hormone-altering medications used primarily for pregnancy prevention. As discussed below, OCs alter endogenous sex hormones and contain exogenous (synthetic) hormones, warranting research of their usage and female cognitive performance.

1.2 Hormonal Effects of Oral Contraceptives

Although OCs exist in progestin-only pills ("mini-pills") for patients in whom estrogen is not tolerated or contraindicated (Zurawin & Ayensu-Coker, 2007), most modern OCs available in the U.S. contain a combination of ethinyl estradiol (synthetic estrogen) and 1 of 10 progestin types (synthetic progesterones) (Amy & Tripathi, 2009).

The progestins in OCs are mainly responsible for their contraceptive effect while the accompanying ethinyl estradiol improve cycle control and augment the progestin's contraceptive efficacy (Zurawin & Ayensu-Coker, 2007). The estradiol/progestin

combination in OCs disrupts a female's natural menstrual cycle, preventing ovulation and the hormonal fluctuations of each cycle phase. As a result, circulating levels of estrogen and progesterone are reduced by at least 50% than those found in regularly menstruating women (Gordon & Lee, 1993). The majority of studies comparing the hormonal profiles of OC users and non-users report lower estrogen, progesterone and testosterone levels in OC users (Kuhl, Gahn, Romberg, Marz & Taubert, 1985; Mordecai, Rubin, & Maki 2008; Schultheis, Dargel, & Rhode, 2003; Thorneycroft et al., 1999; Timmons, Hamadeh, Devries, & Tarnopolsky, 2005) that closely resemble menstrual-like levels (Schultheiss et al., 2003; Thorneycroft et al., 1999). Additionally, OCs contain different progestins with varying degrees of androgenic (masculinizing) or antiandrogenic (anti-masculinizing) biological effects that may further alter female sex hormone levels (Batur, Elder, & Mayer, 2003).

1.2.1 Oral Contraceptive Progestins: Generations and Formulations

The progestins used in OCs are divided into so-called generations based on when they were developed and their biological androgenicity levels. First generation (no longer prescribed for pregnancy prevention), second generation (e.g., *Norethindrone, Levonorgestrel*), third generation (e.g., *Norgestimate, Desogestrel)* and fourth or new generation (e.g., *Drospirenone*) (Kiley & Hammond, 2007).

Second generation progestins are chemically related to testosterone and are considered to be the most biologically androgenic progestins on the market. Third generation progestins were developed later and contain structural modifications to lower their androgenicity and are therefore considered less biologically androgenic. Lastly,

fourth or new generation progestins were created to contain antiandrogenic components and are therefore considered biologically antiandrogenic (Batur et al., 2003; Rowlands, 2003). In conclusion, all but fourth generation (antiandrogenic) progestins are considered biologically androgenic to some degree.

In addition to differences in androgenicity, OCs also differ in the amount of progestins delivered across each pack. Most OCs are either triphasic - delivering increasing amounts of progestins from week to week or monophasic formulations that deliver a constant dose of ethinyl estradiol and progestin across the pack (Batur et al., 2003; Kiley & Hammond, 2007). Historically, monophasic OCs contained 21 active pills and 7 inactive pills to mimic the prototypical 28-day menstrual cycle however, monophasic formulations are also available in 24 active/ 4 inactive pill packs (Zurawin & Ayensu-Coker, 2007) as well as 91-day extended cycle packs (e.g., *Seasonale, Seasonique*) containing 84 active pills and 7 inactive pills.

For the purposes of this study, only monophasic OCs containing the "androgenic" second generation progestins *Levonorgestrel* and *Norethindrone* and the "antiandrogenic" fourth generation progestin *Drospirenone* will be considered as these progestins represent the two extremes of the OC androgenicity spectrum and have not been thoroughly studied within the context of female cognitive performance.

In summary, OCs alter female sex hormones by a) preventing ovulation and the natural hormonal fluctuations of the menstrual cycle, b) lowering sex hormones to menstrual-like levels, and c) exposing users to progestins with varying degrees of biological androgenicity. Given that all OCs prevent hormonal fluctuations and lower sex hormones, warrants study of their androgenicity differences, especially as they relate

to changes in female cognitive performance. Androgenic and antiandrogenic OC formulations are especially relevant to exploring sex hormone related cognitive changes as they represent the two extremes of the OC androgenicity spectrum.

1.3 Oral Contraceptive Progestins: Biological Androgenicity/Antiandrogenicity

At the cosmetic level, signs of a progestin's biological androgenicity or masculinizing effects may include acne, hirsutism (abnormal hair growth), and weight gain. At a pharmacological level, the literature primarily cites the following as markers of a progestin's biological androgenicity: a) binding to androgen receptors, and b) binding to sex hormone binding globulin (SHBG) and subsequent changes on free testosterone levels (Greer, Modugno, Allen & Ness, 2005). Some researchers however, argue that no currently available OC has clinically meaningful androgenic effects (Stanczyk, 2003; Thorneycroft et al., 1999). In the following paragraphs the biological androgenicity parameters and the counterarguments to each parameter will be discussed for the three progestins of interest to the present study: *Levonorgestrel*, *Norethindrone* and *Drospirenone*.

1.3.1 Progestin Androgenicity and Relative Binding Affinity

A progestin's biological androgenicity is mainly determined by measuring its androgenic relative binding affinity (RBA) or capacity to bind to androgen receptors in the body and central nervous system compared with that of testosterone (Carr, 1998; Kumar, Koide, Tsong & Sundaram, 2000). *Levonorgestrel* is chemically related to testosterone and is generally considered in the literature as one of the most androgenic

OC progestins (Sitruk-Ware, 2008). The binding capacity of *Levonorgestrel* to androgen receptors has been gathered from animal studies (Kumar et al., 2000; Phillips, Hahn & Klimek, 1987; Phillips, Demarest, Hahn, Wong & McGuire, 1990) and is reported to be approximately 70% that of testosterone (Kumar et al., 2000). *Norethindrone* is also chemically related to testosterone and considered an androgenic progestin (Stanczyk, 2003) albeit, less androgenic than *Levonorgestrel.* By compiling data from various studies of progestin androgenicity, Greer and colleagues (2005) concluded that *Levonorgestrel* is approximately 1.25 times more androgenic than *Norethindrone* (Greer, Modugno, Allen & Ness, 2005).

Drospirenone on the other hand, was designed to be biologically antiandrogenic by competitively inhibiting binding to androgen receptors in the body and central nervous system (Batur et al., 2003; Krattenmacher, 2000; Rowlands, 2003). Animal studies of the binding capacity of antiandrogenic *Drospirenone* to androgen receptors report binding affinities as low as 0.6% to 2% that of testosterone (Krattenmacher, 2000; Muhn, Krattenmacher, Beier, Elger, & Schillinger, 1995).

Some researchers however, criticize the use of androgen receptor binding data from animal studies as evidence of a progestin's biological androgenicity, citing that these findings may not be relevant to humans given major differences between rats' tissues and female tissues and given that these studies employed considerably higher progestin doses than those found in OCs (Stanczyk, 2003; Thorneycroft et al., 1999). For example, the *Levonorgestrel* dosage needed to cause prostate growth in the male rat is approximately a thousand times higher than the dosage needed to inhibit ovulation in the

female rat; demonstrating that even in rats, the contraceptive dosage of *Levonorgestrel* would have no androgenic effects (Thorneycroft et al., 1999).

1.3.2 Progestin Androgenicity and Sex Hormone Binding Globulin

The second biggest marker of a progestin's biological androgenicity is derived from measuring levels of sex hormone binding globulin (SHBG), the major carrier protein for endogenous testosterone. SHBG binds to free testosterone and this determines the amount of free testosterone that remains in the blood (Van der Vange, Blankenstein, Kloosterboer, Haspels, & Thijssen, 1990). Free testosterone is defined as the fraction of testosterone that is bioavailable- or able to pass from the blood into bodily tissues where it can exert its biological (i.e., androgenic) activity (Baird, Horton, Longcope, & Tait, 1969). Free testosterone levels are therefore considered indicators of androgen concentrations in humans (Cumniings & Wall, 1985). Elevated estrogen levels increase SHBG concentrations leading to greater free testosterone binding and thus a decrease in free testosterone levels. Conversely, elevated androgens (i.e., testosterone) reduce SHBG levels resulting in increased free testosterone levels in the blood (Anderson, 1974; Lobl, 1981).

Second generation oral contraceptives containing progestins like *Levonorgestrel* and *Norethindrone* are considered biologically androgenic because they counteract the increase in SHBG brought on by their ethinyl estradiol component and result in more bioavailable testosterone (Janaud, Rouffy, Upmalis, & Dain, 1992; Makhzangy, Wynn, & Lawrence, 1979; Palatsi et al., 1984; Thorneycroft et al., 1999). Conversely, *Drospirenone* does not bind to SHBG therefore OCs containing *Drospirenone* do not

reduce the SHBG increase brought on by their estradiol component (Krattenmacher, 2000). *Drospirenone* is therefore considered biologically antiandrogenic, as greater SHBG levels equate with more binding of circulating testosterone and therefore less bioavailable free testosterone for action at the receptor level (Sitruk-Ware, 2008).

Although research findings suggest that older "androgenic" progestins result in lower SHBG levels than those observed with newer "antiandrogenic" progestins and estradiol alone, researchers warn against using these parameters as evidence of biological androgenicity, citing that all OCs (irrespective of progestin type and dose) increase SHBG and decrease all endogenous sex hormones including free testosterone levels (Stanczyk, 2003; Thorneycroft et al., 1999). Various research studies report lower free testosterone levels in OC users than in naturally cycling females regardless of progestin generation (Knopp et al., 2001; Thorneycroft et al., 1999; Van der Vange et al., 1990; Van Rooijen, Silveira, Hamsten, & Bremme, 2004; Wiegratz et al., 2003) suggesting that OCs may actually exert hypoandrogenic biological effects (Bachmann, 2002; Thorneycroft et al., 1999). Additionally, the hypoandrogenic effects of OCs may account for their widespread use for the treatment of hyperandrogenicity-related disorders like hirsutism (excessive hair growth) and acne (Stanczyk, 2003; Thorneycroft et al., 1999).

Some researchers however, argue that the standard hormonal assays commonly used in research were designed to measure endogenous estrogen, progesterone and testosterone and are therefore not sensitive enough to accurately detect synthetic OC sex hormones (Hampson & Young, 2007). For example, a 2003 review of ten widely used hormonal assays for testosterone found that only two met reasonable validity criteria for measuring testosterone in females (Taieb et al., 2003).

In summary, the literature recognizes *Levonorgestrel* as the most biologically androgenic progestin, *Norethindrone* as androgenic and *Drospirenone* as an antiandrogenic progestin. Some researchers however, claim that the methods used to establish those androgenicity parameters are invalid in humans and that all OCs (androgenic and antiandrogenic formulations) lower endogenous testosterone levels and are therefore biologically hypoandrogenic. Lastly, there is concern that standard hormonal assays are not sensitive enough to measure synthetic OC sex hormones. Overall, these contradictions in the literature suggest that the true hormonal profiles and biological androgenicity of OC users is unclear. This warrants the study of the cognitive performance patterns of users of highly androgenic and antiandrogenic OCs, as findings of "androgenized" (masculinized) or "antiandrogenized" (antimasculinized) cognitive performance will shed light on the true biological androgenicity of different OC types.

1.4 Oral Contraceptives and Female Performance Across Cognitive Domains

Although the mechanism of action remains unknown, the literature suggests that androgenic OCs have an "androgenizing" or masculinizing effect on female cognitive performance for some measures with established sex differences. In this context, "androgenizing" refers solely to the shift towards male-typical and away from female-typical cognitive performance observed in users androgenic OCs. Additionally, some studies suggest that antiandrogenic OC formulations may have "antiandrogenizing" or antimasculinizing cognitive effects; shifting performance on some cognitive measures towards female-typical patterns. Antiandrogenic OCs are newer and less studied than androgenic OCs. It is important to note that few studies in the OC literature have

controlled for OC formulation; grouping females with heterogeneous hormonal profiles without considering how their hormonal differences may relate to changes in cognitive performance. Additionally, few studies have taken into account progestin generation and subsequent variability in the biological androgenicity of different OCs.

1.4.1Visuospatial Performance and Sex Hormones

Although not all findings are consistent, more accurate visuospatial performance has been observed in users of androgenic OCs compared to non-users (McFadden, 2000; Silverman & Phillips, 1993, Wharton et al., 2008). Visuopatial performance is generally defined as the ability to generate, represent, transform and recall information about an object's shape in space (Linn & Petersen, 1995) and can be broken down into specific categories, among them mental rotation (Linn & Petersen, 1995) and spatial orientation (Lawton & Morrin, 1999). Mental rotation refers to the ability to mentally rotate two or three dimensional figures (Linn & Petersen, 1995) while spatial orientation is defined as a complex set of skills used to locate an object with respect to a point of reference or a system of coordinates (Lawton & Morrin, 1999).

