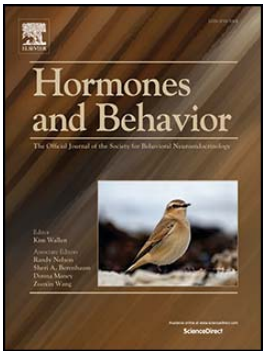

Effects of exogenous testosterone and mating context on men's preferences for female facial femininity

Brian M. Bird, Lisa L.M. Welling, Triana L. Ortiz, Benjamin J.P. Moreau, Steve Hansen, Michael Emond, Bernard Goldfarb, Pierre L. Bonin, Justin M. Carré

PII: S0018-506X(16)30112-X
DOI: doi: 10.1016/j.yhbeh.2016.08.003
Reference: YHBEH 4087

To appear in: *Hormones and Behavior*

Received date: 7 March 2016
Revised date: 2 August 2016
Accepted date: 3 August 2016

Please cite this article as: Bird, Brian M., Welling, Lisa L.M., Ortiz, Triana L., Moreau, Benjamin J.P., Hansen, Steve, Emond, Michael, Goldfarb, Bernard, Bonin, Pierre L., Carré, Justin M., Effects of exogenous testosterone and mating context on men's preferences for female facial femininity, *Hormones and Behavior* (2016), doi: 10.1016/j.yhbeh.2016.08.003

Effects of Exogenous Testosterone and Mating Context on Men's Preferences for Female Facial Femininity

Brian M. Bird[1], Lisa L.M. Welling[2], Triana L. Ortiz[3], Benjamin J.P. Moreau[3], Steve Hansen[4], Michael Emond[1], Bernard Goldfarb[5], Pierre L. Bonin[5] & Justin M. Carré[3*]

[1]Department of Psychology, Laurentian University, Sudbury, Ontario, Canada

[2]Department of Psychology, Oakland University, Rochester, Michigan, United States

[3] Department of Psychology, Nipissing University, North Bay, Ontario, Canada

[4] Physical and Health Education, Nipissing University, North Bay, Ontario, Canada

[5] Northern Ontario School of Medicine, Sudbury, Ontario, Canada

*Correspondence concerning this article should be addressed to:

Justin M. Carré, Ph.D.
Assistant Professor
Department of Psychology
Nipissing University
100 College Dr.
North Bay, Ontario, Canada
Email: justinca@nipissingu.ca

Abstract

Correlational research suggests that men show greater attraction to feminine female faces when their testosterone (T) levels are high. Men's preferences for feminine faces also seem to vary as a function of relationship context (short versus long-term). However, the relationship between T and preferences for female facial femininity has yet to be tested experimentally. In the current paper, we report the results of two experiments examining the causal role of T in modulating preferences for facial femininity across both short and long-term mating contexts. Results of Experiment 1 (within-subject design, $n = 24$) showed that participants significantly preferred feminized versus masculinized versions of women's faces. Further, participants showed a stronger preference for feminine faces in the short versus the long-term context after they received T, but not after they received placebo. Post-hoc analyses suggested that this effect was driven by a lower preference for feminine faces in the long-term context when on T relative to placebo, and this effect was found exclusively for men who received placebo on the first day of testing, and T on the second day of testing (i.e., order X drug X mating context interaction). In Experiment 2 (between-subject design, $n = 93$), men demonstrated a significant preference for feminized female faces in the short versus the long-term context after T, but not after placebo administration. Collectively, these findings provide the first causal evidence that T modulates men's preferences for facial femininity as a function of mating context.

Key Words: Testosterone; facial preferences; femininity; mate preferences; hormones; mating

Effects of Exogenous Testosterone and Mating Context on Men's Preferences for Female Facial Femininity

Introduction

Evidence indicates that humans prefer opposite sex faces that align with sex-typicality (i.e., men prefer feminine faces; women prefer masculine faces) for sexual relationships, where such preferences are thought to represent an adaptive strategy for securing mates with greater immunocompetence or fertility advantages (Gangestad & Scheyd, 2005; Lee et al., 2013; Little et al., 2007; Little et al., 2008; O'Connor et al., 2013; Wheatley et al., 2014). Other evidence suggests that facial preferences may also vary as a function of the perceiver's circulating hormone levels, perhaps helping to facilitate mating goals. For example, women show the greatest preferences for masculinity in men's faces when they are at peak fertility, and when their testosterone levels are high (Bobst et al., 2014; Little & Jones, 2012; Penton-Voak & Perrett, 2000; Welling et al., 2007, see Gildersleeve et al., 2014 for meta analysis), which may function to increase offspring health through transmission of superior genes (Gangestad et al., 2004; Johnston et al., 2001).

In ancestral environments, the ability to determine the quality of a mate from physical appearance would have afforded survival or reproductive advantages to those who exploited these signals (Little et al., 2011; Little, 2014). The finding that men generally prefer feminine faces (e.g., Jones et al., 2007; Komori et al., 2009; O'Connor et al., 2013) and that facial femininity is correlated with judgments of attractiveness and health by opposite sex individuals (Law Smith et al., 2006; Röder et al., 2013) as well as certain health indices and/or estrogen levels (Gray & Boothroyd, 2012; Jones et al., 2015; Thornhill & Gangestad, 2006; van Anders, 2010), longevity (Henderson & Anglin, 2003), and fertility (e.g., Jokela, 2009; Roberts et al., 2003) suggests that facial femininity may represent one such cue.

Recently, researchers have examined factors that map onto variability in men's preferences for facial femininity. For example, men scoring high on sensation seeking demonstrate greater preferences for feminine faces (Jones et al., 2007) and men who rate themselves as more attractive show a greater preference for femininity in short-term versus long-term mating contexts (Burriss et al., 2011). Other recent work has explored the role of men's endogenous testosterone (T) in modulating preferences for facial femininity. To the best of our knowledge, the only study that directly examined this relationship was conducted by Welling et al. (2008), whereby male participants entered the lab on two separate occasions for a facial preferences task, and provided saliva samples for the assessment of T. Each day, participants were asked to rate pairs of masculinized and feminized faces (1 masculinized and 1 feminized per pair) for their degree of attractiveness. Results showed that attractiveness ratings for the feminine female faces (but not feminine male faces) were highest on the day in which the participants had higher basal T-levels, suggesting that men may be more attracted to females who signal greater health or fertility when T-levels are high relative to low. One other study tacitly suggests that men's facial femininity preferences vary as a function of their T-levels: Welling et al. (2013) examined men's facial preferences following a competitive interaction, whereby participants were assigned to win or lose a first-person shooter video game against an unseen male confederate. Results revealed that winners showed an overall greater preference for feminine faces relative to losers. Additionally, for winners, femininity preferences in the short-term context were significantly higher than for the long-term context, whereas this difference was not present among losers. Because T-levels typically rise in winners relative to losers (e.g., Archer, 2006; Carré & Olmstead, 2015), stronger preferences in winners may have been mediated by changes in their T-levels (Welling et al., 2013).

Effective mating strategies are also argued to depend on relationship context. Feminine women are rated as more attractive, more intrasexually competitive, and more willing to engage in short-term mating (Fink et al., 2014). Furthermore, they show a greater interest in unrestricted sexual relationships (Boothroyd et al., 2008), and are perceived as more promiscuous (Brewer & Archer, 2007; Little et al., 2013) and as more likely to seek extra-pair copulations (i.e., cheat on a partner). Thus, differential preferences for feminine women across mating contexts (e.g., Burriss et al., 2011; Little et al., 2011; Little et al., 2013) may represent a trade-off between the likelihood of successfully reproducing with a healthy, feminine partner in a short-term relationship, while avoiding the potential for partner defection in a long-term relationship, as well as avoiding the difficulty of defending a sexually attractive mate from other men. However, the extent to which T-levels influence men's shifts in preferences for facial femininity across mating contexts remains untested.

Although studies investigating basal T-levels and facial preferences can provide important information about hormonal associations with mating preferences, the correlational nature of these studies precludes the possibility of establishing causal relationships. This problem can be overcome by manipulating T levels via pharmacological challenge—a rapidly emerging line of research (reviewed in Bos et al., 2012). Evidence for varied partner preferences across mating contexts, coupled with the limited and strictly correlational research on men's T levels in relation to their femininity preferences, calls for an experimental protocol. Thus, the present paper employed 2 experiments (Experiment 1: within-subjects; Experiment 2: between-subjects) in double blind, placebo-controlled T-administration paradigms, in order to temporarily elevate T-concentrations in healthy young men, and subsequently measure their preferences for female facial femininity across both short- and long-term mating contexts. Based on previously

reviewed work suggesting that feminine faces are associated with judgments of health and fertility (e.g., Law Smith et al., 2006; Röder et al., 2013), as well as other work showing that T-levels are positively associated with mating success (e.g., Peters et al., 2008), and heightened attraction to feminine faces (Welling et al., 2008), men in the present experiments were expected to demonstrate a heightened preference for feminized female faces following T-administration, compared to the placebo condition. Additionally, the preference for feminine female faces in the T condition was expected to be more pronounced for contexts relating to short-term, rather than long-term relationships (Burris et al., 2011; Little et al., 2011; 2013), in light of the potential trade-off between attraction to a healthy and fertile partner who is willing to engage in short-term mating (i.e., more feminine face), and a faithful long-term partner who potentially poses less risk for partner defection (i.e., less feminine face).

Experiment 1

Methods

Participants. Our sample consisted of 30 healthy young men between the ages of 18 and 35 (Mean age = 21.21, *SD* = 2.19) who were part of a larger T-administration protocol at Nipissing University (*n* = 28 Caucasian, *n* = 1 Latin American, *n* = 1 First Nations/Aboriginal). Prior to enrollment in the study, each prospective participant was interviewed to determine his eligibility. Exclusion criteria for participants included the following: receiving prescription medication affecting hormone concentrations; taking performance enhancing substances; current diagnosis of a psychiatric disorder; diagnosed heart condition; and membership on a sports team or organization where T was a banned substance. Participants who qualified for the protocol consented to providing blood samples for future hormonal assay, as well as to having their T-levels temporarily manipulated. The study was approved by the Nipissing University Research

Ethics Board under protocol #140609, and each participant provided informed consent prior to the commencement of the protocol. Because of the inherently heterosexual nature of this protocol (i.e., rating opposite sex faces for partner attractiveness), non-heterosexual participants were removed prior to analysis (n = 2). Finally, data for 4 participants were lost due to computer malfunction. Thus, our final sample size for analyses was n = 24.

Stimuli. In line with previous work investigating sexually-dimorphic face preferences (DeBruine et al., 2006; Jones et al., 2007; Welling et al., 2007, 2008, 2013), the present study used prototype-based image transformations in order to objectively manipulate sexual dimorphism of 2D shape in facial images, creating masculinized and feminized images of the same individual that are matched for other variables (e.g., skin color, identity, texture: Rowland & Perrett, 1995). Briefly, prototype images (i.e., an average male face and an average female face) were created by averaging a group of male and a group of female images via widely-used computational methods in face perception studies (e.g., Jones et al., 2005; Penton-Voak et al., 1999; Welling et al., 2007). Once prototypes are established, individual stimuli are created by adding or subtracting a percentage of the differences in position between the prototype images from the corresponding points on a third face (for technical details see Rowland & Perrett, 1995; Tiddeman, Burt, & Perrett, 2001).

For the present study, 50% of the linear differences in 2D shape between symmetrized male and female prototypes were either added or subtracted from 20 young Caucasian female adults (Mean age = 20.52 year, SD = 2.78), creating 40 images (i.e., 20 pairs, with each pair including one masculinized and one feminized version of the same individual). The resulting images were subjected to a manipulation check in previous work, and were rated by an independent group of observers as representing ecologically valid representations of feminine or

masculine faces (Welling et al., 2007, 2008). See Figure 1 for an example of masculinized and feminized stimuli.

Procedure. Testing for the full protocol occurred across three separate days. Day 1 involved familiarizing participants with the experimental procedures, obtaining informed consent, as well as the administration of a number of demographic and self-report questionnaires as part of the larger protocol. Day 1 took approximately 1 hour to complete.

***Hormone and placebo administration*.** On day 2 of testing, a registered nurse drew 10 mL of blood from the antecubital area of the right arm. Next, participants either received 150 mg of AndroGel®—a topical gel commonly used for hypogonadal men—or equivalent placebo (counter-balanced across participants). AndroGel® or placebo was applied to both upper arm and shoulder areas by a male research assistant blind to the drug condition (application site established based on the recommendations provided by AndroGel®). Additionally, blood samples were drawn at 60 and 120 min post drug administration, alternating between the right and left arms. After 120 min, participants then performed a series of computer-based tasks assessing social perception, cognition, and decision-making abilities over approximately two hours. Assessment of face preferences occurred approximately 3 h 15 min after gel application (M = 191.25 min, SD = 5.7 min). We chose this time-course for the assessment of face preferences as previous pharmacokinetic work indicates that T concentrations begin to rise 2 hours after gel application and peak concentrations occur 3 hours after application (Eisenegger et al., 2013). Moreover, recent evidence suggests that a single administration of T can rapidly (within 45 to 90 min) modulate brain function (see Goetz et al., 2014; van Wingen et al., 2008). Day 3 took place two weeks following Day 2 and was identical in nature to Day 2 described above, with the exception that participants received whichever drug they did not receive on their

original testing day (AndroGel® or placebo). At the conclusion of Day 3 of testing, participants were asked whether they believed they received testosterone on the 2nd or 3rd day of testing. A binomial test indicated that participants were no better than chance at guessing which day they received testosterone (p = .10).

Prior to performing the facial femininity task, participants completed other tasks for hypotheses unrelated to the present study. These tasks included the 'Reading the Mind in the Eyes' task (Carré et al., 2015), 'Pick Your Own Face' task (Welling et al., 2016), risk-preference task, moral decision-making task (Arnocky et al., 2016), emotion recognition task, and selective visual attention tasks (inhibition of return)[1].

***Face preferences task*.** Participants rated 20 pairs of female faces (each pair with one masculinized and one feminized version of the same individual) twice: once for attractiveness as a short-term partner, and once for attractiveness as a long-term partner. The 20 pairs were all evaluated for one context before moving on to the other. Randomization was used for each variable, including the order of context, the order of stimuli, and the side of the screen on which the masculine or feminine version of each pair was presented. Verbal instructions for each participant were as follows:

> "This task requires you to rate 20 pairs of faces for their attractiveness as a long- or short-term relationship. It's important that you understand what we mean by each, so please listen to these definitions. **Short-term relationship**: you are looking for the type of person that would be attractive in a short-term relationship. This implies that the relationship may not last a long time. Examples of this type of relationship would include a single date accepted on the spur of the moment, an affair within a long-term relationship, or a one-night stand.
>
> **Long-term relationship:** you are looking for the type of person that would be attractive in a

[1] Statistically controlling for performance on these other measures did not alter the significance of any of the results.

long-term relationship. Examples of this type of relationship would include someone you may want to move in with, someone you may consider leaving a current partner to be with, or someone you may wish to marry (or enter a relationship on similar grounds as marriage). For each preference task, try not to think too long and hard about which face you're going to choose. We are most interested in your first impressions. The image pairs look very similar, but they are subtly different. You will get one practice trial, and then you will proceed to the main rating task. Please read the instructions carefully on the screen at the beginning of the task prior to beginning. Do you have any questions?"

Following verbal instructions, participants could begin the task. Instructions on the screen prior to the first trial were as follows: "**Short-term relationship:** You will see 20 pairs of facial photographs of women. Please choose which of the two photographs you feel is most ATTRACTIVE for a SHORT-TERM RELATIONSHIP by clicking on the face you prefer. A short-term relationship refers to an uncommitted, purely sexual relationship such as a one-night stand." OR "**Long-term relationship:** You will see 20 pairs of facial photographs of women. Please choose which of the two photographs you feel is most ATTRACTIVE for a LONG-TERM RELATIONSHIP by clicking on the face you prefer. A long-term relationship refers to a committed relationship, such as marriage."

Initial Processing of Data

Hormone assays. Blood samples were assayed for total-T concentrations using commercially-available enzyme immunoassay kits (DRG International). As standard procedure, all samples were assayed in duplicate, and the averages of the duplicates were recorded for statistical analyses. The intra- and inter-assay coefficients of variation were 4.19% and 5.34%, respectively. The analytical sensitivity of the testosterone assay is .085 ng/mL.

Face preferences. For each participant, the number of trials in which the more feminine face from each pair was chosen, was calculated for each context (short- term vs. long- term) and drug (testosterone vs. placebo).

Results and Discussion

Testosterone Concentrations

A 3-Time by 2-Drug repeated-measures ANOVA on T-concentrations was performed [within-subject factors: Time (baseline vs. 60 min vs. 120 min) and Drug (Testosterone vs. Placebo)]. Results revealed main effects of Drug [$F(1, 23) = 29.44$, $p < .001$, $\eta^2_G = .20$][2] and Time [$F(2, 46) = 42.09$, $p < .001$, $\eta^2_G = .19$]. These main effects were qualified by a significant Drug by Time interaction [$F(2, 46) = 36.13$, $p < .001$, $\eta^2_G = .11$]. Post-hoc analyses indicated that T-concentrations were higher after Androgel® compared to placebo at 60 minutes post gel application [$t(23) = 5.38$, $p < .001$, Cohen's $D = 1.17$] and 120 post gel application [$t(23) = 6.94$, $p < .001$, Cohen's $D = 1.42$]. Overall, participants in the AndroGel® condition experienced an average increase of 56.39% in T from baseline to 120 mins. There were no differences in T-concentrations for Androgel® versus placebo prior to gel application [$t(23) = -.05$, $p = .96$] (See Figure 2).

Femininity Preferences

One sample *t*-tests comparing the number of times the feminine versions of the female faces were chosen against the chance value of 10 revealed that participants chose the feminine

[2] Eta-squared (η^2) and partial eta-squared (η^2_p) are not particularly well-suited for making comparisons across studies with different designs (e.g., within-subject design vs. between-subject design; Fritz, Morris & Richler, 2012). The generalized eta-squared (η^2_G) is a more appropriate measure of effect size for repeated measures and/or mixed factor designs and when one wishes to compare effect sizes across different experimental designs (Olejnik & Algina, 2003; Bakeman, 2005). Thus, we report η^2_G as an estimate of effect size for ANOVAs. We also report Cohen's *D* (Cohen, 1988) for simple group comparisons (paired sample t-tests and independent sample t-tests).

face as more attractive across both drugs (T-Day = Testosterone Day, P-Day = Placebo Day) and contexts: T-Day/Short-term [$t(23) = 10.39$, $p < .001$], T-Day/Long-term [$t(23) = 5.05$, $p < .001$], P-Day/Short-term [$t(23) = 12.31$, $p < .001$], P-Day/Long-term [$t(23) = 10.32$, $p < .001$].