The Mental Rotation Task (MRT) by Shepard and Metzler (1971) is a commonly used spatial measure in OC studies because it has been shown to produce the largest and most consistent sex differences in spatial performance favoring males (Linn & Petersen 1985; Masters & Sanders, 1993; Voyer, Voyer, & Bryden, 1995). Increased accuracy in female MRT performance have been observed with lower endogenous sex hormone levels in the menstrual phase (Silverman & Phillips, 1993), higher endogenous testosterone levels (Gouchie & Kimura, 1991; Moffat & Hampson, 1996; Ostatnikova et

al., 2002) synthetic testosterone administration (Aleman et al., 2004) and testosterone treatments (Van Goozen et al., 1994). The majority of OC studies exploring visuospatial performance have focused on the MRT and on users of androgenic OCs (Gordon & Lee, 1993; McFadden, 2000; Mordecai et al., 2008; Rosenberg & Park, 2002; Silverman & Phillips, 1993) and have reported contradictory findings. Silverman and Phillips (1993) and McFadden (2000) reported that OC users outperformed non-users on the MRT whereas the remaining studies reported no differences on the MRT performance of OC users and nonusers (Gordon & Lee, 1993; Mordecai et al., 2008; Rosenberg & Park, 2002)

To date, only two studies have examined OC progestin androgenicity and spatial performance via the MRT. Wharton and colleagues (2008) reported a trend for women using highly androgenic OCs to outperform nonusers but most importantly, women using antiandrogenic OCs were found to perform significantly worse than nonusers; suggesting that the antiandrogenic components of newer OCs may actually hinder male-typical spatial performance. The second study by Griksiene and Ruksensas (2011) did not replicate Wharton's findings, but a similar trend of hindered MRT performance was observed in antiandrogenic OC users.

The Judgement of Line Orientation Test (JLO) (Benton, Varney & Hamsher, 1978) and the Judgement of Line Angle and Position Test (JLAP) (Collaer & Nelson, 2002) are a set of similar paper and pencil spatial orientation measures that have also revealed sex differences favoring males (Collaer & Nelson, 2002; Lindgren & Benton, 1980; Glamser & Turner, 1995). Sex differences favoring males have also been observed in computerized versions of these measures- JLAP-15 (Collaer, Reimers, & Manning,

2007) and CJOLO (Gur et al., 2001; Gur et al., 2010; Gur et al., 2012). To date, the only study to examine JLO performance and OC use did not report significant differences between OC users and non-users (Goyette, McCoy, Kennedy, & Sullivan, 2011). This study however, did not control for OC formulation or progestin generation.

The Raven's Progressive Matrices Test (RPM) (Raven, 1960) is a paper and pencil measure of nonverbal reasoning that has shown consistent performance differences favoring males in two meta-analyses (Irwing & Lynn, 2005; Lynn & Irwing, 2004). A recent computerized version of the RPM has also revealed sex differences favoring males in 18-21 year olds (Gur et al., 2012). Some researchers suggest that these sex differences in performance may be related to the fact that many items on the RPM are expressed in spatial form (e.g., geometric figures) (Abad, Colom, Rebollo, & Escorial, 2004; Hyde, Fennema, & Lamon, 1990), warranting additional research to explore whether there is a relationship between performance on the RPM and OC androgenicity or antiandrogenicity. To date, no studies have examined RPM performance and sex hormone changes or OC use.

1.4.2 Verbal Performance and Sex Hormones

Verbal fluency measures are commonly used to explore sex differences in verbal performance and typically assess participants' ability to rapidly produce words that start with a given letter (e.g. F) over a 60 second period while they monitor word production to ensure no words are repeated (Gourovitch, Goldberg, & Weinberger, 1996). Females generally tend to outperform males on tests of verbal fluency (Halari et al., 2006; Hampson, 1990; Hausmann, Schoofs, Rosenthal & Jordan, 2009; Weiss, Kemmler,

Deisenhammer, Fleischhacker, & Delazer, 2003). Some studies report increased verbal fluency with higher estrogen and progesterone levels (Hampson, 1990; Maki et al., 2002; Rosenberg & Park, 2002; Solís-Ortiz & Corsi-Cabrera, 2008) while others report no relationship between sex hormone levels and verbal fluency performance (Halari et al., 2005; Mordecai et al., 2008). Likewise, some researchers report decreased verbal fluency with testosterone treatments (Van Goozen et al., 1994; Van Goozen, Cohen-Kettenis, Gooren, Frijda, & Van de Poll, 1995) while others report no relationship between testosterone and verbal fluency (Moffat & Hampson, 1996; Slabekoorn et al., 1999). Few studies have explored verbal fluency and OC use. Mordecai et al. (2008) reported no difference in the verbal fluency performance of OC users and non-users but did not control for OC formulation or progestin generation. To date, the only study of verbal fluency to control for OC androgenicity levels reported that naturally cycling females outperformed females using androgenic OCs and females using antiandrogenic OCs (Griksiene & Ruksensas, 2011).

Some research studies report that females tend to outperform males on measures of verbal episodic memory. Verbal episodic memory is commonly assessed by first presenting participants with a word and later asking them to recall or recognize the earlier-presented word (Herlitz & Rehnman, 2008). Verbal episodic memory studies using paper and pencil tests of word recognition report sex differences favoring females when words are presented auditorily (Herlitz, Nilsson, & Bäckman, 1997) and visually (Temple & Cornish, 1993; Zelinski, Gilewski, & Schaie, 1993). Additionally, a series of studies exploring verbal episodic memory for visually presented words via a computerized word recognition task- the Penn Word Memory Test (CPW) have reported

sex differences generally favoring females (Gur et al., 2001; Gur et al., 2010) and specifically among 20-21 year olds (Gur et al., 2012). Verbal episodic memory is not well studied within the context of hormonal fluctuations or OC use. In their 2004 study, O'Reilly and colleagues used a visually presented episodic word memory test and reported no differences in performance across two phases of the menstrual cycle (menstrual and luteal). Using a similar visually presented word recognition task, Wharton and colleagues reported no differences in the verbal episodic memory of females across menstrual phases or OC progestins (androgenic vs. antiandrogenic) (Wharton et al., 2008).

Verbal reasoning via verbal analogies is a less studied verbal subdomain. Some studies report sex differences favoring males on measures of verbal analogies (Lim, 1994; Gur et al., 2012) while others report no differences in performance (Feingold, 1988) or in the fMRI activation patterns of males and females during verbal reasoning (Gur et al., 2000). To date, no study has explored analogical reasoning within the context of sex hormone changes or OC use.

1.4.3 Facial Expression Processing and Sex Hormones

Facial expression processing (FEP) or the ability to decode facial expression is broadly defined to include performance on a variety of tasks including facial expression discrimination, recognition and identification (McClure, 2000). In a 2000 meta-analysis, McClure reported a significant female advantage in FEP performance in infancy, childhood and adolescence (McClure, 2000). The female FEP advantage is also seen on computerized measures of emotion identification (Gur et al., 2010; Gur et al., 2012;

Montagne, Kessels, Frigerio, de Haan, & Perrett, 2005). Additionally, females tend to be more sensitive than males at labeling images of facial expressions that could signal threat such as anger and disgust (Montagne et al., 2005).

Changes in female FEP performance have been observed with cycle-related hormonal fluctuations in studies using computerized emotion identification measures. Two studies using Ekman and Friesen's (1976) Pictures of Facial Affect Series reported conflicting results of increased facial emotion recognition accuracy with higher estrogen and progesterone levels (Pearson & Lewis, 2005) and increased facial emotion recognition accuracy with lower estrogen and progesterone levels (Guapo et al., 2009). As previously discussed, Pearson and Lewis' (2005) findings may have been confounded by relying on a delayed self-report to estimate cycle phase during testing day. A series of studies by Derntl and colleagues (2008a, 2008b, 2013) that used a similar emotion identification measure - the computerized Vienna Emotion Recognition Task developed by Gur and colleagues (2002) also reported improved facial emotion recognition accuracy with lower estrogen and progesterone levels.

In addition to using images of full-blown facial expressions, FEP studies have further explored emotion recognition via images of faces with various degrees of emotion intensity. Using computerized measures of morphed facial expressions, Montagne and colleagues (2005) revealed that females were generally more accurate than males at labeling facial expressions and more sensitive at recognizing subtle facial expressions, particularly for anger and disgust (Montagne et al., 2005). Sensitivity in that study was defined as the average amount of expression intensity needed to correctly identify an emotion. Interestingly, Montagne et al. (2005) noted that whereas females outperformed

males at labeling anger and disgust facial expressions with less emotion intensity, they were as accurate as males when overall performance (i.e., all intensities together) on each expression was analysed. A series of follow-up studies by Hoffman et al. (2010) using a similar paradigm did not replicate the previous finding of a general female advantage for faces displaying intense emotions but it did report increased female accuracy for recognizing subtle facial displays of emotion, particularly for low (40%-50%) to mid (60%-70%) intensity disgust faces (Hoffman, Kessler, Eppel, Rukavina, & Traue, 2010). Lastly, a 2007 study using morphed facial expressions reported decreased FEP accuracy in females, particularly for anger faces, after testosterone administration (Van Honk & Schutter, 2007).

To date, no study has compared FEP performance in OC users and non-users. A recent fMRI study of FEP performance reported differences in the eye movements of OC users and non-users, which may represent differences in FEP performance (Mareckova et al., 2012). No behavioral data was provided.

1.5 Rationale for the Present Study

In summary, the literature strongly suggests that naturally and pharmacologically induced changes in sex hormone levels (estrogen, progesterone, testosterone) are related to changes in performance for cognitive measures with established sex differences. These include visuospatial measures (mental rotation, spatial orientation) verbal measures (verbal fluency, verbal episodic memory) and social cognition measures (facial expression processing). Oral contraceptives alter female hormonal profiles and the literature suggests that these changes may contribute to the observed “androgenizing” and

"antiandrogenizing" effects of OCs on female cognition. Currently, the relationship between OC use and female cognitive performance remains unclear given that studies report conflicting findings and that most studies fail to control for OC formulation and the biological androgenicity of OC progestins.

The goal of the present study was to examine the relationship between the biological androgenicity of oral contraceptives and female cognitive performance by exploring highly androgenic OCs, androgenic OCs and the relatively new and less studied antiandrogenic OCs. By exploring antiandrogenic OCs, this study aimed to contribute to the sparse literature on biologically antiandrogenic progestins and female cognitive performance. We hoped that findings would help to clarify whether the observed trend of antiandrogenized (female-typical) MRT performance in users of antiandrogenic OCs is also present on other cognitive measures containing visuospatial components as well as other measures where androgenized (male-typical) patterns are observed with androgenic formulations. The present study enriches the OC literature by providing data for or against established progestin androgenicity classifications. For example, findings consistent with the study's hypotheses would lend support to the established progestin classifications (i.e., second generation OCs are biologically androgenic, fourth generation OCs are biologically antiandrogenic). Conversely, findings that are inconsistent with the study's hypotheses would lend support to the counterargument that all OCs (regardless of progestin generation) are in fact biologically hypoandrogenic. Lastly, the present study aimed to contribute to the OC literature by examining female cognitive performance using previously unexplored (matrix reasoning, analogical reasoning, facial expression processing) and underexplored (spatial

orientation, verbal episodic memory) cognitive measures. Lastly, the present study addressed the need to parse out the relationship between OC progestins and cognitive performance by clarifying whether or not (and which) OCs have "androgenizing" effects on female cognition.

1.6 Aims and Hypotheses of the Present Study

Aim 1: To investigate the effect of OC progestin androgenicity on participants' performance on measures containing visuospatial components – the Penn Line Orientation Test (PLOT) and the Penn Matrix Reasoning Test (PMAT).

Hypothesis 1

Based on previous research, it was hypothesized that participants using OCs with higher levels of androgens (i.e., highly androgenic, androgenic) would exhibit more accurate performance than those using OCs with lower levels of androgens (i.e., antiandrogenic) and controls on the PLOT and the PMAT. More specifically, it was hypothesized that participants in the highly androgenic and androgenic groups would perform similarly to each other. Further, the highly androgenic and androgenic groups would perform better than controls, who would perform better than the antiandrogenic group on the PLOT and the PMAT.

Aim 2: To investigate the effect of OC progestin androgenicity on participants' performance on measures containing verbal components – the Penn Word Memory Test (CPW) and the Short Penn Verbal Reasoning Test (SPVRT).

Hypotheses 2

It was hypothesized that controls would exhibit more accurate performance on the CPW than participants using OCs. Furthermore, no differences in performance were expected among OC users.

Hypothesis 3

Given that males tend to outperform females on verbal reasoning measures, it was hypothesized that users of highly androgenic and androgenic OCs would outperform controls on the SPVRT. More specifically, it was hypothesized that participants in the highly androgenic and androgenic groups would perform similarly to each other. Furthermore, controls were expected to outperform users of antiandrogenic OCs on this task.

Aim 3: To investigate the effect of OC progestin androgenicity on participants' performance on measures containing facial expression processing (FEP) components – the Penn Emotion Recognition Task (ER40) and the Measured Emotion Differentiation Test (MEDF).

Hypothesis 4

It was hypothesized that users of antiandrogenic OCs would perform more accurately than controls, who would perform more accurately than users of highly androgenic and androgenic OCs on the ER40 and the MEDF. More specifically, it was hypothesized that, compared to user of highly androgenic and androgenic OCs, users of

antiandrogenic OCs and controls would exhibit more accurate performance for anger faces on the ER40 and the MEDF.

CHAPTER 2: METHODS

2.1 Procedures

Participant data were obtained from the Philadelphia Neurodevelopmental Cohort Database (PNC) gathered from a large collaborative study between the Center for Applied Genomics at Children's Hospital of Philadelphia (CHOP) and the Brain Behavior Laboratory at the University of Pennsylvania (PENN). The study from which the data were obtained, recruited and tested 9,498 youths ages 8-21 years from the greater Philadelphia area who attended CHOP or CHOP-affiliated clinics and volunteered to participate in genomic studies (Gur et al., 2012).

Once eligibility was determined, the assessment session was scheduled to occur either at home, in the laboratory, or in the community (e.g., quiet room at a public library). Participants met with a trained research coordinator who explained the research procedure and obtained informed consent. Participants then completed a structured clinical interview (GOASSESS) where treatment history including medication use was obtained and entered directly into an interface on the coordinator's Macbook Pro laptop computer.