A 2-Drug by 2-Context by 2-Order of Drug Administration mixed ANOVA [within-subject factors: Drug (Testosterone vs. Placebo); Context (Short-Term vs. Long-Term); between-subject factor: Order of Drug Administration (T then P vs. P then T)] was conducted to test for differences in the frequency of trials in which the feminine face was selected as more attractive as a function of context and drug condition, and whether the order in which the drug was administered influenced the pattern of findings. Results revealed a main effect for Context [F (1, 22) = 7.21, $p = .01$, $\eta^2_G = .04$], whereby participants demonstrated a stronger preference for feminine faces in the short-term relative to the long-term mating context. There were no main effects of drug condition [$F(1, 22) = 2.88$, $p = .10$, $\eta^2_G = .017$] or Order of Drug Administration [[$F(1, 22) < .01$, $p = .99$, $\eta^2_G < .001$]. There were no Drug by Order [$F(1, 22 = .91$, $p = .35$, $\eta^2_G = .005$] or Context by Order [$F(1, 22) = .90$, $p = .35$, $\eta^2_G = .005$] interactions. However, there was a significant Drug by Context interaction [$F(1, 22) = 5.28$, $p = .031$, $\eta^2_G = .013$]. Unexpectedly, we also observed a significant Drug by Context by Order of Drug Administration interaction [$F(1, 22) = 14.01$, $p = .001$, $\eta^2_G = .033$]. Analyses split by order of drug administration indicated that the Drug by Context interaction was specific to those receiving P on the first test session and T on the second test session [$F(1, 11) = 13.32$, $p = .004$, $\eta^2_G = .17$)]. Specifically, there was a stronger preference for facial femininity in the short-term mating context versus long-term mating context after T [$t(11) = 3.54$, $p = .005$, Cohen's $D = 1.11$), but not P ($t(11) = -.12$, $p = .91$, Cohen's $D = -.03$]. This effect was driven by a weaker preference for facial femininity in the long-term mating context for T ($M = 13.50$, $SE = 1.27$) relative to P ($M =$

16.75, *SE* = .85; $t(11) = -2.81$, $p = .017$, Cohen's $D = -1.27$). There was no Drug by Context interaction among those who received T on the first test session and P on the second test session [$F(1, 11) = 1.66$, $p = .22$, $\eta^2_G = .01$)] (See Figure 3).

Results of Experiment 1 indicate that 1) regardless of mating context, participants preferred feminine female faces significantly more than masculine female faces; 2) preferences for feminine female faces were significantly higher in the short-term context than the long-term context; 3) this effect was particularly robust after T administration; and 4) the effect of T on preferences for facial femininity in short-term versus long-term mating contexts was exclusively found among men who received P on the first test session, and T on the second test session.

Although there are considerable strengths associated with within-subject designs (e.g., increased power of having participants serve as their own control), there are also limitations. For instance, it is difficult to interpret the significant Order of Drug Administration by Drug by Context interaction observed in the current study. One possibility is that this order-dependent effect is spurious, especially in light of the small sample sizes used when splitting analyses by order of drug administration ($n = 12$ men per condition). Another possibility is that the novelty of the research environment (blood draws, drug administration, travelling from Urologist's office to novel research environment) may have increased stress levels, and this may have in turn ultimately blocked the context-dependent effect of T. Indeed, a growing body of work indicates that endogenous T positively correlates with several behavioral outcomes (e.g., dominance, risk-taking), but only among individuals with relatively low cortisol levels (Mehta & Josephs, 2010; Mehta et al., 2015). Another limitation of Experiment 1 was the absence of a blood sample directly before the facial femininity task. We based our timing of behavioral assessment on previous research indicating that a single 150 mg dose of Androgel® led to increased T

concentrations in healthy young men for up to 7 hours after drug administration (Eisenegger et al., 2013). Nevertheless, results from Experiment 1 indicate that T concentrations peaked more rapidly compared to previous work (60 mins vs. 180 mins; Eisenegger et al., 2013).

To address these limitations, we employed a second experiment using a between-subjects design to examine the extent to which the Drug x Context interaction is robust, while simultaneously ruling out any potential order effects. In Experiment 2, we collected additional blood samples throughout the protocol, including a final blood draw immediately prior to the facial rating task to verify that blood serum levels remained significantly elevated directly before testing.

Experiment 2

Method

Participants. Our sample consisted of 120 healthy young men between the ages of 18 and 35 (Mean age = 25.27 years, *SD* = 4.98) who were part of a larger T-administration protocol run at a medical research facility in Sudbury, Ontario. Subjects were recruited from advertising on local media sites, through medical research participant databases, as well as through local colleges and universities. Prior to enrollment in the study, each prospective participant was interviewed to determine his eligibility. Exclusionary criteria were identical to Experiment 1. Participant ethnicities were self-reported as follows: 77.5% Caucasian, 13.1% First Nations/Aboriginal, 4.1% Asian, 1.7% Latin American, and 3.3% Other. Each participant provided informed consent prior to the commencement of the protocol. Because of the inherently heterosexual nature of this protocol (i.e., rating opposite sex faces for partner attractiveness), non-heterosexual participants were removed prior to analysis (*n* = 11). Additionally, individuals

were removed who had femininity ratings more than 3 standard deviations below the mean (n = 2), did not complete the task (n = 3), or only rated faces for one of the two contexts (n = 11). Therefore, the final sample size for the present study was 93 (T n = 48; Placebo n = 45).

Stimuli

The feminine and masculine pairs were identical to those used in Experiment 1. The order of presentation of the stimuli and screen-side of presentation were randomized in a similar manner to that used in Experiment 1.

Procedure

Testing for the full protocol occurred in a single session. Participants reported to the laboratory at either 10:00am or 1:00pm. Upon arrival, participants completed informed consent and had the opportunity to ask any additional questions about the study. Following this introduction, participants completed a battery of online demographic and self-report questionnaires as part of the larger protocol.

Hormone and placebo administration. After the completion of the online questionnaires, participants received their initial blood draw, where a phlebotomist drew 10 mL of blood. Blood samples were allowed to clot and then were centrifuged at 3000 rpm, after which serum samples were extracted and then stored in -60°C refrigeration until assayed. Next, participants were randomly assigned to one of two experimental conditions: 150 mg of AndroGel®, or equivalent placebo. Drug condition (AndroGel® or placebo) was fully randomized across participants. Regardless of drug condition, a male research assistant who was blind to the experimental condition applied topical gel to the upper arm and shoulder area. After gel application, participants rested for 1 hour, after which they received their second blood draw, and then performed a series of computer-based tasks assessing social perception, cognition, and

decision-making abilities over approximately a one-hour span. The third and fourth blood draws were spread out over the rest of the protocol, with the final blood draw occurring directly before the facial femininity task that occurred at approximately 2 hours and 15 minutes post drug administration (M = 133.03 min, SD = 10.38 min). At the end of the protocol, participants were asked if they thought they had received T or placebo. A binomial test indicated that participants were precisely at chance level (p = 1.0) for correctly identifying whether or not they had received T.

Prior to the facial femininity task, participants completed other tasks for hypotheses unrelated to the present study. These tasks included the 'Pick Your Own Face' task (Welling et al., 2016), Point Subtraction Aggression Paradigm (Carré et al., in press), Balloon Analogue Risk Taking task, a risk-preference task, and an emotion recognition task[3].

Facial preferences task. The tasks and instructions were identical in nature to those reported in Experiment 1. Briefly, participants rated 20 pairs of female faces (each pair with one masculinized and one feminized version of the same individual) twice: once for attractiveness as a short-term partner, and once for attractiveness as a long-term partner. Twenty pairs were all rated for one relationship context before moving on to the other context. As with Experiment 1, each variable was randomized, including the order of context, the order of stimuli, and the side of the screen on which the masculine or feminine version of each pair was presented.

Initial Processing of Data

Hormone Assays. Using commercially-available enzyme immunoassay kits (DRG International), blood serum samples were assayed for total T concentrations. All samples were assayed in duplicate, with the average of the duplicates being recorded for statistical analyses.

[3] Statistically controlling for performance on these other measures did not alter the significance of any of the results.

The intra- and inter-assay coefficients of variation were 7.38% and 16.03%, respectively. The analytical sensitivity of the testosterone assay is .085 ng/mL.

Face preferences. As in Experiment 1, the number of trials in which the more feminine face from each pair was selected as attractive was calculated for each relationship context (short-term and long-term) and each experimental drug condition (T or placebo).

Results and Discussion

Testosterone Concentrations

A 4-Time by 2-Drug mixed-measures ANOVA on T-concentrations was performed **[**within subject factor: Time; between subject factor: Drug]. Results revealed main effects of Drug [$F(1,91) = 16.85$ $p < .001$, $\eta^2_G = .13$] and Time [$F(3, 273) = 30.80$, $p < .001$, $\eta^2_G = .06$]. These main effects were qualified by a significant Drug by Time interaction [$F(3, 273 = 21.55$, $p < .001$, $\eta^2_G = .04$]. Post-hoc analyses indicated that T concentrations were higher after Androgel® compared to placebo at blood sample 2 [$t(91) = 4.63$, $p < .001$, Cohen's $D = .97$], blood sample 3 [$t(91) = 4.59$, $p < .001$, Cohen's $D = .96$], and blood sample 4 [$t(91) = 4.36$, $p < .001$, Cohen's $D = .91$]. Importantly, blood sample 4 occurred directly before the femininity preferences task in order to confirm that blood serum T levels were in fact elevated prior to completing the task. Overall, participants in the AndroGel® condition experienced an average increase in T of 52.61% from baseline to 180 mins after administration, which is a relatively large effect size (Cohen's $D = 1.27$). As expected, there were no differences in T concentrations for Androgel® versus placebo prior to gel application [$t(91) = .52$, $p = .60$, Cohen's $D = .11$] (See Figure 4).

Femininity Preferences

One sample t-tests comparing the number of times the feminine versions of the female faces were chosen against the chance value of 10 revealed that participants chose the feminine face as more attractive across both drugs and contexts [T/Short-term ($t(47) = 17.74$, $p < .001$), T/Long-term ($t(44) = 11.64$, $p < .001$), P/Short-term ($t(44) = 14.01$, $p < .001$), P/Long-term ($t(44) = 14.28$, $p < .001$)].

A 2-Drug by 2-Context mixed measures ANOVA [between-subject factor: Drug (T versus Placebo); within-subject factor: Context (Short-Term versus Long-Term)] was conducted to test for differences in the number of trials in which the feminine face was selected as more attractive, as a function of drug and relationship context. Results revealed no main effect of Drug [$F(1, 91) = 1.48$, $p = .228$, $\eta^2_G = .014$], but a main effect of Context [$F(1, 91) = 4.48$, $p = .037$, $\eta^2_G = .007$], whereby participants demonstrated a stronger preference for feminine faces in the short-term relative to long-term mating context. This main effect was qualified by a significant Drug x Context interaction [$F(1, 91) = 9.89$, $p = .002$, $\eta^2_G = .014$]. Post-hoc analyses revealed that after T, participants showed a significantly higher preference for facial femininity in the short-term ($M = 16.77$, $SE = .43$) versus the long-term context ($M = 15.52$, $SE = .48$; $t(91) = 3.78$, $p < .001$, Cohen's $D = .49$). In contrast, participants in the placebo condition did not show a stronger preference for facial femininity in the short-term ($M = 16.76$, $SE = .44$) versus the long-term context ($M = 17.00$, $SE = .49$; $t(91) = .71$, $p = .48$, Cohen's $D = -.13$). Further analyses indicated that participants receiving T demonstrated a weaker preference for facial femininity in the long-term context ($M = 15.46$, $SE = .46$) compared to participants receiving placebo ($M = 17.00$, $SE = .49$; $t(91) = 2.18$, $p = .033$, Cohen's $D = -.46$). In contrast, there was no difference in

preference for facial femininity in the short-term context after T relative to placebo [$t(91) = .03$, $p = .98$, Cohen's $D = .01$] (See Figure 5).

Results of Experiment 2 are consistent with the findings from Experiment 1 (for men who received P on the first test session and T on the second test session). Specifically, participants who received T significantly preferred feminine faces more in the short-term relative to the long-term context. Also, as in Experiment 1 (for men who received P on the first test session and T on the second test session), this effect was driven by a lower preference for feminine faces in the long-term context.

General Discussion

The experiments presented here are the first to test the causal effects of exogenous T on men's preferences for facial femininity across both short- and long-term mating contexts. In both Experiment 1 and 2, initial face preference analyses suggested that regardless of mating context, participants preferred the feminine faces significantly more than the masculine faces—an effect that aligns with previous findings showing that men indeed show a preference for feminine versus masculine faces (e.g., Jones et al., 2007; Welling et al., 2008, 2013). For main analyses, results revealed that participants showed a stronger preference for feminine faces in the short-term context versus the long-term context after T relative to placebo, although in Experiment 1, this effect was found only among men who receive P on the first test session and T on the second test session.

Although participants on T preferred feminine faces more in the short- than the long-term mating context, this effect appears to be driven by a smaller preference for feminine faces in the long-term context. This somewhat surprising finding requires consideration of a number of

factors for interpretation. Given that T may increase interest in uncommitted sex (e.g., Puts et al., 2015), it is possible that an acute rise in T makes men less attuned to women in committed, long-term contexts, and thus women's characteristics in this context could be less salient. In other words, the smaller preferences for feminine faces in the long-term context by men who received T could represent a lower level of general *interest* in long-term mating, rather than a lower preference for femininity, per se. Should this be the case, it might be expected that men's preferences for feminine faces would be closer to chance level (i.e., 10 out of 20 faces) in the long-term when they had received T; although femininity preferences in this case were indeed closer to chance, the results of one sample t-tests confirmed that the preferences for feminine faces were still *significantly* above chance, so this explanation may not tell the whole story.

Another possibility for the pattern of results is the presence of ceiling effects for T-treated men. Should participants have been given a greater range of morph percentages to rate, and at levels that make feminization less salient (e.g., 15% feminized, 30% feminized), results might have shown that men receiving T preferred feminine faces significantly more so in the short-term versus the long-term context, but also significantly more so than preferences identified for either context when on placebo. In other words, if there are indeed ceiling effects present, the results found in the present study might actually be underestimated.

A third, more speculative possibility, is that T influences sensitivity to infidelity cues. Recent investigations have shown that near peak fertility, women's faces are not only rated as more attractive by observers, but are also characteristically more feminine in appearance (Oberzaucher et al., 2012; Puts et al., 2013; Roberts et al., 2004). Feminine female faces, both as composites and real faces, are accurately rated by men as having a more unrestricted sociosexuality (i.e., more likely to pursue short-term relationships; Boothroyd et al., 2011;

Boothroyd et al., 2008; but see Campbell et al., 2009), and near ovulation, some women report a greater sex drive, as well as interest in, or fantasy about, extra-pair partners (Gangestad et al., 2002; Gangestad et al., 2010; Haselton & Gangestad, 2006; see Gangestad & Thornhill, 2008, for review). Further, men show an increase in jealous mate-guarding for women near ovulation (Gangestad et al., 2002), which is of evolutionary importance, given that ovulation reflects a period in which a man's reproductive probability may be compromised by partner defection (Buss & Haselton, 2005). Although speculative, when considered in the context of this evidence, the present findings that men on T show a significantly greater preference for feminine female faces in the short-term versus long-term contexts (but no difference in context preference when on placebo) could suggest the possibility that T may increase men's sensitivity to infidelity cues, perhaps triggering careful decision-making regarding the trade-off between a healthy partner with good genes (i.e., feminine faces), and personality/potential for defection (i.e., a more masculine face offering reduced likelihood of cheating) when it comes to selecting a long-term mate. The present design did not allow a direct test of this hypothesis, so this remains entirely speculative as one of many possibilities for the pattern observed. Future research will be needed to ascertain whether men's perceptions of putative signals of infidelity risk vary as a function of their own T levels, and whether any such effects are predictive of men's preferences for feminine facial femininity. Moreover, any such hypothesis would need to be contrasted with other likely alternatives, such as men on T feeling more interest in short-term mating, and as such, less interest in long-term mating.

Of note is that the effects of Experiment 2 are consistent with those of Experiment 1 (for men who received P on the first test session and T on the second test session), despite measurement of face preferences occurring 1 hour earlier in Experiment 2 (Experiment 2 =

~2h15min post gel application versus Experiment 1 = ~3h15 min post gel application). Previous T administration studies have assessed behavior around 3 to 4 hours after peak T concentrations occur (see Bos et al., 2013, for review). However, the current findings indicate that assessment of behavior at 2 hours or 3 hours post T administration yields similar results. Thus, it appears that T may exert relatively rapid effects on face perception. Whether assessment of behavior at earlier time points (e.g., 1 hour post administration) would reveal similar findings is an important question for future investigations. Research in animal models indicates that T can exert rapid, likely non-genomic effects on brain function and behavior (reviewed in Foradori, Weiser, & Handa, 2008), and therefore, the presence of effects in humans earlier than 2 hours post administration is a possibility for investigation.

The present findings show both similarities to, and differences from, previous work. For instance, Welling et al. (2013) found that following a video game contest in which the outcome was unknowingly predetermined, winners (but not losers) showed significantly greater facial femininity preferences in the short-term versus the long-term contexts. As previously mentioned, there is evidence that winners of competitions experience a rise in T relative to losers (Archer, 2006; Carré & Olmstead, 2015). Although not directly tested, if the findings in Welling et al. (2013) were mediated by competition-induced T dynamics, then the present finding that feminine face preference was significantly higher in the short- versus the long-term context when participants were on T (but not placebo) aligns with their findings. However, the present study differs from Welling et al. (2013) in that we did not find a difference between short-term femininity preferences on T Day, relative to short-term femininity preferences on placebo day. One potential explanation for this apparent discrepancy is that those winning the competition may have not only experienced increased T levels, but may have also experienced increased

perceptions of their own attractiveness or masculinity/dominance (Welling et al., 2013; Welling et al., 2016)—factors implicated in men's mating success (e.g., Rhodes et al., 2005). Indeed, recent evidence suggests that competition outcome can modulate self-perceptions on sexually-relevant dimensions (e.g., dominance: Watkins & Jones, 2012).

Welling et al. (2008) found that across two separate testing days, men showed greater preference for feminized versus masculinized faces on the day in which the men's T levels were higher. This study did not account for mating context (short- or long-term), which has now been shown to be an important consideration for men's facial preferences (e.g., Welling et al., 2013), and thus may partially account for the discussed differences. The present study provides further evidence to suggest that mating context is an important consideration for mate preference research—particularly with respect to hormonal influences—and thus should also be considered, where possible, in future studies.

Limitations and Future Directions

A number of important limitations to our study should be noted. First, the findings from Experiment 1 indicated that the Drug by Context interaction was only found for men who received P on the first day of testing, and T on the second day of testing. Experiment 2, which employed a between-subjects design, revealed a similar Drug by Context interaction. We speculated that the novelty of the research context, which may increase stress and uncertainty, may in part underlie the order effect found in Experiment 1. However, if this were the case, we should not have observed a similar Drug by Context interaction in Experiment 2, given that participants in Experiment 2 were also exposed to a novel experimental situation. Aside from differences in experimental design (within- vs. between-subjects), a few key differences between Experiment 1 and Experiment 2 are worth mentioning. First, Experiment 1 assessed face

preferences approximately 3 hours after drug administration, whereas Experiment 2 assessed face preferences approximately 2 hours after drug administration. In light of known rapid effects of steroid hormones on sexual perception in animal models (e.g., Lord, Bond & Thompson, 2009), it remains possible that more robust effects of T on preferences for facial femininity may occur at earlier time points after drug administration (as seen in Experiment 2). However, this would not explain why effects of T on femininity preferences were found for men in Experiment 1 who received T on the second day of testing. Another important methodological difference is that participants in Experiment 1 were tested in a novel environment (Urology Clinic) for the first 2 hours of testing, and then tested in another novel laboratory environment (University setting) for the remainder of the experiment. In Experiment 2, participants were tested in the same location for the duration of the study. Thus, perhaps the additional novelty in Experiment 1 may have enhanced stress, which in turn could have blunted the effects of T on femininity preferences for participants who received T on the first day of testing. Habituation to the experimental procedures may have dampened stress levels, allowing T to exert an influence on femininity preferences. Although we did not assess objective indices of stress (e.g., cortisol), correlational work suggests that individual differences in T concentrations positively predict numerous outcome measures (e.g., dominance, risk-taking), but only among individual with relatively low cortisol concentrations (Mehta & Josephs, 2010; Mehta et al., 2015). This explanation is unlikely, however, given that order of drug administration did not moderate effects of exogenous T on other socio-cognitive processes assessed using the same dataset (Arnocky et al., 2016; Carré et al., 2015; Welling et al., 2016). Clearly, more research will be needed to determine the extent to which order of drug administration plays a key role in modulating facial femininity preferences and the factor(s) that may underlie such order effects.