For the testing part of the study, the research coordinator maintained standardized testing procedures by sitting at the table next to the participant and reading the testing instructions as they appeared on the computer screen. A professional testing environment was maintained by minimizing potential distractors (e.g., cell phones, TV). Using the coordinator's laptop computer, participants completed a 1-hour computerized neurocognitive battery (CNB) developed by the University of Pennsylvania (Gur et al., 2001, 2010) containing 12 tests from a sampling of cognitive domains. The tests were

administered in the following fixed order designed to maintain participant's engagement and reduce fatigue: reading test, motor praxis, emotion identification, continuous performance, face memory, word memory, working memory, conditional exclusion, emotion differentiation, finger taping, matrix reasoning, spatial memory, verbal reasoning, age differentiation, and line orientation. A subset of the tests was of interest to the current study (See Table 1).

2.1.1 Data Consolidation

Data for the present study was provided by PENN and was organized and analyzed by this study's investigator. Participants' performance on the standardized reading test (WRAT4) was provided as standardized scores and raw scores. Standardized scores were used in lieu of raw scores. Participants' performance on the measures of interest was provided as both raw scores and age-based normative values. Because participants were administered one of multiple versions of the measures of interest, with either the same number (CPW, SPVRT, ER40) or a different number (PLOT, PMAT, MEDF) of stimuli, percent correct scores were used in lieu of raw scores.

2.1.2 Data Consolidation for Oral Contraceptive Status

Participants were grouped into one of three OC groups based on their OC's progestin: "highly androgenic" (*Levonorgestrel*) (*n*=34), "androgenic" (*Norethindrone*) (*n*=51) and "antiandrogenic" (*Drospirenone*) (*n*=49). All monophasic antiandrogenic OCs contained 3 mg of *Drospirenone* and all androgenic OCs contained 1 mg of *Norethindrone* (See Table 2). Initially, there were four types of monophasic highly

androgenic OCs: one-month and three-month packs containing 0.15 mg of *Levonorgestrel* and one-month and three-month packs containing 0.10 mg of *Levonorgestrel*. Because these groups did not significantly differ on demographic information or outcome variables of interest, they were collapsed into a single group labeled "highly androgenic" OCs (See Figure 2).

2.1.3 Data Consolidation for Level of Lifetime Depressive Symptoms

Participant data included level of lifetime depressive symptoms endorsed at time of study participation: no depressive symptoms, one or more depressive symptoms, subthreshold depressive disorder (falling short of minimal DSM-IV criteria), threshold depression (reporting no distress or impairment but otherwise meeting DSM-IV criteria) and depression with significant impairment/distress. For the purposes of the present study level of depressive symptoms were grouped as follows: "no depressive symptoms," "mild/moderate depressive symptoms," (one of more symptoms and subthreshold depression), and "severe depressive symptoms" (threshold depression and depression with significant impairment/distress).

2.2 Participants

A subset of one hundred and seventy nine participants from the PNC database were included in the present study. Participants were English-speaking females between the ages of 15 and 21 years (M =17.94, SD = 1.53) who reported using monophasic OCs containing the progestins *Levonorgestrel, Norethindrone* or *Drospirenone* (see Table 2) or who reported not using any kind of hormonal contraceptive at the time of study

participation. Participants were excluded from the study if they reported a serious birth defect or major illness affecting the central nervous system (e.g., epilepsy, hydrocephalus, brain tumor, brain damage, meningitis), endocrine disorders (hypothyroidism, Congenital Adrenal Hyperplasia (CAH)), major health conditions (e.g., Leukemia, HIV, AIDs) and developmental conditions (e.g., Autism, Down's Sydrome, Turner's Syndrome). Additionally, participants were excluded from the study if they reported taking prenatal vitamins, as this could be an indication of possible pregnancy during study participation.

The sample included 147 Caucasian (82.1%), 13 African American (7.3 %) and 19 multi-race participants (19%). One hundred sixty six participants identified themselves as non-Hispanic (92.7%) and 13 participants identified themselves as Hispanic/Latino (7.3%). Participants' level of education ranged from 7 to 16 years with a mean of 11.68 (*SD* = 1.60). Parental education was on average 14.70 years (*SD* = 2.24) for mothers and 14.66 (*SD* = 2.66) for fathers. Participants' word reading ability was in the average range for age (*M* = 103.94, *SD* = 12.81). Eighty-one participants (45.3 %) were identified as having "no depressive symptoms," 38 participants (21.2%) were identified as having "mild/moderate depressive symptoms," and 60 participants (33.5%) were identified as having "severe depressive symptoms." Table 3 displays the demographic characteristics of the study sample stratified by group. As previously explained, participants who reported taking OCs of interest to the present study were grouped into one of three OC groups: "highly androgenic OCs" (n= 34), "androgenic OCs" (n=51), and "antantiandrogenic" OCs (n=49).

Participants who reported not taking any kind of hormonal contraceptive at the time of the study were grouped into the control group. Because there were substantially greater participants that met criteria for the control group than the OC groups, a subset of 45 participants meeting criteria for controls, were ultimately included in this study. Controls were selected for inclusion in this study using a random-number table and were matched to participants in the OC groups on age, race, ethnicity, medical history, psychiatric history, lifetime history of depression, education level, mother's education, father's education, and word reading performance (See Table 3).

2.3 Measures

2.3.1 Demographic Information

Demographic information was collected by PENN for each participant using a *Demographics Questionnaire* and a structured clinical interview (i.e., GOASSESS). The following information was obtained: age, sex, race, ethnicity, current medications, medical history, psychiatric history, and current level of depressive symptoms.

2.3.2 Standardized Reading Test

A brief (i.e., approximately 3 minutes) standardized reading test from the Wide Range Achievement Test (WRAT4, Wilkinson & Robertson, 2006) was administered by PENN to provide an estimate of IQ and determine participants' ability to complete the Computerized Neurocognitive Battery (CNB).

2.3.3 Computerized Neurocognitive Measures

The measures included in the present study were part of a 1-hour computerized neurocognitive battery (CNB) administered to participants as part of the collaborative study between CHOP and PENN. The CNB consisted of 14 tests assessing five neurobehavioral functions, among them episodic memory (e.g., CPW), complex cognition (e.g., sPVRT, PMAT, PLOT), and social cognition (ER40, MEDF) (Gur et al., 2012).

Penn Line Orientation Test (PLOT) (Gur et al., 2012)

The PLOT was included as a measure of participants' spatial orientation performance. The PLOT is a short (i.e., approximately 5 minutes) computerized measure of spatial orientation inspired by Benton's Judgment of Line Orientation test (Benton et al., 1978). The PLOT contains either 12 or 24 trials in which participants are shown a pair of lines with different orientations; one red (fixed) and the other blue (moveable). Participants are asked to rotate the blue line by clicking on one of two buttons that rotate the line by 3, 6, or 9 degrees (depending on the item) either clockwise or counterclockwise. Throughout the test, there are variations in the length of the blue line (long or short), possible degree of movement in each click, and relative location of both lines on the screen. Scores are provided for total correct trials as well as broken down by blue line length, possible degree per click, and for each of the four possible relative line positions. In a recent study by Gur and colleagues (2012) involving a subset of the participants from the PNC from which the current participants were drawn, the PLOT

demonstrated sensitivity to sex differences, with males outperforming females overall (*p < .0001, ES = - 0.23*) and in the 18 through 21 age groups (Gur et al., 2012). Overall task accuracy (% correct) and reaction time (ms) were the variables of interest for the current study.

Penn Matrix Reasoning Test (PMAT) (Gur et al., 2012)

The PMAT was developed as a measure of a nonverbal reasoning. Given the spatial components of the task, it is used in this study as a measure of visuospatial performance. The PMAT is a short (i.e., approximately 5 minutes) computerized version of the Raven's Progressive Matrices (Raven, 1960) which measures abstraction and mental flexibility. The PMAT is a multiple-choice test that requires participants to reason by geometric analogy and conceptualizing design, and numerical relationships. Questions range in difficulty from very easy to increasingly complex. Participants are presented with patterns made up of 2X2, 3X3, and 1X5 arrangements and are asked to click on the square that best fits the missing square of the pattern. Each question has 5 response choices and the measure contains either 18 or 24 questions and 3 adaptive bonus questions based on the participant's performance. The questions are presented in order of increasing difficulty and the test is discontinued after 5 incorrect responses (to any 5 items) at which point the 3 bonus questions are presented. The final score represents the total number of correct responses. In the recent PNC study by Gur and colleagues (2012) the PMAT showed sensitivity to sex differences, with males outperforming females overall *(p = .0059, ES = - 0.04*) and with males showing better accuracy in the 18

through 21 age groups (Gur et al., 2012). Overall task accuracy (% correct) and reaction time (ms) were the variables of interest for the current study.

Short Penn Verbal Reasoning Test (sPVRT) (Gur et al., 2010)

The sPVRT was used as a measure of participants' verbal reasoning. The sPVRT is a short (i.e., approximately 2.5 minutes) computerized measure of verbal reasoning by analogy. It consists of 15 multiple choice verbal analogy problems from the Educational Testing Service (ETS) factor-referenced test kit. The final score represents the total number of correct responses. The sPVRT was shortened from an earlier version (Penn Verbal Reasoning Test (PVRT)) containing 30 multiple-choice analogies (Gur, Gur, Obrist, Skolnick, & Reivich, 1987), which was found to be sensitive to sex differences (Gur et al., 2001). The sPVRT has demonstrated high internal consistency (Cronbach's α = .90) and sensitivity to sex differences (Gur et al., 2010). In the recent PNC study by Gur and colleagues (2012) the sPVRT showed sensitivity to sex differences, with males outperforming females overall ($p = .0111$, $ES = -0.02$) and with males showing better accuracy in the 18 to 19 age group (Gur et al., 2012). Overall task accuracy (% correct) was the variable of interest for the current study.

Penn Word Memory Test (CPW) (Gur et al., 1993)

The CPW was used as a measure of participants' verbal episodic memory. The CPW is a short (i.e., approximately 3 minutes) computerized measure of word memory. The CPW requires participants to memorize 20 target words, which are presented one a time at the rate of one word per second. Participants are then shown a series of words

(one a time) containing the 20 target words mixed with 20 novel foils equated for frequency, length, concreteness and imageability. Participants are then asked to decide whether they have seen the word before by clicking on 1 of 4 buttons presented at the bottom of the screen: “definitely yes”, “probably yes”, “probably no” and “definitely no”. The CPW is scored based on the number of correct and incorrect responses, divided into true/false positives/negatives. The CPW has demonstrated high internal consistency (Cronbach‘s $\alpha = .90$) (Gur et al., 2010) and sensitivity to sex differences favoring females (Gur et al., 2001; Gur et al., 2010). In the recent PNC investigation by Gur and colleagues (2012) the CPW showed sensitivity to sex differences favoring females ($p = .0268$, $ES = 0.13$), with females showing more accurate performance in the 20 to 21 age group (Gur et al., 2012). Overall task accuracy (% correct) and reaction time (ms) were the variables of interest for the current study.

Penn Emotion Recognition Task (ER40) (Gur et al., 2010)

The ER40 was used as a measure of participants’ emotion recognition. The ER40 is a short (i.e., approximately 2.5 mintues) computerized measure of facial expression/affect identification. In the ER40 participants are shown 40 color photographs of faces balanced for sex, age, and ethnicity. There are four female faces and four male faces for each emotion. The faces are presented one at a time and participants are asked to determine what emotion each face is showing from a multiple choice format: “anger,” “fear,” “happiness,” “sadness,” or “no feeling.” Scores are based on the number of total correct identifications including correct responses for male versus female faces, correct responses per emotion and false positives per emotion. The ER40 has demonstrated high

internal consistency (Cronbach's α = .92) and sensitivity to sex differences favoring females (Gur et al., 2010). In the Gur and colleagues (2012) PNC investigation, the ER40 showed sensitivity to sex differences with females outperforming males overall (*p < .0001, ES = 0.25*) however, no significant sex differences were reported in the 18 through 21 age groups (Gur et al., 2012). Overall task accuracy (% correct) and reaction time (ms) were the variables of interest for the current study.

Measured Emotion Differentiation Test (MEDF) (Gur et al., 2012)

The MEDF was used as a measure of participants' emotion intensity differentiation. The MEDF is a short computerized measure of facial expression/emotion differentiation (administration time: *M = 3.1* minutes, *SD = 0.8*). In the MEDF, participants are presented with a pair of images of the same individual expressing the same emotion ("anger," "fear," "happiness," "sadness"), one more intense than the other or of equal intensity. The participant is then asked to click on the face that displays the more intense emotion or indicate whether both images have equal intensity. The MEDF36 consists of either 36 or 60 random untimed trials, some showing no emotional difference and most showing emotional differences in increments of 10% ranging from 10% to 60%. Scores are based on the number of correct responses. In the PNC study by Gur and colleagues (2012) the ER40 showed sensitivity to sex differences with females outperforming males overall (*p < .0001, ES = 0.25*) however, no significant sex differences were reported in the 18 through 21 age groups (Gur et al., 2012). Overall task accuracy (% correct) and reaction time (ms) were the variables of interest for the current study.

CHAPTER 3: RESULTS

3.1 Demographic Analyses

Demographic analyses were conducted in order to determine if there were any pre-existing differences among the groups that would need to be controlled for in subsequent analyses. Chi Square analyses revealed no significant differences between groups on race, ethnicity, medical history, psychiatric history and level of depressive symptoms. One-way ANOVAs revealed no significant differences between the groups on age, education, or reading test performance (WRAT4).