Another limitation of the current experiments is that we used a forced-choice paradigm whereby participants chose either a masculinized or a feminized version of a woman's face as their preference for either a short- or long-term partner. However, it is conceivable that some participants would prefer neither face if given the choice. The option to select neither face could provide more insight into the potential for T to reduce general interest for mating in long-term contexts (e.g., if participants on T more often than not selected "neither" as their preference rather than a lower preference for femininity, it would provide some support for the view that T decreases general interest in long-term mating). In a similar respect, participants were not asked any questions about motivation for short- or long-term relationships, or given other tasks that address the broader topic of sexual motivation. To establish the mechanisms underlying the effects of the present study, future investigations will need to employ a wider variety of tasks, including those that measure participant interest and motivation to engage in each type of mating (short- or long-term partnerships), in addition to their preferences for facial femininity across these contexts. Further, while we replicate findings across two independent studies in the present investigation, the degree to which small differences in facial femininity preferences for long-term mates translates into observable behavior remains unclear; this may prove an important consideration for future work, in addition to the potential functional role(s) of hormone-mediated facial preference differences for long-term mates discussed above.

Although Experiment 1 had a relatively small sample size, it is consistent with other single T-administration studies conducted in women examining effects on social, cognitive, and behavioral processes (see Bos et al., 2013, for review); further, the within-subject design makes it a more powerful test of intra-individual variation in facial preferences. Experiment 2 bolstered these findings using a between-subjects design with a larger sample size of 93 men. Future T-

administration studies may consider larger sample sizes in order to examine the extent to which individual difference factors moderate effects of T on preferences for facial femininity. For instance, previous work indicates that people scoring relatively high on sensation seeking demonstrate greater preferences for feminine faces (Jones et al., 2007). Similarly, males rating themselves relatively high on attractiveness also demonstrate a stronger preference for feminine faces (Burriss et al., 2011). Therefore, T may have a strong effect on preferences for female facial femininity, but this may be reserved for those individuals who perceive themselves to be highly attractive, or who are high in trait levels of sensation seeking.

Conclusion

The findings from both Experiment 1 and Experiment 2 provide the first evidence that a single administration of T can rapidly modulate preferences for female facial femininity in a mating-context dependent manner. Across studies, men showed a decreased preference for feminine female faces in the long-term relative to short-term context when they were administered T, but this difference was not present when they were administered placebo. While the study design prevented the confident identification of underlying psychological mechanisms, future studies can seek to extend the present findings and further contrast potential influences of T on general mating motivation across short- and long-term contexts.

Acknowledgements

The authors wish to thank Brandy Labine, Nathan Olmstead, Erika Ruddick, Sara Taylor and Zachary Root for their help with data collection. We also thank Mrs. Terri Moreau (Exp #1) and Medicor Research Inc (Exp #2) for overseeing the medical components of the protocol (e.g.,

blood draws and gel application). Finally, we also thank Algonquin Pharmasave for help with preparing Androgel® and Placebo gels (in particular, Meghan Heintz and Jen Zou).

Funding Statement

This project was supported by a Natural Sciences and Engineering Research Council (NSERC) Canada Graduate Scholarship (CGS-M) to B.M.B, as well as an NSERC Discovery Grant (to J.M.C) and a Northern Ontario Heritage Fund Corporation (NOHFC) grant (to J.M.C.).

References

Archer, J. (2006). Testosterone and human aggression: an evaluation of the challenge hypothesis. *Neuroscience & Biobehavioral Reviews*, *30*(3), 319-345. http://dx.doi.org/10.1016/j.neubiorev.2004.12.007

Arnocky, S., Taylor, S.M., Olmstead, N.A., & Carré, J.M. (2016). The effects of exogenous testosterone on men's moral decision-making. *Adaptive Human Behavior and Physiology*, 1–13. doi: 10.1007/s40750-016-0046-8

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379–384. doi: 10.3758/BF03192707

Bobst, C., Sauter, S., Foppa, A., & Lobmaier, J.S. (2014). Early follicular testosterone level predicts preference for masculinity in male faces – but not for women taking hormonal contraception. *Psychoneuroendocrinology, 41*, 142 – 150. doi: 10.1016/j.psyneuen.2013.12.012

Boothroyd, L. G., Cross, C. P., Gray, A. W., Coombes, C., & Gregson-Curtis, K. (2011). Perceiving the facial correlates of sociosexuality: Further evidence. *Personality and Individual Differences*, *50*(3), 422-425. doi:10.1016/j.paid.2010.10.017

Boothroyd, L. G., Jones, B. C., Burt, D. M., DeBruine, L. M., & Perrett, D. I. (2008). Facial correlates of sociosexuality. *Evolution and Human Behavior*, *29*(3), 211-218. doi: 10.1016/j.evolhumbehav.2007.12.009

Bos, P. A., Panksepp, J., Bluthé, R., & Honk, J. v. (2012). Acute effects of steroid hormones and neuropeptides on human social–emotional behavior: A review of single administration studies. *Frontiers in Neuroendocrinology, 33*(1), 17-35. doi:10.1016/j.yfrne.2011.01.002

Brewer, G., & Archer, J. (2007). What do people infer from facial attractiveness? *Journal of Evolutionary Psychology*, *5*(1), 39-49. doi: 10.1556/JEP.2007.1002

Burriss, R.P., Welling, L.L.M., Puts, D.A. (2011). Men's attractiveness predicts their preferences for facial femininity when judging for short-term, but not long-term, partners. *Personality and Individual Differences*, *50*(5), 542-546. doi: 10.1016/j.paid.2010.11.022

Buss, D. M., & Haselton, M. (2005). The evolution of jealousy. *Trends in Cognitive Sciences*, *9*(11), 506-507.

Campbell, L., Cronk, L., Simpson, J. A., Milroy, A., Wilson, C. L., & Dunham, B. (2009). The association between men's ratings of women as desirable long-term mates and individual differences in women's sexual attitudes and behaviors. *Personality and Individual Differences*, *46*(4), 509–513.

Carré, J. M., & Olmstead, N. A. (2015). Social neuroendocrinology of human aggression: Examining the role of competition-induced testosterone dynamics. *Neuroscience*, *286*, 171–186. http://dx.doi.org/10.1016/j.neuroscience.2014.11.029

Carré, J.M., Ortiz, T.L., Labine, B., Moreau, B.J., Viding, E., Neumann, C.S., & Goldfarb, B. (2015). Digit ratio (2D:4D) and psychopathic traits moderate the effect of exogenous testosterone and socio-cognitive processes in men. *Psychoneuroendocrinology*, *62*, 319–326. doi: 10.1016/j.psyneuen.2015.08.023

Carré, J. M., Geniole, S. N., Ortiz, T. L., Bird, B. M., Videto, A., & Bonin, P. L. (in press). Exogenous testosterone rapidly increases aggressive behavior in dominant and impulsive men. *Biological Psychiatry*. doi: 10.1016/j.biopsych.2016.06.009

Carré, J. M., Putnam, S. K., & McCormick, C. M. (2009). Testosterone responses to competition predict future aggressive behaviour at a cost to reward in men. *Psychoneuroendocrinology*, *34*(4), 561-570. doi: 10.1016/j.psyneuen.2008.10.018

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.

DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., ... & Tiddeman, B. P. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1592), 1355-1360. doi: 10.1098/rspb.2005.3445

Eisenegger, C., von Eckardstein, A., Fehr, E., & von Eckardstein, S. (2013). Pharmacokinetics of testosterone and estradiol gel preparations in healthy young men. *Psychoneuroendocrinology*, *38*(2), 171-178. http://dx.doi.org/10.1016/j.psyneuen.2012.05.018

Fink, B., Klappauf, D., Brewer, G., & Shackelford, T. K. (2014). Female physical characteristics and intra-sexual competition in women. *Personality and Individual Differences*, *58*, 138-141. doi: 10.1016/j.paid.2013.10.015

Foradori, C. D., Weiser, M. J., & Handa, R. J. (2008). Non-genomic actions of androgens. *Frontiers in Neuroendocrinology, 29*(2), 169-181. doi:10.1016/j.yfrne.2007.10.005

Frederick, D. A., & Haselton, M. G. (2007). Why is muscularity sexy? Tests of the fitness indicator hypothesis. *Personality and Social Psychology Bulletin*, *33*(8), 1167-1183. doi:10.1177/0146167207303022

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, *141*(1), 2–18. http://dx.doi.org/10.1037/a0024338

Gangestad, S. W., & Scheyd, G. J. (2005). The Evolution of Human Physical Attractiveness. *Annual Review of Anthropology*, *34*523-548. doi:10.1146/annurev.anthro.33.070203.143733

Gangestad, S. W., Simpson, J. A., Cousins, A. J., Garver-Apgar, C. E., & Christensen, P. N. (2004). Women's Preferences for Male Behavioral Displays Change Across the Menstrual Cycle. *Psychological Science*, *15*(3), 203-206. doi:10.1111/j.0956-7976.2004.01503010.x

Gangestad, S. W., Thornhill, R., & Garver-Apgar, C. E. (2010). Fertility in the cycle predicts women's interest in sexual opportunism. *Evolution and Human Behavior*, *31*(6), 400-411. doi:10.1016/j.evolhumbehav.2010.05.003

Gangestad, S. W., Thornhill, R., & Garver, C. E. (2002). Changes in women's sexual interests and their partner's mate–retention tactics across the menstrual cycle: evidence for shifting conflicts of interest. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *269*(1494), 975-982. doi: 10.1098/rspb.2001.1952

Gangestad, S. W., & Thornhill, R. (2008). Human oestrus. *Proceedings of the Royal Society of London B: Biological Sciences*, *275*(1638), 991-1000. doi: 10.1098/rspb.2007.1425

Geniole, S.N., MacDonell, E.T., & McCormick, C.M. (in press). The point subtraction aggression paradigm as a laboratory tool for investigating the neuroendocrinology of aggression and competition. *Hormones and Behavior.*

Gildersleeve, K., Haselton, M. G., & Fales, M. R. (2014). Do women's mate preferences change across the ovulatory cycle? A meta-analytic review. *Psychological Bulletin, 140*(5), 1205-1259. doi:10.1037/a0035438

Goetz, S. M., Tang, L., Thomason, M. E., Diamond, M. P., Hariri, A. R., & Carré, J. M. (2014). Testosterone rapidly increases neural reactivity to threat in healthy men: a novel two-step pharmacological challenge paradigm. *Biological Psychiatry, 76*(4), 324-331. http://dx.doi.org/10.1016/j.biopsych.2014.01.016

Gray, A. W., & Boothroyd, L. G. (2012). Female facial appearance and health. *Evolutionary Psychology, 10*(1), 66-77.

Haselton, M. G., & Gangestad, S. W. (2006). Conditional expression of women's desires and men's mate guarding across the ovulatory cycle. *Hormones and Behavior, 49*(4), 509-518. doi: 10.1016/j.yhbeh.2005.10.006

Henderson, J. J., & Anglin, J. M. (2003). Facial attractiveness predicts longevity. *Evolution and Human Behavior, 24*(5), 351–356. http://dx.doi.org/10.1016/S1090-5138(03)00036-9

Johnston, V. S., Hagel, R., Franklin, M., Fink, B., & Grammer, K. (2001). Male facial attractiveness: Evidence for hormone-mediated adaptive design. *Evolution and Human Behavior, 22*(4), 251-267. doi:10.1016/S1090-5138(01)00066-6

Jokela, M. (2009). Physical attractiveness and reproductive success in humans: evidence from the late 20th century United States. *Evolution and Human Behavior, 30*(5), 342–350. http://dx.doi.org/10.1016/j.evolhumbehav.2009.03.006

Jones, B. C., DeBruine, L. M., Little, A. C., Conway, C. A., Welling, L. L. M., and Smith, F. (2007). Sensation seeking and men's face preferences. *Evolution and Human Behavior, 28*, 439-446. doi: 10.1016/j.evolhumbehav.2007.07.006

Jones, B. C., Hahn, A. C., Fischer, C. I., Wincenciak, J., Kandrik, M., Roberts, S. C., Little, A. C., & DeBruine, L. M. (2015). Facial coloration tracks changes in women's estradiol. *Psychoneuroendocrinology, 56*, 29-34. Doi: 10.1016/j.psyneuen.2015.02.021

Jones, B. C., Little, A. C., Boothroyd, L., DeBruine, L. M., Feinberg, D. R., Smith, M. L., ... & Perrett, D. I. (2005). Commitment to relationships and preferences for femininity and apparent health in faces are strongest on days of the menstrual cycle when progesterone level is high. *Hormones and Behavior, 48*(3), 283-290. doi:10.1016/j.yhbeh.2005.03.010

Komori, M., Kawamura, S., & Ishihara, S. (2009). Effect of averageness and sexual dimorphism on the judgment of facial attractiveness. *Vision Research, 49*(8), 862-869. doi:10.1016/j.visres.2009.03.005

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*(3), 390–423.

Law Smith, M. J., Perrett, D. I., Jones, B. C., Cornwell, R. E., Moore, F. R., Feinberg, D. R., ... & Hillier, S. G. (2006). Facial appearance is a cue to oestrogen levels in women. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1583), 135-140. doi: 10.1098/rspb.2005.3296

Lee, A. J., Dubbs, S. L., Kelly, A. J., von Hippel, W., Brooks, R. C., & Zietsch, B. P. (2013). Human facial attributes, but not perceived intelligence, are used as cues of health and resource provision potential. *Behavioral Ecology, 24*(3), 779-787. doi:10.1093/beheco/ars199

Lejuez, C.W., Read, J.P., Kahler, C.W., Richards, J.B., Ramsey, S.E., Stuart, G.L., … & Brown, R.A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue

Risk Task (BART). *Journal of Experimental Psychology: Applied, 8*(2), 75–84. doi: 10.1037//1076-898X.8.2.75

Little, A. C. (2014). Facial attractiveness. *WIREs Cognitive Science, 5*, 621–634. doi: 10.1002/wcs.1316

Little, A. C., Connely, J., Feinberg, D. R., Jones, B. C. & Roberts, S. C. (2011). Human preference for masculinity differs according to context in faces, bodies, voices, and smell. *Behavioral Ecology*, *22*(4), 862-868.

Little, A. C., & Jones, B. C. (2012). Variation in facial masculinity and symmetry preferences across the menstrual cycle is moderated by relationship context. *Psychoneuroendocrinology*, *37*(7), 999-1008. doi:10.1016/j.psyneuen.2011.11.007

Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: evolutionary based research. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1571), 1638-1659. doi: 10.1098/rstb.2010.0404

Little, A. C., Jones, B. C., Feinberg, D. R., & Perrett, D. I. (2013). Men's strategic preferences for femininity in female faces. *British Journal of Psychology*. doi: 10.1111/bjop.12043, decision times, and emotional salience. *Journal of Behavioral Decision Making, 27*(1), 57–65. doi: 10.1002/bdm.1782

Lord, L. D., Bond, J., & Thompson, R. R. (2009). Rapid steroid influences on visually guided sexual behavior in male goldfish. *Hormones and behavior*, *56*(5), 519–526. doi:10.1016/j.yhbeh.2009.09.002

Lotto, L., Manfrinati, A., & Sarlo, M. (2014). A new set of moral dilemmas: Norms for moral acceptability

Mazur, A., & Booth, A. (1998). Testosterone and dominance in men. *Behavioral and Brain*

Sciences, *21*(03), 353–363.

Mehta, P. H., & Josephs, R. A. (2006). Testosterone change after losing predicts the decision to compete again. *Hormones and Behavior*, *50*(5), 684-692. doi:10.1016/j.yhbeh.2006.07.001

Mehta, P. H., & Josephs, R. A. (2010). Testosterone and cortisol jointly regulate dominance: Evidence for a dual-hormone hypothesis. *Hormones and Behavior*, *58*(5), 898–906. doi:10.1016/j.yhbeh.2010.08.020

Mehta, P. H., Welker, K. M., Zilioli, S., & Carré, J. M. (2015). Testosterone and cortisol jointly modulate risk-taking. *Psychoneuroendocrinology*, *56*, 88–99. doi:10.1016/j.psyneuen.2015.02.023

Miller, S. L., Maner, J. K., & McNulty, J. K. (2012). Adaptive attunement to the sex of individuals at a competition: The ratio of opposite- to same-sex individuals correlates with changes in competitors' testosterone levels. *Evolution and Human Behavior*, *33*(1), 57-63. doi:10.1016/j.evolhumbehav.2011.05.006

Oberzaucher, E., Katina, S., Schmehl, S. F., Holzleitner, I. J., Mehu-Blantar, I., & Grammer, K. (2012). The myth of hidden ovulation: Shape and texture changes in the face during the menstrual cycle. *Journal of Evolutionary Psychology*, *10*(4), 163-175. doi: http://dx.doi.org/10.1556/JEP.10.2012.4.1

O'Connor, J. J., Fraccaro, P. J., Pisanski, K., Tigue, C. C., & Feinberg, D. R. (2013). Men's preferences for women's femininity in dynamic cross-modal Stimuli. *PloS one*, *8*(7), e69531.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychological Methods*, *8*(4), 434–447.

http://dx.doi.org/10.1037/1082-989X.8.4.434

Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K., & Minamisawa, R. (1999). Menstrual cycle alters face preference. *Nature*, *399*(6738), 741-742. doi:10.1038/21557

Penton-Voak, I. S., & Perrett, D. I. (2000). Female preference for male faces changes cyclically: Further evidence. *Evolution and Human Behavior, 21*(1), 39-48. doi:10.1016/S1090-5138(99)00033-1

Peters, M., Simmons, L. W., & Rhodes, G. (2008). Testosterone is associated with mating success but not attractiveness or masculinity in human males. *Animal Behaviour, 76*(2), 297-303. doi:10.1016/j.anbehav.2008.02.008

Pollet, T. V., der Meij, L., Cobey, K. D., & Buunk, A. P. (2011). Testosterone levels and their associations with lifetime number of opposite sex partners and remarriage in a large sample of American elderly men and women. *Hormones and Behavior, 60*(1), 72-77. doi:10.1016/j.yhbeh.2011.03.005

Puts, D. A., Bailey, D. H., Cárdenas, R. A., Burriss, R. P., Welling, L. L., Wheatley, J. R., & Dawood, K. (2013). Women's attractiveness changes with estradiol and progesterone across the ovulatory cycle. *Hormones and Behavior, 63*(1), 13-19. http://dx.doi.org/10.1016/j.yhbeh.2012.11.007

Puts, D.A., Pope, L.E., Hill, A.K., Cardenas, R.A., Welling, L.L.M., Wheatley, J.R., & Breedlove, M.S. (2015) Fulfilling desire: evidence for negative feedback between men's testosterone, sociosexual psychology, and sexual partner number. *Hormones and Behavior, 70,* 14–21. doi: S0018-506X(15)00014-8 [pii] 10.1016/j.yhbeh.2015.01.006

Roberts, S. C., Havlicek, J., Flegr, J., Hruskova, M., Little, A. C., Jones, B. C., ... & Petrie, M.

(2004). Female facial attractiveness increases during the fertile phase of the menstrual cycle. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 271*(Suppl 5), S270-S272. doi: 10.1098/rsbl.2004.0174

Röder, S., Fink, B., & Jones, B. C. (2013). Facial, olfactory, and vocal cues to female reproductive value. *Evolutionary Psychology*, *11*(2), 392-404.