3.1.1 Age as a Possible Covariate

Previous research using the Philadelphia Neurodevelopmental Cohort Database (PNC) from which these data were obtained suggests that age is significantly correlated with performance accuracy on the measures of interest (Gur et al., 2012). Therefore, in the present study age was carefully considered as a possible confounding factor. To investigate the relationship between age and the outcome variables, bi-variate correlations were conducted. Age was significantly correlated with participants' performance on the PLOT ($r = .21$, $p = .006$), the PMAT ($r = .27$, $p < .001$), and the SPVRT ($r = .16$, $p = .035$). Due to these significant relationships, age was further investigated as a possible covariate. First, one-way ANOVAs were performed to examine possible differences in PLOT, PMAT, and SPVRT performance across groups. Results revealed no significant differences among groups on the PLOT ($F(3,175) = 0.11$, $p = .954$, *partial* $\eta 2 = .002$), the PMAT ($F(3,175) = 1.33$, $p = .267$, *partial* $\eta 2 = .022$) or the SPVRT ($F(3,175) = 1.29$, $p = .280$, *partial* $\eta 2 = .022$).

Then, ANCOVAs were performed for each of these outcome variables to investigate the impact of OC androgenicity while controlling for age. Results of these ANCOVAs revealed that age was not a covariate as group status continued to not be a significant predictor of outcome variables.

Given that age did not significantly alter the relationship between OC status and the outcome variables, subsequent analyses were conducted using percent correct scores rather than age-based normative values. Although previous PNC studies used aged-based normative values in their analyses (Gur et al., 2010; Gur et al., 2012), percent correct scores were better suited for the present study. First, available age-based normative values were derived from both male and female participants and the present study was interested in female performance only. Furthermore, normative scores were only available for overall performance accuracy and the present study was interested in anger performance accuracy on the ER40 and the MEDF.

3.1.2 Depressive Symptoms as a Possible Covariate

Given that the rate of lifetime severe depressive symptoms in this sample was much higher (33.5%) than what would be expected in the general population of 17-18 year old adolescents (15.4%) (Merikangas et al., 2010), depressive symptoms were examined as a possible confounding factor. First, separate one-way ANOVAs were performed to examine performance on outcome variables across levels of depressive symptoms. Results revealed that participants' ability to correctly identify intensity of anger faces on the MEDF differed based on level of depressive symptoms ($F(2,176) = 3.87$, $p = .023$, *partial* $\eta 2 = .042$). However, a subsequent 3X4 factorial ANOVA did not

reveal a significant interaction between group status and level of depressive symptoms, suggesting that level of depressive symptoms was not a significant covariate ($F(6,178) = .59$, $p = .738$, *partial* $\eta 2 = .021$). Therefore, level of depressive symptoms was not included as a factor in subsequent analyses.

3.2 Hypothesis 1

The first aim of this study was to investigate the effect of OC progestin androgenicity on participants' performance on measures containing visuospatial components. Hypothesis 1 stated that the "highly androgenic" OC group and the "androgenic" OC group would perform more accurately than controls, who would perform more accurately than the "antiandrogenic" OC group on outcome measures containing visuospatial components (PLOT, PMAT). Descriptive statistics for the PLOT and the PMAT are summarized in Table 4. To examine the effects of OC androgenicity on visuospatial performance, a series of one-way ANOVA were conducted using PLOT and PMAT total percent accuracy as dependent variables. Results indicated that scores on the PLOT did not significantly differ among groups ($F(2,178) = .11$, $p = .954$, *partial* $\eta 2 = .002$). Similarly, scores on the PMAT did not significantly differ among groups ($F(2,178) = 1.33$, $p = .267$, *partial* $\eta 2 = .022$).

3.3 Hypothesis 2

The second aim of this study was to investigate the effect of OC progestin androgenicity on participants' performance on measures containing verbal components. Hypothesis 2 stated that controls would perform more accurately than all OC groups on

the CPW. Descriptive statistics for the CPW are summarized on Table 4. To examine the effects of OC androgenicity on CPW performance, a one-way ANOVA was conducted using total percent accuracy as the dependent variable. Results indicated that scores on the CPW did not significantly differ among groups ($F(2,178) = .71$, $p = .551$, *partial η2* =.012).

3.4 Hypothesis 3

Related to Aim 2, hypothesis 3 stated that the "highly androgenic" OC group and the "androgenic" OC group would perform more accurately than controls, who would perform more accurately than the "antiandrogenic" OC group on the SPVRT. Descriptive statistics for the SPVRT are summarized on Table 4. To examine the effects of OC androgenicity on SPVRT performance, a one-way ANOVA was conducted using total percent accuracy as the dependent variable. Results indicated that scores on the SPVRT did not significantly differ among groups ($F(2,178) = 1.29$, $p = .280$, *partial η2* =.022). However, although the difference among groups was not significant, the pattern of mean scores was generally consistent with the hypothesis.

3.5 Hypothesis 4

The third aim of this study was to investigate the effect of OC progestin androgenicity on participants' performance on measures containing FEP components. Hypothesis 4 stated that the "antiandrogenic" OC group would perform more accurately than controls, who would perform more accurately than the "highly androgenic" and the

"androgenic" OC groups on outcome measures containing FEP components (ER40, MEDF).

Given that the ER40 and the MEDF are similar measures of FEP, a one-way MANOVA was performed to examine the effect of OC androgenicity on composite FEP performance. Results indicated a statistically significant difference in FEP processing among groups, $F(6,348) = 2.23$, $p < .05$; *Wilk's* $\Lambda = 0.927$, *partial* $\eta 2 = .037$). To follow up, univariate one-way ANOVAs were conducted using ER40 and MEDF total percent accuracy as separate dependent variables. Results indicated that differences in ER40 performance accuracy among groups approached significance ($F(2,178) = 2.06$, $p = .107$, *partial* $\eta 2 = .034$). Similarly, scores on the MEDF approached significance ($F(2,178) = 2.46$, $p = .064$, *partial* $\eta 2 = .040$). Given the exploratory nature of this study, p values between .06 and .10 were considered borderline significant and further examined. Post hoc comparisons using Tukey's HSD indicated no significant differences between groups on ER40 performance. However on the MEDF, participants in the "highly androgenic" OC group ($M = 81.85$, $SD = 8.21$) significantly outperformed participants in the control group ($M = 77.05$, $SD = 8.25$), $p < .05$ (See Table 4).

Related to hypothesis 4, we expected that "antiandrogenic" OC group and controls would perform more accurately than the "highly androgenic" and the "androgenic" OC groups for anger faces on the ER40 and the MEDF. To examine the effects of OC androgenicity on anger performance, additional one-way ANOVAs were conducted using ER40 and MEDF anger percent accuracy as dependent variables. Results indicated that anger scores did not significantly differ among groups on the ER40

($F(2,178) = 1.35$, $p = .259$, *partial η2* =.023) or the MEDF ($F(2,178) = .83$, $p = .480$, *partial η2* =.014).

3.6 Power

A power analysis using the guidelines set forth by Cohen (1992) indicated that this study was adequately powered to detect a medium effect size of $d = .25$ for a one-way ANOVA with 4 groups, $\alpha = .05$, power = .80. The intended sample size for the study was a total 180 participants; 45 participants per group. The actual sample size was 179 participants; 45 controls, 49 "antiandrogenic" OC users, 51 "androgenic" OC users, and 34 "highly androgenic" OC users. Given the study's marginally significant findings with small to moderate effect sizes, OCs appear to have generally small effects on female neurocognitive performance. In order to detect these small effect sizes, a sample size of 1,096 would be required.

CHAPTER 4: DISCUSSION

The goal of this study was to examine the relationship between OC progestin androgenicity and female performance on measures with visuospatial (PLOT, PMAT), verbal (CPW, SPVRT) and facial expression processing (ER40, MEDF) components. Findings revealed a marginally significant trend of FEP performance differences across groups, specifically that participants using "highly androgenic" OCs were more accurate at determining facial expression intensity on the MEDF than naturally cycling controls. However, the current study failed to reveal a significant relationship between OC androgenicity and female performance on measures containing visuospatial and verbal components.

4.1 Review of Findings

4.1.1 Facial Expression Processing Performance

Based on previous findings of a female advantage on computerized measures of emotion recognition (Gur et al., 2010; Gur et al., 2012; Montagne et al., 2005) and decreased FEP accuracy in females after testosterone administration (Van Honk et al., 2007), it was expected that OC androgenicity would hinder performance on measures with FEP components. Furthermore, we expected decreased performance accuracy for anger faces with androgenic OCs ("highly androgenic", "androgenic") given previous findings of decreased accuracy at labeling anger facial expressions in males compared to females (Hampson et al., 2006; Montagne et al., 2005) and in females after a single dose of testosterone (Van Honk et al., 2007). This was the first study to examine OC androgenicity and female FEP performance both overall and specifically for anger faces.

Although there was a significant among-groups difference in performance accuracy on composite ER40 and MEDF performance, analyses of the individual measures revealed marginally significant findings. Our hypothesis of hindered FEP performance with OC androgenicity was unsupported. On the contrary, findings revealed that "highly androgenic" OC users performed more accurately than naturally cycling controls on the MEDF.

Although the unexpected finding that "highly androgenic" OC users outperformed controls may be spurious, we entertain the notion that it may be consistent with past findings of increased FEP performance with lower sex hormone levels. Contrary to established parameters of OC androgenicity, some researchers argue that OCs lower all natural sex hormones and therefore have no clinically meaningful androgenic effects on cognition (Stanczyk, 2003; Thorneycroft et al., 1999). This argument is supported by a recent meta-analysis showing that OCs significantly suppress natural testosterone levels regardless of progestin generation (Zimmerman, Eijkemans, Coelingh Bennink, Blankenstein & Fauser, 2014).

Following this argument, we could speculate that the OC users in our sample had lower sex hormone levels than naturally cycling controls. As such, our study's finding and pattern of mean scores on the MEDF are consistent with previous research showing increased FEP performance with lower estrogen and progesterone levels (i.e., follicular phase) (Guapo et al., 2009; Derntl et al., 2008a; Derntl et al., 2008b; Derntl et al., 2013).

Our finding however, cannot fully support a position that OCs in general facilitate FEP performance by lowering natural sex hormones as only "highly androgenic" OC users outperformed naturally cycling controls on the MEDF. Progestin androgenicity

does not fully account for these findings either otherwise both androgenic OC groups ("highly androgenic" and "androgenic") would have outperformed controls on the MEDF. Given that the major ingredient difference among the OCs used in the present study was progestin type (all OCs had comparable doses of ethinyl estradiol) suggests further exploration of another progestin effect - progestogenicity. Whereas the scope of the present study was progestin androgenicity, it should be noted that the OC progestins investigated (*Levonorgestrel, Norethindrone, Drospirenone*) have varying levels of progestogenicity. A progestin's progestogenicity is measured by its relative binding affinity (RBA) to progesterone receptors in the body and central nervous system. Compared to natural progesterone, *Levornogestrel* has a progestogenic RBA of 150%, *Norethindrone* 75% and *Drospirenone* 35% (Schindler, et al., 2008). Because of differences in progestogenicity, *Levornogestrel* may be used in much smaller dosages (i.e., 0.10 – 0.15 mg) to obtain a progestogenic effect equivalent to that of *Norethindrone* (i.e., 1 mg) and *Drospirenone* (i.e., 3 mg) (Greet et al., 2005). Nonetheless, the OC progestins included in the present study are derived from different chemical compounds, which may differentially influence FEP performance accuracy. Given the complex biological effects of OC progestins (i.e., androgenicity, progestogenicity) and the lack of research into OC progestogenicity and female cognition, we cannot interpret participants' MEDF performance beyond stating that progestogenicity may have played a role.

As previously discussed, performance differences among groups was only observed on the measure of facial emotion intensity (MEDF) but not on the measure of emotion recognition (ER40). Furthermore, the pattern of mean scores on the ER40 did not reveal a similar relationship (i.e., naturally cycling controls were more accurate than

"antiandrogenic" OC users). One way to account for these findings is to conjecture that the MEDF may be more sensitive than the ER40 at detecting differences in performance accuracy related to sex hormone fluctuations.

The ER40 and the MEDF were designed to measure emotion recognition and emotion intensity differentiation respectively. The primary difference between these measures lies in the emotional stimuli presented. The ER40 contains images of full-blown emotions (100% intensity) whereas the MEDF contains images with varying levels of emotion intensity (i.e., 0% - 60% difference). Interestingly, our findings of performance differences on the MEDF but not the ER40 parallel previous findings of sex differences in emotion recognition. A series of FEP studies by Hoffman et al. (2010) reported that females were more accurate than males at recognizing subtle facial emotions but both sexes were equally accurate when full-blown emotions were presented (i.e., 100% intensity). Hoffman and colleagues argued that the female FEP advantage may be mediated by their ability to recognize facial emotions under conditions of subtle emotional information (Hoffman et al., 2010). This, they argued, explains why meta-analyses of FEP studies showing a female advantage on average report considerably small effect sizes and why some studies fail to detect sex differences in FEP performance altogether (Hoffman et al., 2010). Given that the present study identified between-group differences on the MEDF but not on the ER40 and with a slightly larger small-moderate effect size on the MEDF than the ER40 suggests that sex hormone alterations (i.e., via OC use) may affect a woman's ability to interpret subtle facial expressions but may not hinder overall emotion recognition accuracy.

Lastly, our hypothesis of hindered facial anger accuracy with OC androgenicity was not supported nor were the pattern of mean scores on both measures consistent with the study's hypothesis. There were no significant differences in emotion recognition of facial anger (ER40) or emotion intensity differentiation of facial anger (MEDF). Interestingly, the pattern of mean scores on each measure was contradictory as naturally cycling controls recognized anger faces more accurately than OC users on the ER40 and less accurately than OC users on the MEDF. As previously discussed, the MEDF may be more sensitive than the ER40 at picking up performance differences related to hormone fluctuations. Although these differences may be spurious, considering MEDF anger performance patterns more representative of FEP changes with OC use would lend additional support to arguments against established OC androgenicity parameters. Increased anger sensitivity with OC use would support theories that OC users have lower natural testosterone levels than naturally cycling females.