Rowland, D. A., & Perrett, D. I. (1995). Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications, 15*(5), 70-76. doi:10.1109/38.403830

Slatcher, R. B., Mehta, P. H., & Josephs, R. A. (2011). Testosterone and self-reported dominance interact to influence human mating behavior. *Social Psychological and Personality Science*, *2*(5), 531-539. doi:10.1177/1948550611400099

Thornhill, R., & Gangestad, S. W. (2006). Facial sexual dimorphism, developmental stability, and susceptibility to disease in men and women. *Evolution and Human Behavior, 27*(2), 131–144. doi: 10.1016/j.evolhumbehav.2005.06.001

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications, 21*(5), 42-50. doi:10.1109/38.946630

Van Anders, S. M. (2010). Gonadal steroids and salivary IgA in healthy young women and men. *American Journal of Human Biology*, *22*(3), 348-352. doi: 10.1002/ajhb.20997

van Anders, S. M., Hamilton, L. D., & Watson, N. V. (2007). Multiple partners are associated with higher testosterone in North American men and women. *Hormones and Behavior, 51*(3), 454-459. doi:10.1016/j.yhbeh.2007.01.002

van Wingen, G. A., Zylicz, S. A., Pieters, S., Mattern, C., Verkes, R. J., Buitelaar, J. K., & Fernández, G. (2009). Testosterone increases amygdala reactivity in middle-aged women

to a young adulthood level. *Neuropsychopharmacology*, *34*(3), 539-547. doi:10.1038/npp.2008.2

Watkins & Jones (2012). Priming men with different contest outcomes modulates their dominance perceptions. *Behavioral Ecology.* doi:10.1093/beheco/arr221

Welling, L. M., Jones, B. C., DeBruine, L. M., Conway, C. A., Law Smith, M. J., Little, A. C., & ... Al-Dujaili, E. S. (2007). Raised salivary testosterone in women is associated with increased attraction to masculine faces. *Hormones and Behavior*, *52*(2), 156-16. doi:10.1016/j.yhbeh.2007.01.010

Welling, L. L., Jones, B. C., DeBruine, L. M., Smith, F. G., Feinberg, D. R., Little, A. C., & Al-Dujaili, E. A. (2008). Men report stronger attraction to femininity in women's faces when their testosterone levels are high. *Hormones and Behavior*, *54*(5), 703-708. doi: 10.1016/j.yhbeh.2008.07.012

Welling, L. L., Moreau, B. J., Bird, B. M., Hansen, S., & Carré, J. M. (2016). Exogenous testosterone increases men's perceptions of their own physical dominance. *Psychoneuroendocrinology*, *64*, 136–142. http://dx.doi.org/10.1016/j.psyneuen.2015.11.016

Welling, L. L., Persola, L., Wheatley, J. R., Cárdenas, R. A., & Puts, D. A. (2013). Competition and men's face preferences. *Personality and Individual Differences*, *54*(3), 414-419. doi:10.1016/j.paid.2012.10.014

Wheatley, J. R., Apicella, C. A., Burriss, R. P., Cárdenas, R. A., Bailey, D. H., Welling, L. L., & Puts, D. A. (2014). Women's faces and voices are cues to reproductive potential in industrial and forager societies. *Evolution and Human Behavior*. doi: 10.1016/j.evolhumbehav.2014.02.006

Figure Captions:

Figure 1. Example of feminized (left) and masculinized stimuli.

Figure 2. Testosterone concentrations as a function of drug condition in Experiment 1. * $p <$ .001 for difference between Androgel® and Placebo conditions.

Figure 3. Frequency of feminine faces selected across all trials as a function of drug, context, and order of drug administration. A significant drug by context interaction was found among individuals who received placebo on their first day and testosterone on their second day (shown left), but not for those who received testosterone on their first day and placebo on their second day (shown right). * $p <$.01 ** $p <$ .05

Figure 4. Testosterone concentrations as a function of drug condition in Experiment 2. * $p <$ .001 for difference between Androgel® and Placebo conditions.

Figure 5. Frequency of feminine faces selected across all trials as a function of drug and context in Experiment 2. Error bars represent the *SE*. * $p <$.01 ** $p <$ .05

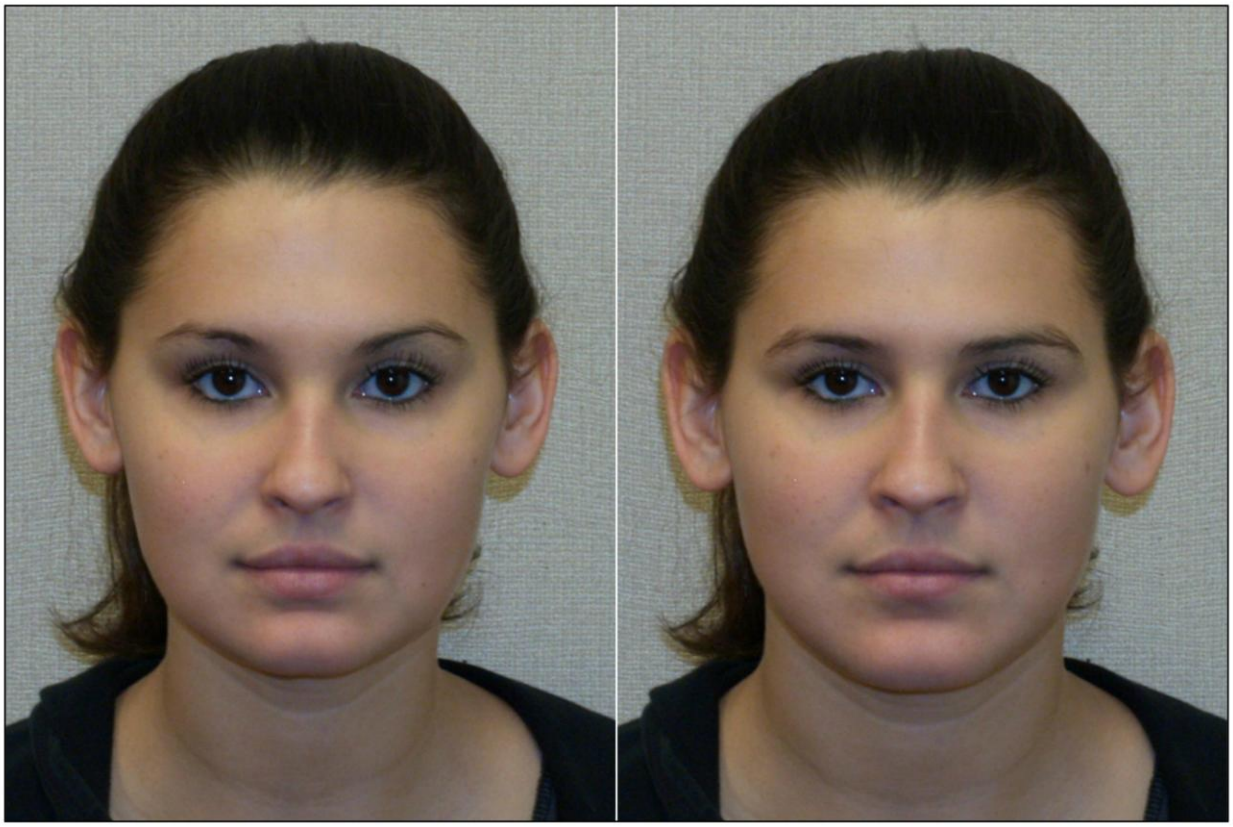

Fig. 1

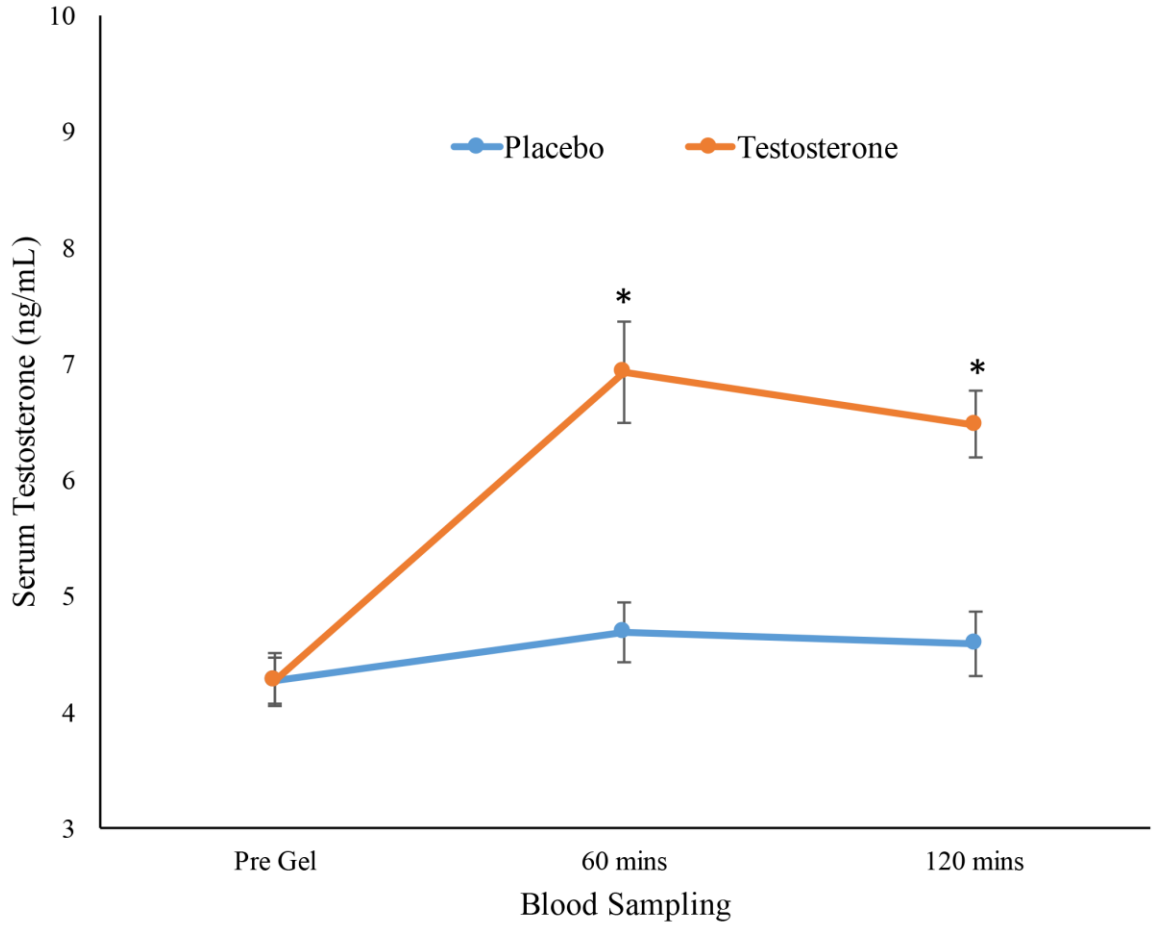

Fig. 2

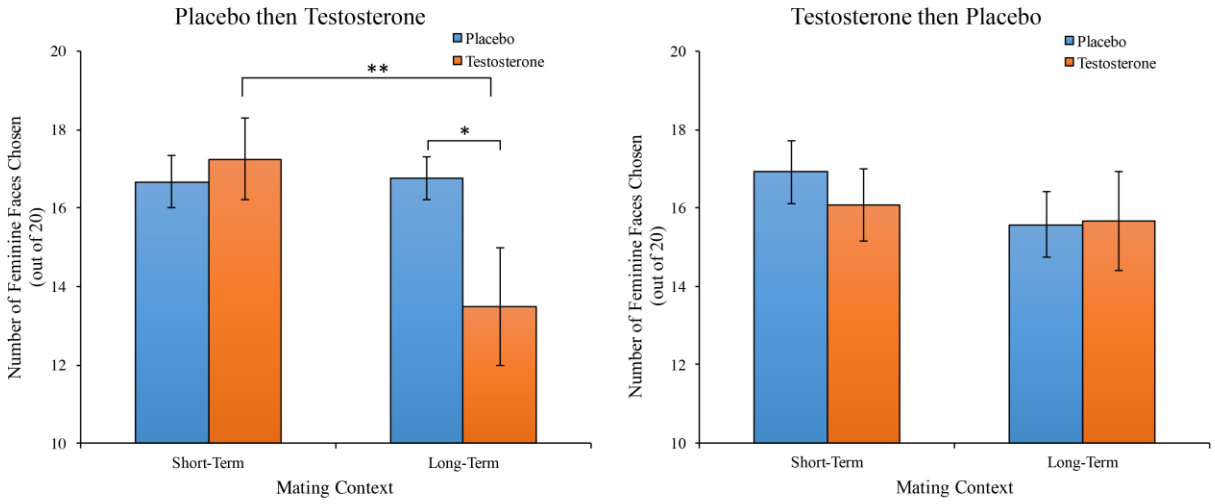

Fig. 3

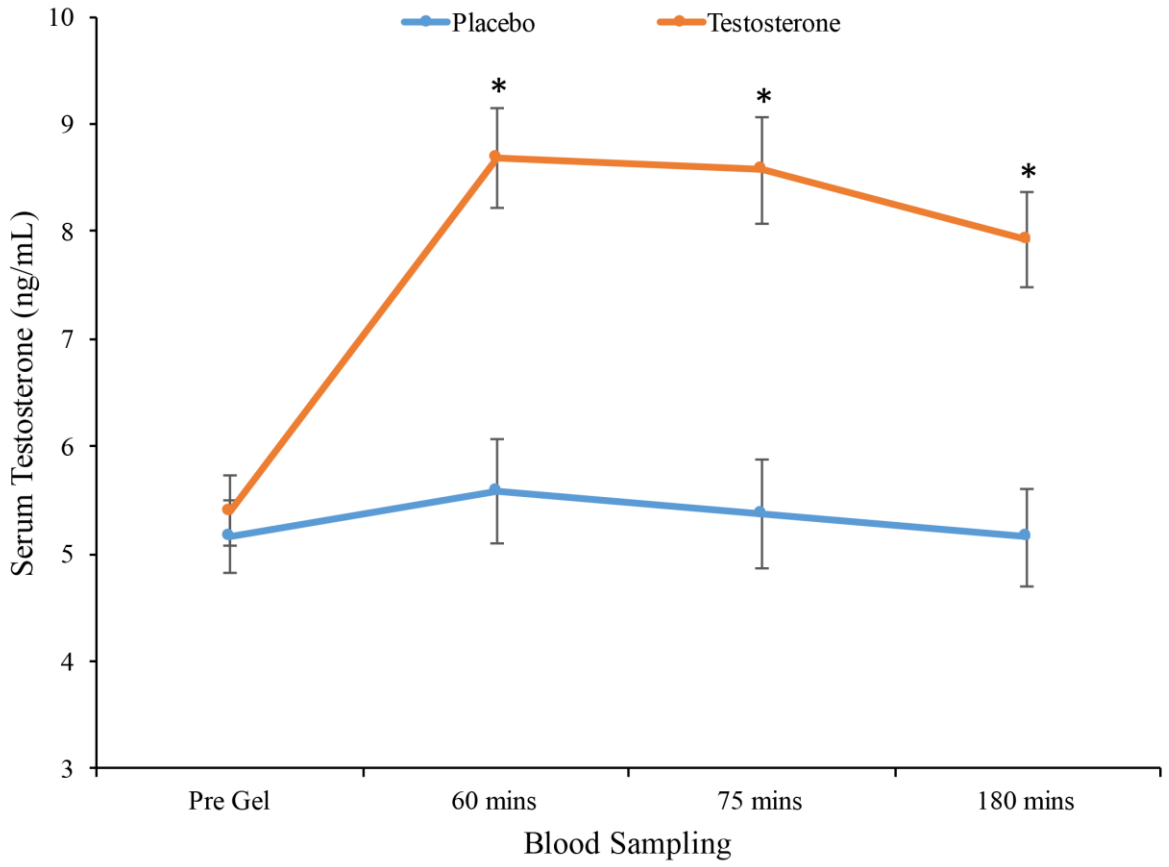

Fig. 4

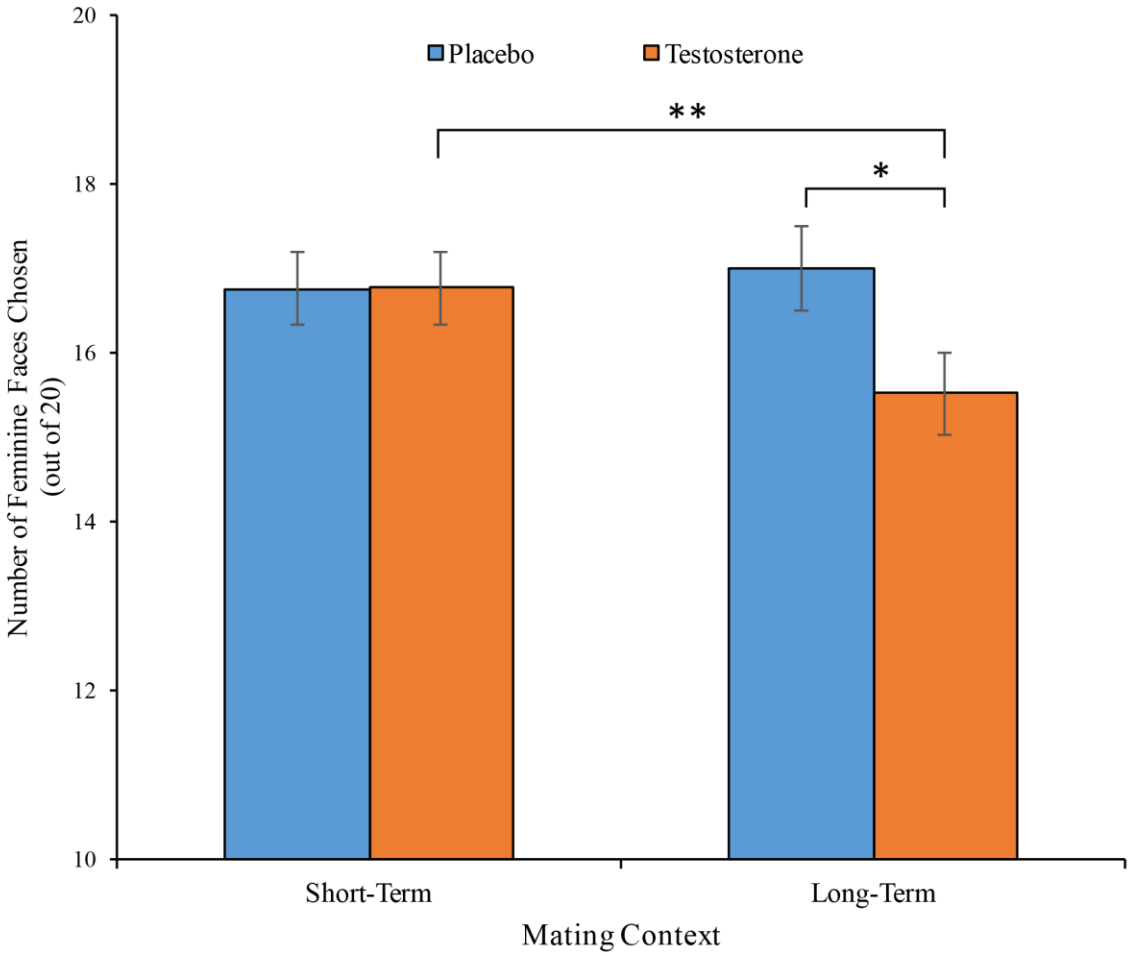

Fig. 5

Highlights

-Men demonstrate a stronger preference for feminine female faces for short-term versus long-term mating contexts.

-In Exp 1 (n = 24, within-subject, cross-over design), a single dose of testosterone (T) caused an increase in preferences for feminine female faces in short-term versus long-term mating context – but this effect was only found for men who received placebo on the first test session, and T on the second test session

-In Exp 2 (n = 93, between-subject design), a single dose of T caused an increase in preferences for feminine female faces in short-term versus long-term mating contexts.

-These are the first studies to investigate the causal role of T in modulating men's preferences for feminine female faces across different mating contexts.