Multiple explanations may account for our lack of significant findings. First, some facial expressions (i.e., disgust) may be affected by sex hormone changes more than others (i.e., anger). Whereas two FEP studies to date have reported a female advantage for subtle disgust facial expressions (Hoffman et al., 2010; Montagne et al., 2005), only one of those studies reported a female advantage for subtle anger facial expressions (Montagne et al., 2005). The second study to replicate significant performance differences for anger expressions among females achieved this after administering a dose of sublingual testosterone shown to increase female testosterone levels to male-like levels (Van Honk et al., 2007). These findings suggest that in females, anger FEP performance

may be sensitive to much higher testosterone thresholds (i.e., male-like levels) than any afforded by the androgenic effects of OC progestins.

4.1.2 Visuospatial Performance

Based on the hypothesis that OC androgenicity would facilitate visuospatial performance, it was expected that users of "highly androgenic" and "androgenic" OCs would perform more accurately on measures with visuospatial components. Furthermore, we expected to see decreased performance accuracy on users of "antiandrogenic" OCs based on previous research showing hindered spatial rotation performance with antiandrogenic OCs (Griksiene & Ruksensas, 2011, Wharton et al., 2008).

Our hypotheses were largely unsupported. There were no differences in visuospatial performance between OC users and non-users. Although female spatial orientation has been examined within the context of OC use (Goyette et al., 2010), this was the first study to control for OC progestin androgenicity. Consistent with Goyette and colleagues' findings (2011), this study did not observe differences in spatial orientation performance between naturally cycling females and OC users.

Similarly, despite the geometric components of the PMAT, no relationship between OC androgenicity and performance accuracy was observed. This is the first study to date, to examine female matrix reasoning performance and OC androgenicity.

Given that most studies reporting significant visuospatial performance differences in OC users employ a mental rotation task (MRT) suggests that some aspects of visuospatial processing (i.e., spatial rotation) may be more sensitive to hormonal fluctuations than others (i.e., spatial orientation, reasoning by geometric analogy).

4.1.3 Verbal Performance

Based on previous findings of females outperforming males on verbal episodic memory tasks (Gur et al., 2001; Temple & Cornish, 1993; Zelinski et al., 1993), it was expected that non-OC users would outperform OC users on the CPW.

Despite a lack of significance, findings were consistent with previous research reporting no differences in the verbal episodic memory of OC users and naturally cycling controls (Wharton et al., 2008).

Based on previous findings of males outperforming females on measures of verbal reasoning by analogy (Lim, 1994; Gur et al., 2012), it was expected that OC androgenicity would facilitate verbal reasoning performance. It was hypothesized that users of "highly androgenic" and "androgenic" OCs would exhibit more accurate verbal reasoning performance than controls who were expected to perform more accurately than users of "antiandrogenic" OCs. No study to date had examined female analogical reasoning and androgenicity.

Our hypothesis was unsupported however; the pattern of mean scores was generally consistent with the hypothesis of a negative relationship between OC androgenicity and female verbal analogical reasoning. An examination of effect sizes revealed a small effect size, suggesting that the effects of OC androgenicity on analogical reasonins are not clinically significant.

Given that most studies reporting verbal performance differences in females with sex hormone changes use a verbal fluency measure suggests that some aspects of verbal processing (i.e., verbal fluency) may be more sensitive to hormonal fluctuations than others (verbal episodic memory, analogical reasoning).

4.2 Limitations

Several of the study's limitations stem from using a pre-existing data set. First, using a database limited the study to a quasi-experimental design, as participants could not be randomly assigned to their respective groups. Therefore, although the groups were relatively equal in demographic information, the investigator could not account for any pre-existing differences that may have led participants to choose an "androgenic" OC versus an "antiandrogenic" one. Oral contraceptives were primarily designed for pregnancy prevention however they may be prescribed for noncontraceptive reasons. Common "off-label" indications of OCs include the treatment of menstrual disorders (e.g., dysmenorrhea, amenorrhea, irregular cycles, excessive bleeding), hyperandrogenic disorders (e.g., acne, hirsutism), and gynecological conditions like endometriosis and polycystic ovary syndrome (Dayal & Barnhart, 2001). Additionally, several consumer-related factors may affect the OC formulations or brand females use. For example, one female may choose "antiandrogenic" 24/4 formulation OC *Yaz* over a typical 21/7 formulation OC because *Yaz* will result in shorter menstrual periods whereas another female may choose *Yaz* because studies show that *Drospirenone*-containing OCs result in less weight gain and water retention (Bonnema & Spencer, 2011). Similarly, most of the females that choose extended cycle (3-month) "highly androgenic" OCs like *Seasonale* do so because of the convenience of having fewer menstrual periods however, a large number of females avoid extended cycle OCs because of the common misconception that bleeding once a month is "necessary" and "normal" (Andrist et al., 2004). Cost can also be a determining factor as some OC brands may require higher copays and some formulations (i.e. 3-month packs) may require users to pay 3-months'

worth of copayments up-front. Lastly, other variables like media advertisements and the availability of OC samples at clinics and gynecologists' offices (which may or may not be driven by pharmaceutical companies) should also be explored.

The present study was also limited in that there were no available data from which to directly measure or estimate participants' hormonal status. No information was obtained regarding OC use therefore OC group assignment was solely based on alleged progestin androgenicitiy. Because the biological androgenicity of OC progestins was not verified via hormonal assays, the investigator could not ascertain whether in fact a participant was truly in the "highly androgenic" versus the "androgenic" hormone range. Similarly, the study did not obtain information to estimate the cycle phase controls, potentially mixing together women with very different hormonal profiles (e.g., low hormone women in the follicular phase and high hormone women during the luteal phase).

Additional limitations included that the OC group showing a nearly significant trend on the MEDF ("highly androgenic") contained the smallest number of participants and that the study's sample may have not have been representative as participants reported higher rates of lifetime depressive symptoms than those found in the general population.

4.3 Directions for Future Research

4.3.1 Study Design

Future investigations of the relationship between OC androgenicity and female neurocognitive performance may benefit from following a true experimental design in which neurocognitive performance is assessed before and after random assignment to an

OC group. Future studies should examine participants' estrogen, progesterone and testosterone levels via hormonal assays before grouping OC users into androgenicity categories. Additionally, historical information regarding OC use (e.g., why OCs were prescribed, why the specific OC brand was chosen, length of use, etc) should be obtained in an effort to uncover any pre-existing differences among different OC users that may contribute to neurocognitive performance differences.

4.3.2 Further Exploration of Depression and Oral Contraceptives

Although depression was beyond the scope of the present study and was not found to be a significant factor in participants' neurocognitive performance, it remains a variable of interest given the higher prevalence of lifetime depressive symptoms in OC users compared to controls. Of specific concern was the possibility that other variables related to depression may have played a role in participants' performance.

Early studies of OCs and depression concluded that depression was a side effect of OC use (Herzberg, Johnson, & Brown, 1970) however, recent population-based studies report either no relationship between OCs and depression (Joffe, Cohen, & Harlow, 2003; Duke, Sibbritt, & Young, 2007) or that OCs reduce level of depressive symptoms (Toffol, Heikinheimo, Koponen, Luoto, & Partonen, 2011; Keyes et al., 2013). The true link between OC use and depression may lie in their relationship to premenstrual dysphoric disorder (PMDD) - a severe form of premenstrual syndrome (PMS) characterized by emotional symptoms such as irritability, mood swings and depressed mood surfacing during the luteal phase of the menstrual cycle and disappearing shortly after menstruation (Soares, Cohen, Otto, & Harlow, 2001). Premenstrual dysphoric

disorder (PMDD) and depression share a significant comorbidity (Soares et al., 2001), to the extent that PMDD is regarded as a variant of depression and is categorized in the DSM-IV as a "depressive disorder not otherwise specified" (Landen & Eriksson, 2003). While the role of sex hormones in the etiology of PMDD and depression remains unclear, researchers theorize that the sex hormone estrogen may play a role. Specifically, that estrogen may mediate serotonergic functions and therefore precipitous changes in estrogen levels may trigger depressive symptoms (Joffe & Cohen, 1998; Keyes et al., 2013). Because of their stabilizing effects on sex hormones, OCs have been widely prescribed as an off-label treatment of PMDD (Freeman et al., 2001). Newer OCs containing the progestin *Drospirenone* (e.g., *Yaz, Yasmin*) have even been recognized by the FDA as effective treatments of PMDD and have been marketed accordingly (Schindler, 2013). This may in part explain why in the present study, users of "antiandrogenic" *Drospirenone*-based OCs reported the highest levels of severe lifetime depressive symptoms.

As previously discussed, future studies should gather historical data on why OCs (or particular types of OCs) were prescribed as off-label uses like the treatment of PMDD may elucidate pre-existing differences among groups which may contribute to performance differences. Additionally, this information would allow researchers to create subgroups of OC users based on disorder-driven categories (e.g., PMDD treatment), which may help to further clarify the relationship between sex hormones and female cognition. For example, females who were prescribed OCs for PMDD may be more cognitively sensitive to sex hormone changes than females who where prescribed OCs solely for contraceptive purposes.

4.3.3. Further Exploration of Facial Expression Processing Measures

The present study explored female performance on two measures containing FEP components – emotion recognition (ER40) and emotion intensity differentiation (MEDF). Although findings suggest the MEDF may be more sensitive to hormonal fluctuations than the ER40, these measures should be jointly explored as their design may allow for further clarification of the potentially moderating role of emotion intensity in facial expression recognition. Additionally, ideally disgust facial expressions would be included as previous research has shown differences in FEP performance for disgust expressions with hormonal fluctuations. Lastly, FEP performance for low to mid intensity emotions on the MEDF should be examined to determine if there are any between-group differences that may have been missed when the total intensities for each emotion was assessed.

4.4 Clinical Implications

4.4.1 The Role of Emotion Intensity in Facial Expression Processing

Various studies report differences in female FEP performance with natural and pharamacologically induced hormonal changes. To date the only study of female FEP performance and OC focused on imaging and did not provide behavioral data of participants' performance (Mareckova et al., 2012). Additionally, the present study was the first to research OC androgenicity and FEP performance. The study's nearly significant trend of performance differences for facial expressions with varying emotion intensities (MEDF) but not for full-blown facial emotions (ER40) contributes evidence that some aspects of FEP (i.e., emotion intensity differentiation) that may be more

sensitive to hormonal fluctuations than others (i.e., emotion recognition). The study's findings also contribute to a developing theory in the FEP literature that the female FEP advantage lies in the processing of subtle facial emotional cues.

4.4.2 Contributions to the Verbal and Visuospatial Literature

The majority of studies examining female cognitive performance and sex hormone changes focus on verbal and spatial measures of verbal fluency and spatial orientation respectively. The present study contributes to the existing literature on female verbal cognitive performance by replicating previous null findings in verbal episodic memory across OC users and controls. Additionally, it expands the literature by being the first study to examine OC use and OC androgenicity and female analogical reasoning. Similarly, the present study contributes to the body of research into female visuospatial performance by replicating previous null findings in spatial orientation performance across OC users and controls and by being the first study to examine OC androgenicity and spatial orientation. Lastly, the present study expands visuospatial processing research by examining OC use and OC androgenicity and female matrix reasoning performance.

4.4.3 Cognitive Performance and Progestin Androgenicity Parameters

Overall, findings of performance patterns on the MEDF similar to those seen in low hormone females contribute behavioral data that supports arguments against established OC androgenicity parameters. Additionally, patterns of mean scores inconsistent with either OC androgenicity or hypoandrogenicity arguments, highlight the

need to explore the progestational components of oral contraceptives to gain a more in-depth understanding of the synergistic effects of natural and synthetic sex hormones on female cognition. Lastly, the study's findings of mostly small effect sizes suggest that the effects of all OC's on cognition may be clinically negligible.

List of References

Abad, F. J., Colom, R., Rebollo, I., & Escorial, S. (2004). Sex differential item functioning in the Raven's Advanced Progressive Matrices: Evidence for bias. *Personality and Individual Differences, 36,* 1459-1470. doi: 10.1016/S0191-8869(03)00241-1

Aleman, A., Bronk, E., Kessels, R., Koppeschaar, H., & Van Honk, J. (2004). A single administration of testosterone improves visuospatial ability in young women. *Psychoneuroendocrinology, 29,* 612–617. doi: 10.1016/S0306-4530(03)00089-1

Amy, J. J., & Tripathi, V. (2009). Contraception for women: An evidence based overview. *British Medical Journal, 339,* b2895.

Anderson, D. C. (1974). Sex hormone binding globulin. *Clinical Endocrinology, 3,* 69-96. doi: 10.1111/j.1365-2265.1974.tb03298.x

Andrist, L. C., Arias, R. D., Nucatola, D., Kaunitz, A. M., Musselman, B. L., Reiter, S., … Emmert, S. (2004). Women's and provider's attitudes toward menstrual suppression with extended cycle oral contraceptives. Contraception, 70(5), 359-363. doi: 10.1016/j.contraception.2004.06.008

Ansbacher, R. (1991). Interchangeability of low-dose oral contraceptives: Are current equivalent testing measures adequate to ensure therapeutic equivalency? *Contraception, 43(2),* 139-147. doi: 10.1016/0010-7824(91)90041-D

Bachmann, G. A. (2002). The hypoandrogenic woman: pathophysiologic overview. *Fertility and Sterility, 77 (4),* S72-S76. doi: 10.1016/S0015-0282(02)03003-0

Baird, D. T., Horton, R., Longcope, C., & Tait, J. F. (1969). Steroid dynamics under steady-state conditions. *Recent Progress in Hormone Research, 25,* 611–664.

Batur, P., Elder, J., & Mayer, M. (2003). Update on contraception: Benefits and risks of the new formulations. *Cleveland Clinic Journal of Medicine, 70,* 681–696. doi: 10.3949/ccjm.70.8.681

Benton, A. L., Varney, N. R., & Hamsher, K. D. (1978). Visuospatial judgment: A clinical test. *Archives of Neurology, 35,* 364–367

Bonnema, R., & Spencer, A. L. (2011). Health issues in oral contraception: risks, side effects and health benefits. *Expert Review of Obstetrics & Gynecology, 6(5),* 551-557. doi: 10.1586/eog.11.49

Carr, B. R. (1998). Uniqueness of oral contraceptive progestins. *Contraception, 58,* 23S-27S.

Collaer, M. L., & Nelson, J. D. (2002). Large visuospatial sex difference in line judgment: Possible role of attentional factors. *Brain & Cognition, 49,* 1-12. doi: 10.1006/brcg.2001.1321

Collaer, M. L., Reimers, S., & Manning, J. T. (2007). Visuospatial performance on an internet line judgment task and potential hormonal markers: Sex, sexual orientation, and 2D: 4D. *Archives of Sexual Behavior, 36,* 177-192. doi: 10.1007/s10508-006-9152-1

Cumniing, D. C., & Wall, S. R. (1985). Non-sex hormone-binding globulin-bound testosterone as a marker for hyperandrogenism. *Journal of Clinical Endocrinology and Metabolism, 61,* 873-876. doi: 10.1210/jcem-61-5-873

Darney, P. (1995). The androgenicity of progestins. *American Journal of Medicine, 98(1),* S104-S110. doi: 10.1016/S0002-9343(99)80067-9

Dayal, M., & Barnhart, K. T. (2001). Noncontraceptive benefits and therapeutic uses of the oral contraceptive pill. *Seminars in Reproductive Medicine, 19(4),* 295-304. doi: 10.1055/s-2001-18637

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,* 97-104. doi: 10.1016/j.yhbeh.2012.10.009

Derntl, B., Kryspin-Exner, I., Fernbach, E., Moser, E., & Habel, U. (2008a). Emotion recognition accuracy in healthy young females is associated with cycle phase. *Hormones and Behavior, 53,* 90-95. doi: 10.1016/j.yhbeh.2007.09.006

Derntl, B., Windischberger, C., Robinson, S., Lamplmayr, E., Kryspin-Exner, I., Gur, R., … Habel, U. (2008b). Facial emotion recognition and amygdala activation are associated with menstrual cycle phase. *Psychoneuroendocrinology, 33,* 1031-1040. doi: 10.1016/j.psyneuen.2008.04.014

Duke, J. M., Sibbritt, D. W., & Young, A. F. (2007). Is there an association between the use of oral contraceptives and depressive symptoms in young Australian women? *Contraception, 75(1),* 27-31. doi: 10.1016/j.contraception.2006.08.002

Ekman, P., & Friesen, W.V. (1976). *Pictures of Facial Affect.* Palo Alto: Consulting Psychologists Press.

Erlanger, D. M., Kutner, K. C., & Jacobs, A. R. (1999). Hormones and cognition: Current concepts and issues in neuropsychology. *Neuropsychology Review, 9,* 175-207. doi: 10.1023/A:1021634622577

Faul, F. (2009). GPower (Version 3.1.2). [Computer software]. Universitat Kiel.

Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist, 43 (2),* 95-103. doi: 10.1037/0003-066X.43.2.95

Fitzgerald, C., Feichtinger, W., Spona, J., Elstein, M., Ludicke, F., Muller, U., & Williams, C. (1994). A comparison of the effects of two monophasic low dose oral contraceptives on the inhibition of ovulation. *Advances in Contraception, 10,* 5-18. doi: 10.1007/BF01986524

Freeman, E. W., Kroll, R., Rapkin, A., Pearlstein, T., Brown, C., Parsey, K., … Foegh, M. (2001). Evaluation of a unique oral contraceptive in the treatment of premenstrual dysphoric disorder. *Journal of Women's Health and Gender-Based Medicine, 10(6),* 561-569. doi: 10.1089/15246090152543148

Glamser, F. D., & Turner, R. W. (1995). Youth sport participation and associated sex differences on a measure of spatial ability. *Perceptual & Motor Skills, 81,* 1099-1105. doi: 10.2466/pms.1995.81.3f.1099

Golbs, S., Nikolov, R., & Zimmermann, T. (2001). Pharmacology of Nortestosterone Derivatives. *Menopause Review, 6,* 31–37.

Gordon, H. W. & Lee, P. A. (1993). No difference in cognitive performance between phases on the menstrual cycle. *Psychoneuroendrocrinology, 18,* 521-531. doi: 10.1016/0306- 4530(93)90045-M

Gouchie, C., & Kimura, D. (1991). The relationship between testosterone levels and cognitive ability patterns. *Psychoneuroendocrinology, 16(4),* 323-334. doi: 10.1016/0306- 4530(91)90018-O

Gourovitch, M. L., Goldberg, T. E., & Weinberger, D. R. (1996). Verbal fluency deficits in patients with schizophrenia: Semantic fluency is differentially impaired as compared with phonologic fluency. *Neuropsychology, 10 (4),* 573-577. doi: 10.1037/0894-4105.10.4.573

Goyette, S. R., McCoy, J. G., Kennedy, A., & Sullivan, M. (2012). Sex differences on the judgement of line orientation task: A function of landmark presence and hormonal status. *Physiology and Behavior, 105,* 1045-1051. doi: 10.1016/j.physbeh.2011.11.018

Greer, J. B., Modugno, F., Allen, G. O., & Ness, R. B. (2005). Androgenic progestins in oral contraceptives and the risk of epithelial ovarian cancer. *American College of Obstetricians & Gynecologists, 105 (4),* 731- 740. doi: 10.1097/01.AOG.0000154152.12088.48

Griksiene, R. & Ruksenas, O. (2011). Effects of hormonal contraceptives on mental rotation and verbal fluency. *Psychoneuroendocrinology, 36,* 1239-1248. doi: 10.1016/j.psyneuen.2011.03.001

Guapo, V. G., Graeff, F. G., Zani, A. C. T., Labate, C. M., Dos Reis, R. M., & Del-Ben C. M. (2009). Effects of sex hormonal levels and phases of the menstrual cycle in the processing of emotional faces. *Psychoneuroendocrinology, 34,* 1087- 1094. doi: 10.1016/j.psyneuen.2009.02.007

Guay, A. T. (2002). Screening for androgen deficiency in women: methodological and interpretive issues. *Fertility and Sterility, 77(Suppl. 4),* 83–88. doi: 10.1016/S0015-0282(02)02965-5

Gur, R. C., Alsop, D. Glahn, D., Petty, R., Swanson, C. L., Maldjian, J. A., … Gur, R. E. (2000). An fMRI study of sex differences in regional activation to a verbal and spatial task. *Brain and Language, 74(2),* 157-170. doi: 10.1006/brln.2000.2325

Gur, R. C., Gur, R. E., Obrist, W. D., Skolnick, B. E., & Reivich, M. (1987) Age and regional cerebral blood flow at rest and during cognitive activity. *Archives of General Psychiatry, 44,* 617–621. doi: 10.1001/archpsyc.1987.01800190037006

Gur, R. C., Jaggi, J. L., Ragland, J. D., Resnick, S. M., Shtasel, D., Muenz, L., & Gur, R. E. (1993). Effects of memory processing on regional brain activation: Cerebral blood flow in normal subjects. *International Journal of Neuroscience, 72,* 31–44. doi: 10.3109/00207459308991621

Gur, R. C., Ragland, D., Moberg, P. J., Turner, T. H., Bilker, W. B., Kohler, C., … Gur, R. E. (2001). Computerized neurocognitive scanning: I. methodology and validation in healthy people. *Neuropsychopharmacology, 25,* 766-776. doi: 10.1016/S0893-133X(01)00278-0

Gur, R. C., Richard, J., Calkins, M. E., Hansen, J. A., Bilker, W. B., Loughead, J., … Gur, R. E. (2012). Age group and sex differences in performance on a computerized neurocognitive battery in children age 8-21. *Neuropsychology, 26(2),* 251-266. doi: 10.1037/a0026712

Gur, R. C., Richard, J., Hughett, P., Calkins, M. E., Macy, L., Bilker, W. B., . . . Gur, R. E. (2010). A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: Standardization and initial construct validation. *Journal of Neuroscience Methods, 187,* 254–262. doi: 10.1016/j.jneumeth.2009.11.017

Gur, R. C., Sara, R., Hagendoorn, M., Marom, O., Hughett, P., Macy, L., … Gur, R. E. (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neu-rocognitive studies. *Journal of Neuroscience Methods, 115(2),* 137–143. doi: 10.1016/S0165-0270(02)00006-7

Halari, R., Sharma, T., Hines, M., Andrew, C., Simmons, A., & Kumari, V. (2006). Comparable fMRI activity with differential behavioural performance on mental rotation and overt verbal fluency tasks in healthy men and women. *Experimental Brain Research, 169,* 1-14. doi: 10.1007/s00221-005-0118-7

Hall, J. K., Hutton, S. B, & Morgan, M. J. (2010). Sex differences in scanning faces: Does attention to the eyes explain female superiority in facial expression recognition? *Cognition and Emotion, 24 (4),* 629-637. doi: 10.1080/02699930902906882

Hampson, E. (1990). Estrogen-related variations in human spatial and articulatory-motor skills. *Psychoneuroendocrinology, 15,* 97-111. doi: 10.1016/0306-4530(90)90018-5

Hampson, E., & Young, E. A. (2007). Methodological issues in the study of hormone-behavior relations in humans: understanding and monitoring the menstrual cycle. In J. B. Becker, K. J. Berkley, N. Geary, E. Hampson, J.P. Herman, & E. Young (Eds.) *Sex Differences in the Brain: From Genes to Behavior* (pp. 63-78). New York, NY: Oxford University Press, Inc.

Hausmann, M., Slabbekoorn, D., Van Goozen, S. H., Cohen-Kettenis, P. T., & Gunturkun, O. (2000). Sex hormones affect spatial abilities during the menstrual cycle. *Behavioral Neuroscience, 114*(6), 1245–1250. doi: 10.1016/S0028-3932(00)00045-2

Hausmann, M., Schoofs, D., Rosenthal, H. E., & Jordan, K. (2009). Interactive effects of sex hormones and gender stereotypes on cognitive sex differences – A psychobiosocial approach. *Psychoneuroendocrinology, 34,* 389-401. doi: 10.1016/j.psyneuen.2008.09.019

Herlitz, A., Nilsson, L. G., & Bäckman, L. (1997). Gender differences in episodic memory. *Memory and Cognition, 25,* 801–811. doi: 10.3758/BF03211324

Herlitz, A., & Rehnman, J. (2008). Sex differences in episodic memory. *Current Directions in Psychological Science, 17(1),* 52-55. doi: 10.1111/j.1467-8721.2008.00547.x

Herzberg, B. N., Johnson, A. L., & Brown, S. (1970). Depressive symptoms and oral contraceptives. *British Medical Journal, 4,* 142-145.

Hirshman, E., Merritt, P., Wang, C. C. L., Wierman, M., Budescu, D. V., Kohrt, W., … Templin, J. L. (2004). Evidence that androgenic and estrogenic metabolites contribute to the effects of dehydroepiandrosterone on cognition in postmenopausal women. *Hormones & Behavior, 45,* 144–155. doi: 10.1016/j.yhbeh.2003.09.008

Hoffman, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, H. C. (2010). Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. *Acta Psychologica, 135,* 278-283. doi: 10.1016/j.actpsy.2010.07.012

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107,* 139-155. doi: 10.1037/0033-2909.107.2.139

Irwing, P. & Lynn, R. (2005). Sex differences in means and variability on the progressive matrices in university students: A meta-analysis. *British Journal of Psychology, 96,* 505-524. doi: 10.1348/000712605X53542

Janaud, A., Rouffy, J., Upmalis, D., & Dain, M. (1992). A comparison study of lipid and androgen metabolism with triphasic oral contraceptive formulations containing norgestimate and levonorgestrel. *Acta Obstetrica et Gynecologica Scandinavica, 71 (S156),* 33-38.

Joffe, H., & Cohen, L. S. (1998). Estrogen, serotonin and mood disturbance: Where is the therapeutic bridge. *Biological Psychiatry, 44(9),* 798-811. doi: 10.1016/S0006-3223(98)00169-3

Joffe, H., Cohen, L. S., & Harlow, B. L. (2003). Impact of oral contraceptive pill use on premenstrual mood: predictors of improvement and deterioration. *American Journal of Obstetrics and Gynecology, 189(6),* 1523-1530. doi: 10.1016/S0002-9378/(03)00927-X

Kelso, W. M., Nicholls, M. E. R., & Warne, G. L. (1999). Effects of prenatal androgen exposure on cerebral lateralization in patients with congenital adrenal hyperplasia (CAH). *Brain and Cognition, 40(1),* 153–156.

Keyes, K. M., Cheslack-Postava, K., Westhoff, C., Heim, C. M., Haloossim, M., Walsh, K., & Koenen, K. (2013). Association of hormonal contraceptive use with reduced levels of depressive symptoms: A national study of sexually active women in the United States. *American Journal of Epidemiology, 178(9),* 1378-1388. doi: 10.1093/aje/kwt188

Kiley, J., & Hammond, C. (2007). Combined oral contraceptives: A comprehensive Review. *Clinical Obstetrics and Gynecology, 50*(4), 868-877. doi: 10.1097/GRF.0b013e318159c06a

Kimura, D. (1996). Sex, sexual orientation and sex hormones influence human cognitive function. *Current Opinion in Neurobiology, 6(2),* 259-263. doi: 10.1016/S0959-4388(96)80081-X

Knopp, R. H., Broyles, F. E., Cheung, M., Moore, K., Marcovina, S., & Chandler, W. L. (2001). Comparison of the lipoprotein, carbohydrate and hemostatic effects of phasic oral contraceptives containing desogestrel or levonorgestrel. *Contraception, 63,* 1-11. doi: 10.1016/S0010-7824(00)00196-7

Krattenmacher, R. (2000). Drospirenone: pharmacology and pharmacology and pharmacokinetics of a unique progestogen. *Contraception, 62,* 29-38. doi: 10.1016/S0010-7824(00)00133-5

Kuhl, H., Gahn, G., Romberg, G., Marz, W., & Taubert, H. D. (1985). A randomized cross-over comparison of two low-dose oral contraceptives upon hormonal and metabolic parameters: Effects upon sexual hormone levels. *Contraception, 31 (6),* 583-593. doi: 10.1016/0010-7824(85)90058-7

Kumar, N., Koide, S. S., Tsong, Y. Y., & Sundaram, K. (2000). Nestorone: A progestin with a unique pharmacological profile. *Steroids, 65,* 629–636. doi: 10.1016/S0039-128X(00)00119-7

Landen, M., & Eriksson, E. (2003). How does premenstrual dysphoric disorder relate to depression and anxiety disorders? *Depression and Anxiety, 17,* 122-129. doi: 10.1002/da10089

Lawton, C. A., & Morrin, K. A. (1999). Gender differences in pointing accuracy in computer-simulated 3D mazes. *Sex Roles, 40,* 73-92. doi: 10.1023/A:1018830401088

Lim, T. K. (1994). Gender-related differences in intelligence: Application of confirmatory factor analysis. *Intelligence, 19,* 179-194. doi: 10.1016/0160-2896(94)90012-4

Lindgren, S. D., & Benton, A. L. (1980). Developmental patterns of visuospatial judgment. *Journal of Pediatric Psychology,* 5, 217–225. doi: 10.1093/jpepsy/5.2.217

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterisation of gender differences in spatial abilities: A meta-analysis. *Child Development, 56,* 1479-1498.

Lobl, T. J. (1981). Androgen transport proteins: physical properties, hormonal regulation, and possible mechanism of T, CBG, and ABP action. *Archives of Andrology, 7,* 133-151.

Lynn, R., & Irwing, P. (2004). Sex differences on the progressive matrices: A meta-analysis. *Intelligence, 32(5),* 481-498. doi: 10.1016/j.intell.2004.06.008

Makghzangy, M. N., Wynn, V., & Lawrence, D. M. (1979). Sex hormone binding globulin capacity as an index of oestrogenicity or androgenicity in women on oral contraceptive steroids. *Clinical Endocrinology, 10,* 39-45. doi:10.1111/j.1365-2265.1979.tb03031.x

Maki, P. M., Rich, J. B., & Rosenbaum, R. S. (2002). Implicit memory varies across the menstrual cycle: estrogen effects in young women. *Neuropsychologia, 40,* 518-529.

Mareckova, K., Perrin, J. S., Khan, I. N., Lawrence, C., Dickie, E., McQuiggan, D. A., & Paus, T. (2012). Hormonal contraceptives, menstrual cycle and brain responses to faces. *Social Cognitive and Affective Neuroscience Advanced Access.* doi: 10.1093/scan/nss128

Massafra, C., De Felice, C., Agnusdei, D. P., Gioia, D., & Bagnoli, F. (1999). Androgens and osteocalcin during the menstrual cycle. *Journal of Clinical Endocrinology and Metabolism, 84(3),* 971–974. doi: 10.1210/jc.84.3.971

Masters, M. S., & Sanders, B. (1993). Is the gender difference in mental rotation disappearing? *Behavior Genetics, 23,* 337–341. doi: 10.1007/BF01067434

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin, 126 (3),* 424-453. doi: 10.1037/0033-2909.126.3.424

McFadden, D. (2000). Masculinizing effects on otoacoustic emissions and auditory evoked potentials in women using oral contraceptives. *Hearing Research, 142(1–2),* 23-33. doi: 10.1016/S0378-5955(00)00002-2

Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., . . . Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication – Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry, 49(10),* 980-989. doi: 10.1016/j.jaac.2010.05.017

Miyahira, A., Morita, K., Yamaguchi, H., Nonaka, K., & Maeda, H. (2000). Gender differences of exploratory eye movements: a life span study. *Life Sciences, 68(5),* 569- 577. doi: 10.1016/S0024-3205(00)00963-2

Moffat, S. D., & Hampson, E. (1996). A curvilinear relationship between testosterone and spatial cognition in humans: Possible influence of hand preference. *Psychoneuroendocrinology, 21,*323–337. doi: 10.1016/0306-4530(95)00051-8

Montagne, B., Kessels, E. F., De Haan, E. H. F., & Perrett, D. I. (2005). Sex differences in the perception of affective facial expressions: Do men really lack emotional sensitivity? *Cognitive Processing, 6,* 136-141. doi: 10.1007/s10339-005-0050-6

Mordecai, K. L., Rubin, L. H., & Maki, P. M. (2008). Effects of menstrual cycle phase and oral contraceptive use on verbal memory. *Hormones and Behavior, 54,* 286-293. doi: 10.1016/j.yhbeh.2008.03.006

Muhn, P., Krattenmacher, R., Beier, S., Elger, W., & Schillinger, E. (1995). Drospirenone: A novel progestogen with antimineralcorticoid and antiandrogenic activity. Pharmacological characterization in animal models. *Contraception, 51,* 99-110. doi: 10.1016/0010-7824(94)00015-O

O'Reilly, M. A., Cunningham, C. J., Lawlor, B. A., Walsh, C. D. & Rowan, M. J. (2004). The effect of the menstrual cycle on electrophysiological and behavioral measures of memory and mood. *Psychophysiology, 41,* 592-603. doi: 10.1111/j.1469-8986.2004.00194.x

Ostatnikova, D., Putz, Z., Celec, P., & Hodosy, J. (2002). May testosterone levels and their fluctuations influence cognitive performance in humans? *Scripta Medica, 75*(5), 245-254.

Palatsi, R., Hirvensalo, E., Liukko, P., Malmiharju, T., Riihiluoma, P., & Yiostalo, P. (1984). Serum total and unbound testosterone and sex hormone binding globulin (SHBG) in female acne patients treated with two different oral contraceptives. *Acta Dermato-venereologica, 64 (6),* 517-523.

Pearson, R., & Lewis, M. B. (2005). Fear recognition across the menstrual cycle. *Hormones and Behavior, 47,* 267-271. doi: 10.1016/j.yhbeh.2004.11

Phillips, A., Demarest, K., Hahn, D. W., Wong, F., & McGuire, J. L. (1990). Progestational and androgenic receptor binding af- finities and in vivo activities of norgestimate and other progestins. *Contraception, 41,* 399–410. doi: 10.1016/0010-7824(90)90039-X

Phillips, A., Hahn, D. W., & Klimek, S. (1987). A comparison of the potencies and activities of progestogens used in contraceptives. *Contraception, 36 (2),* 181-192. doi: 10.1016/0010-7824(87)90013-8

Puts, D. A., Cardenas, R. A., Bailey, D. H., Burriss, R. P., Jordan, C. L., & Breedlove, S. M. (2010). Salivary testosterone does not predict mental rotation performance in men or women. *Hormones and Behavior, 58*(2), 282-289. doi: 10.1016/j.yhbeh.2010.03.005

Raven, J. C. (1960) *Guide to the standard progressive matrices.* London: H.K. Lewis.

Rosenberg, L., & Park, S. (2002). Verbal and spatial functions across the menstrual cycle in healthy young women. *Psychoneuroendocrinology, 27,* 835-841. doi: 10.1016/S0306-4530(01)00083-X

Rothman, M. S., Carlson, N. E., Xu, M., Wang, C., Swerdolf, R., Lee, P., … Wierman, M. E. (2011). Reexamination of testosterone, dehydrotestosterone, estradiol and estrone levels across the menstrual cycle and in postmenopausal women measured by liquid chromatography-tandem mass spectrometry. *Steroids, 76,* 177-182. doi: 10.1016/j.steroids.2010.10.010

Rowlands, S. (2003). Newer progestogens. *Family Planning and Reproductive Health Care, 29,* 13-16. doi: 10.1783/147118903101197188

Salonia, A., Pontillo, M., Nappi, R. E., Zanni, G., Fabbri, F., Scavini, M., … Montorsi, F. (2008). Menstrual cycle-related changes in circulating androgens in healthy women with self- reported normal sexual function. *Journal of Sexual Medicine, 5(4),* 854–863. doi: 10.1111/j.1743-6109.2008.00791.x

Schindler, A. E., Campagnoli, C., Druckman, R., Huber, J., Pasqualini, J. R., Schweppe, K. W., & Thijssen, J. H. H. (2008). Reprint of classification and pharmacology of progestins. *Maturitas, 61(1-2),* 171-180. doi: 10.1016/j.maturitas.2003.09.014

Schindler, A. E. (2013). Non-contraceptive benefits of oral hormonal contraceptives. *Endocrinology and Metabolism, 11(1),* 41-47. doi: 10.5812/ijem.4158

Schultheiss, O. C., Dargel, A., & Rohde, W. (2003). Implicit motives and gonadal steroid hormones: effects of menstrual cycle phase, oral contraceptive use, and relationship status. *Hormones and Behavior, 43,* 293-301. doi: 10.1016/S0018-506X(03)00003-5

Shepard, R., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701-703. doi: 10.1126/science.171.3972.701

Silverman, I., & Phillips, K. (1993). Effects of estrogen changes during the menstrual cycle on spatial performance. *Ethology and Sociobiology, 14,* 257-270. doi: 10.1016/0162-3095(93)90021-9

Sitruk-Ware, R. (2008). Reprint of pharmacological profile of progestins. *Maturitas, 61 (1-2),* 151-157. doi: 10.1016/j.maturitas.2008.11.011

Sitruk-Ware, R., & Michelle, D. R. (Eds.). (2000). *Progestins and antiprogestins in clinical practice.* New York: Marcel Decker.

Slabbekoorn, D., Van Goozen, S. H., Megens, J., Gooren, L. J., & Cohen- Kettenis, P. T. (1999). Activating effects of cross-sex hormones on cognitive functioning: a study of short-term and long-term hormone effects in transsexuals. *Psychoneuroendocrinology, 24,* 423-447. doi: 10.1016/S0306-4530(98)00091-2

Soares, C. N., Cohen, L. S., Otto, M. W., & Harlow, B. L. (2001). Characteristics of women with premenstrual dysphoric disorder (PMDD) who did or did not report history of depression: A preliminary report from the Harvard study of moods and cycles. *Journal of Women's Health and Gender-Based Medicine, 10(9),* 873-878. doi: 10.1089/152460901753285778

Solís-Ortiz, S., & Corsi-Cabrera, M. (2008). Sustained attention is favored by progesterone during early luteal phase and visuo-spatial memory by estrogens during ovulatory phase in young women. *Psychoneuroendocrinology, 33,* 989-998. doi: 10.1016/j.psyneuen.2008.04.003

Spona, J., Elstein, M., Feichtinger, W., Sullivan, H., Ludicke, F., Muller, U., & Dusterberg, B. (1996). Shorter pill-free interval in combined OCs decreases follicular development. *Contraception, 54,* 71-77.

Stanczyk, F. Z. (2003). All progestins are not created equal. *Steroids, 68,* 879-890. doi: 10.1016/j.steroids.2003.08.003

Taieb, J., Mathian, B., Millot, F., Patricot, M. C., Mathieu, E., Queyrel, N., … Boudou, P. (2003). Testosterone measured by 10 immunoassays and by isotope-dilution gas chromatography-mass spectrometry in sera from 116 men, women and children. *Clinical Chemistry, 49,* 1381-1395. doi: 10.1373/49.8.1381

Temple, C. M., & Cornish, K. M. (1993). Recognition memory for words and faces in schoolchildren: A female advantage for words. *British Journal of Developmental Psychology, 11,* 421-426. doi: 10.1111/j.2044-835X.1993.tb00613.x

Thorneycroft, I. H., Stanczyk, F. Z., Bradshaw, K. D., Ballagh, S. A., Nichols, M. & Weber, M. E. (1999). Effects of low-dose oral contraceptives on androgenic markers and acne. *Contraception, 60,* 255-262. doi: 10.1016/S0010-7824(99)00093-1

Timmons, B. W., Hamadeh, M. J., Devries, M. C., & Tarnopolsky, M. A. (2005). Influence of gender, menstrual phase, and oral contraceptive use on immunological changes in response to prolonged cycling. *Journal of Applied Physiology, 99,* 979-985. doi: 10.1152/japplphysiol.00171.2005

Toffol, E., Heikinheimo, O., Kopponen, P., Luoto, R. & Partonen, T. (2011). Hormonal contraception and mental health: results of a population-based study. Human *Reproduction, 26(11),* 2085-3093. doi: 10.1093/humrep/der269

Van der Vange, N., Blankenstein, M. A., Kloosterboer, H. J., Haspels, A. A., & Thijssen, J. H. H. (1990). Effects of seven low-dose combined oral contraceptives on sex hormone binding globulin, corticosteroid binding globulin, total and free testosterone. *Contraception, 41(4),* 345-352. doi: 10.1016/0010-7824(90)90034-S

Van Goozen, S. H. M., Cohen-Kettenis, P. T., Gooren, L. J. G., Frijda, N. H., & Van de Poll, N. E. (1995). Gender differences in behavior: activating effects of cross-sex hormones. *Pyschoneuroendocrinology, 20 (4),* 343-363. doi: 10.1016/0306-4530(94)00076-X

Van Goozen, S. H. M., Cohen-Kettenis, P. T., Gooren, L. J. G., Frijda, N. H., & Van de Poll, N. E. (1994) Activating effects of androgens on cognitive performance: Causal evidence in a group of female-to-male transsexuals. *Neuropsychologia, 32,* 1153-1157. doi: 10.1016/0028-3932(94)90099-X

Van Honk, J., & Schutter, D. J. L. G. (2007). Testosterone reduces conscious detection of signals serving social correction: Implications for antisocial behavior. *Psychological Science, 18,* 663-667. doi: 10.1111/j.1467-9280.2007.01955.x

Van Rooijen, M., Silveira, A., Hamsten, M. D., & Bremme, K. (2004). Sex hormone-binding globulin- A surrogate marker for the prothrombotic effects of combined oral contraceptives. *American Journal of Obsetrics and Gynecology, 190,* 332-337. doi: 10.1016/S0002-9378(03)00950-5

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117,* 250-270. doi: 10.1037/0033-2909.117.2.250

Weiss, E. M., Kemmler, G., Deisenhammer, E. A., Fleischhacker, W. W., & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and Individual Differences, 35,* 863-875. doi: 10.1016/S0191-8869(02)00288-X

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,* 156-164. doi: 10.1037/1064-1297.16.2.156

Wiegratz, I., Kutschera, E., Lee, J. H., Moore, C., Mellinger, U., Winkler, U. H., & Kuhl, H. (2003). Effect of four different oral contraceptives on various sex hormones and serum-binding globulin. *Contraception, 67,* 25-32. doi: 10.1016/S0010-7824(02)00436-5

Wilkinson, G. S., & Robertson, G. J. (2006). Wide-Range Achievement Test – Fourth Edition. Lutz, FL: Psychological Assessment Resources.

Zelinski, E. M., Gilewski, M. J., & Schaie, K. W. (1993). Individual differences in cross-sectional and 3-year longitudinal memory performance across the adult life span. *Psychology and Aging, 8,* 176-186. doi: 10.1037/0882-7974.8.2.176

Zimmerman, Y., Eijkemans, M. J. C., Coelingh Bennink, H. J. T., Blankenstein, M. A., & Fauser, B. C. J. M. (2014). The effect of combined oral contraception on testosterone levels in healthy women: a systematic review and meta-analysis. *Human Reproduction Update, 20(1),* 76-105. doi: 10.1093/humupd/dmt038

Zurawin, R., & Ayensu-Coker, L. (2007). Innovations in contraception: A review. *Clinical Obstetrics and Gynecology, 50*(2), 425-439. doi: 10.1097/GRF.0b013e31804b1be6

Figure 1. *Cognitive Performance Patterns Across the Menstrual Cycle*

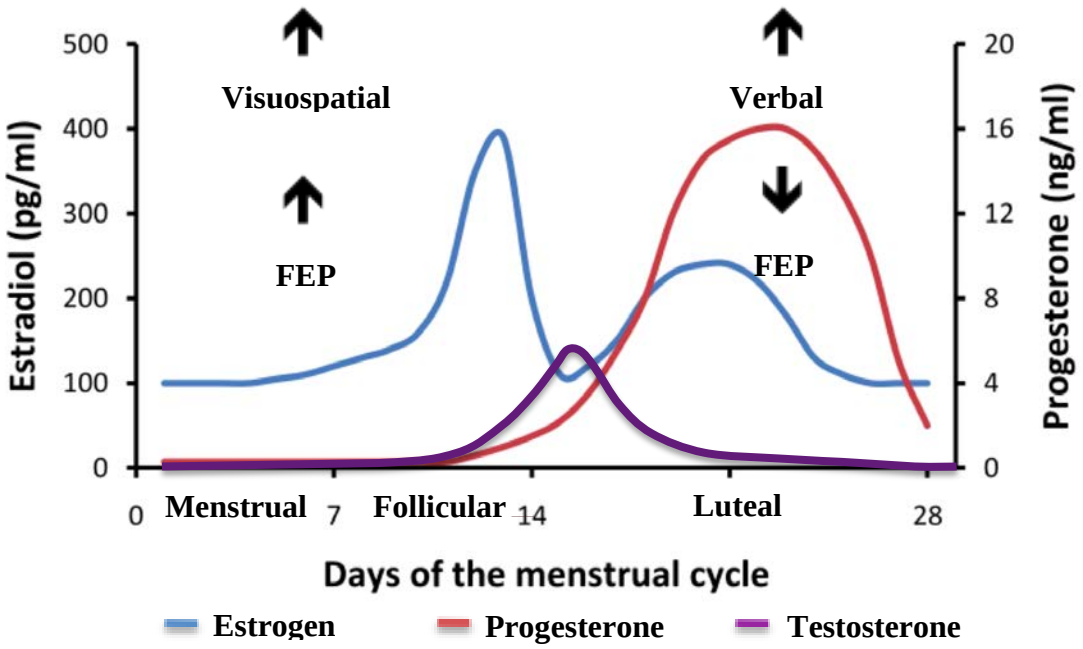


Table 1. *List of Measures and Variables of Interest*

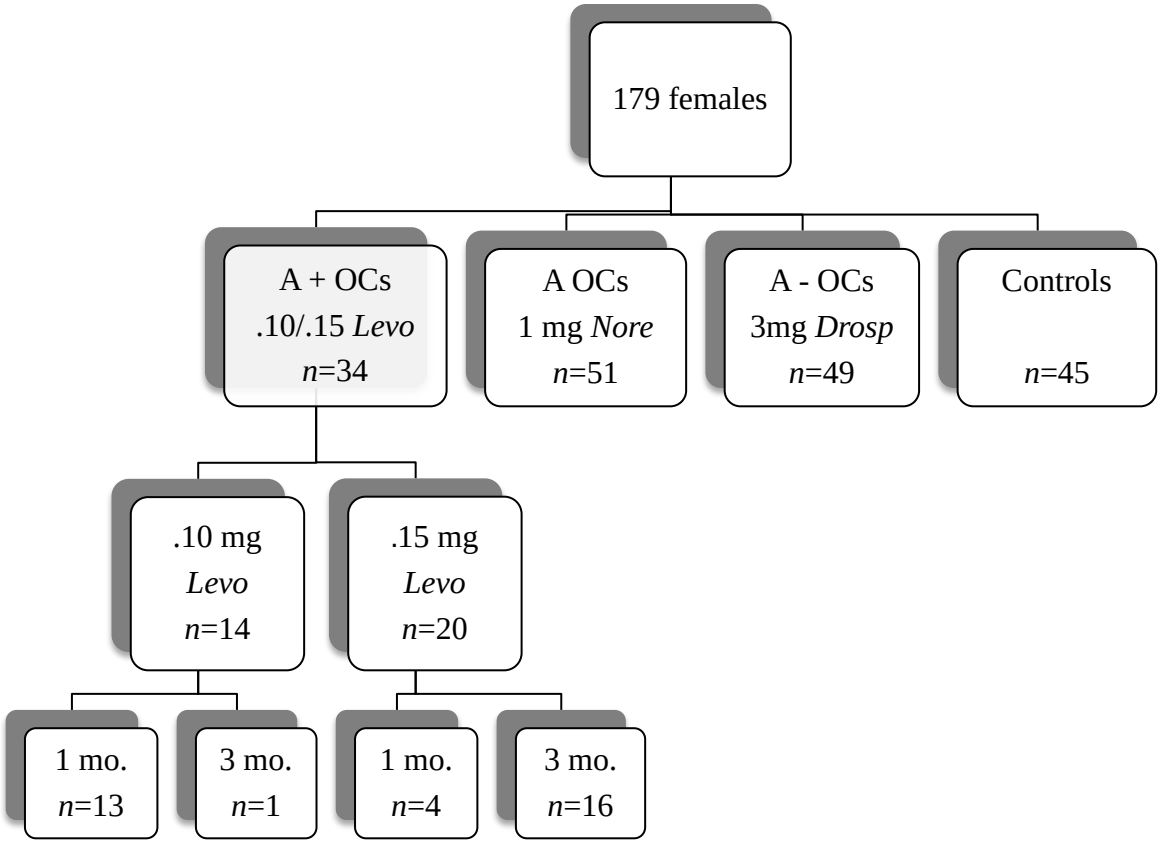
Domain	Measure	Variable of Interest
Visuospatial		
	PLOT	Total % Correct (raw)
	PMAT	Total % Correct (raw)
Verbal		
	sPVRT	Total % Correct (raw)
	CPW	Total % Correct (raw)
Facial Expression Processing		
	ER40	Total % Correct (raw)
		Anger % Correct(raw)
	MEDF	Total % Correct (raw)
		Anger % Correct(raw)

Table 2. *List of Oral Contraceptives by Group*

Antiandrogenic (*n*=49)	Androgenic (*n*=51)	Highly Androgenic (*n*=34)	
Drospirenone 3 mg	*Norethindrone* 1 mg	*Levonorgestrel* .15 mg	.10 mg
Yaz (*n*=25)	Loestrin (*n*=42)	Seasonique* (*n*=11)	Aviane (*n*=7)
Yasmin (*n*=10)	Junel (*n*=4)	Jolessa* (*n*=3)	Alesse (*n*=3)
Ocella (*n*=9)	Microgestin (*n*=4)	Seasonale* (*n*=2)	Lutera (*n*=2)
Gianvi (*n*=4)	Estrostep (*n*=1)	Levora (*n*=2)	Sronyx (*n*=1)
Beyaz (*n*=1)		Portia (*n*=2)	LoSeasonique* (*n*=1)

Note: * Extended cycle (3-mo.) formulations

Figure 2. *Participants*



Note: *Levo = Levonorgestrel*
Nore = Norethindrone
Drosp = Drospirenone

Table 3. *Demographic Characteristics by Group*

Variable	Control (*n*=45)	Anti (*n*=49)	Andro (*n*=51)	Highly Andro (*n*=34)	*p*
Age	17.78 (1.69)	18.02 (1.33)	18.06 (1.46)	17.85 (1.73)	.791
Education	11.53 (1.80)	11.65 (1.38)	11.80 (1.52)	11.71 (1.78)	.874
Mother's Education	14.71 (2.50)	14.75 (2.08)	14.61 (2.23)	14.74 (2.22)	.990
Father's Education	14.89 (3.02)	14.83 (2.99)	14.14 (2.17)	14.91 (2.29)	.442
WRAT-IV (SS)	105.93 (13.07)	101.92 (13.61)	103.14 (12.26)	105.44 (12.03)	.398
Race					.485
Caucasian	38 (84.4%)	40 (81.7%)	43 (84.3%)	26 (76.5%)	
Af. American	3 (6.7%)	6 (12.2%)	2 (3.9%)	2 (5.9%)	
Multi-Race	4 (8.9%)	3 (6.1%)	6 (11.8%)	6 (17.6%)	
Ethnicity					.655
Non-Hispanic	42 (93.3%)	44 (89.9%)	49 (96.1%)	31 (91.2%)	
Hispanic/Latino	3 (6.7%)	5 (10.1%)	2 (3.9%)	3 (8.8%)	
Lifetime Depressive Symptoms					.365
None	24 (53.3%)	16 (32.7%)	27 (52.9%)	14 (41.2%)	
Mild/Moderate	7 (15.6%)	12 (24.6%)	11 (21.6%)	8 (23.5%)	
Severe	14 (31.1%)	21 (42.9%)	13 (25.5%)	12 (35.3%)	

Note: Data are presented as means (standard deviations) or number (percentage)

Table 4. *Means and Standard Deviations for Outcome Measures*

Variable	Anti (*n*=49)	Control (*n*=45)	Andro (*n*=51)	Highly Andro (*n*=34)	Total (*n*=179)
PLOT	47.02 (17.96)	45.83 (19.34)	46.73 (18.15)	48.16 (15.96)	46.86 (18.88)
PMAT	64.09 (16.23)	62.56 (18.42)	66.69 (19.52)	58.70 (19.9)	63.42 (18.52)
CPW	94.69 (5.97)	93.56 (8.60)	93.63 (5.94)	95.37 (5.54)	94.23 (6.64)
SPVRT	84.90 (9.98)	85.04 (14.22)	87.58 (11.70)	89.22 (10.38)	86.52 (11.76)
ER40 (Total)	84.69 (5.51)	86.33 (7.16)	87.25 (6.69)	87.72 (4.98)	86.41 (6.28)
ER40 (Anger)	68.11 (17.50)	74.17 (17.35)	73.28 (16.59)	73.90 (15.50)	72.21 (16.89)
MEDF (Total)	79.28 (7.45)	77.05 (8.25)	78.61 (7.73)	81.85 (8.21)	79.01 (7.98)
MEDF (Anger)	80.68 (11.12)	79.66 (12.73)	81.27 (10.82)	83.73 (12.14)	81.17 (11.64)

Note: Data are presented as means (standard deviations)

Table 5. *One-Way ANOVAs for Effects of Androgenicity on Outcome Measures*

Variable	$F(2, 178)$	p	η^2
PLOT	.11	.954	.002
PMAT	1.33	.267	.022
CPW	.71	.551	.012
SPVRT	1.29	.280	.022
ER40-Total	2.06	.107	.034
ER40-Anger	1.35	.259	.023
MEDF-Total	2.46	.064	.040
MEDF-Anger	.83	.480	.014

Table 6. *One-Way MANOVA for Effects of Androgenicity on Facial Expression Processing*

Variables	$F(6, 348)$	p	*Wilk's Λ*	η^2
ER40-Total MEDF-Total	2.23	.040	.927	.037