Title

Meta-analytic and Experimental Investigations of Shifts in Women's Mate Preferences and Attractiveness across the Ovulatory Cycle

Permalink

https://escholarship.org/uc/item/0qb2g751

Author

Gildersleeve, Kelly

Publication Date

2014

UNIVERSITY OF CALIFORNIA

Los Angeles

Meta-analytic and Experimental Investigations of Shifts in Women's Mate Preferences and Attractiveness across the Ovulatory Cycle

A Dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Psychology

by

Kelly Ann Gildersleeve

2014

ABSTRACT OF THE DISSERTATION

Meta-analytic and Experimental Investigations of Shifts in Women's Mate Preferences and Attractiveness across the Ovulatory Cycle

by

Kelly Ann Gildersleeve

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2014

Professor Martie G. Haselton, Chair

The high-fertility period of the ovulatory cycle is the only time when sex can result in conception. In nonhuman mammals, this period is often marked by dramatic changes in females' social interactions, especially with males. For example, two widespread patterns are that females exhibit increased sexual interest in certain males at high fertility relative to low fertility, and males exhibit increased sexual interest in females currently exhibiting cues of high fertility. Scientific interest in whether the ovulatory cycle has similar impacts on human social behavior has surged in the past two decades, producing a large body of published evidence that is largely supportive of this possibility. Two prominent findings in this literature are that women's sexual attraction to men possessing characteristics historically associated with high genetic quality increases at high relative to low fertility and that women's attractiveness to men also increases at this time. However, studies have varied widely in the methods they have used to examine these effects, findings have been somewhat mixed, and additional nonreplications could remain unpublished. In addition, several important questions have yet to be answered. For example, can women detect cues of high fertility in other women, and do they perceive these cues as attractive?

To address these questions, I conducted two meta-analyses and a laboratory study. The meta-analyses quantitatively evaluated and synthesized evidence across published and unpublished findings for a) changes in women's mate preferences and b) changes in women's attractiveness and other possible fertility cues across the ovulatory cycle. Both analyses revealed robust support in the extant empirical literature for the hypothesized cycle shifts. The laboratory study examined women's perceptions of other women's high- and low-fertility body scents and found that, like men, women perceive other women's high-fertility body scents as more attractive than their low-fertility scents. Taken together, these studies have potentially important implications for understanding the role of the ovulatory cycle—and psychological adaptations that evolved or have been maintained in the context of cyclic variation in human female fertility—in human social cognition, motivation, and behavior.

The dissertation of Kelly Ann Gildersleeve is approved.

Daniel M. T. Fessler

Kerri L. Johnson

Ted F. Robles

Martie G. Haselton, Committee Chair

University of California, Los Angeles

2014

TABLE OF CONTENTS

VITA

EDUCATION

2006 B.A., Psychology, Minor: Art
Alma College, Alma, Michigan
Summa Cum Laude and Honors in Psychology

AWARDS AND HONORS

2013	Human Behavior and Evolution Society New Investigator Award
2013	UCLA Harold H. Kelley Award: Best Social Psychology Research Paper
2012	UCLA Graduate Division Dissertation Year Fellowship
2010	UCLA Graduate Research Mentorship
2009	NSF Graduate Research Fellowship, Honorable Mention
2008 - 2010	NSF-IGERT Interdisciplinary Relationship Science Program Traineeship
2007	NSF-IGERT Interdisciplinary Relationship Science Program Recruitment Award
2007	UCLA University Fellowship
2002 - 2006	Alma College Dean's List
2003	Alma College Distinguished Trustee Award
2002	Alma College Trustee Honors Scholarship

RESEARCH FUNDING

2013 UCLA Academic Senate Grant, *Examining links between intimate relationship quality and fertility in humans* (co-written with Dr. Martie Haselton). Examines links between women's relationships, hormones, and fertility. Funded hormone assays ($10,000).

2011 Grant from American Institute for Bisexuality, *Women's Attraction and Hormone Responses to Other Women's Body Scents* (co-written with Dr. Martie Haselton). Examines lesbian, bisexual, and heterosexual women's responses to scent cues of high fertility. Funded lab and field studies, hormone assays, and genotyping ($90,000).

2008 NSF-IGERT Interdisciplinary Relationship Science Program Research Grant, *Scent Cues of Ovulation* ($1,000)

2008 UCLA Psychology Department Master’s Thesis Research Grant, *Scent Cues of Ovulation* ($800)

PUBLICATIONS

Fales, M.R., Gildersleeve, K., & Haselton, M.G. (*in press*). Exposure to perceived male rivals raises men’s testosterone on fertile relative to non-fertile days of their partner’s ovulatory cycle. *Hormones and Behavior.*

Lukaszewski, A.W., Larson, C.M., Gildersleeve, K., Roney, J.R., & Haselton, M.G. (*in press*). Condition-dependent physical features predict mating strategies and past sexual behavior: Evidence from multiple independent samples. *Evolution and Human 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*. Advance online publication.

* Human Behavior and Evolution Society New Investigator Award, 2013

* UCLA Harold H. Kelley Award for Best Social Psychology Basic Research Paper, 2013

Gildersleeve, K., DeBruine, L., Haselton, M.G., Frederick, D.A., Penton-Voak, I.S., Jones, B.C., & Perrett, D.I. (2013). Shifts in women’s mate preferences across the cycle: A critique of Harris (2011) and Harris (2012). *Sex Roles, 69,* 525 – 535.

Larson, C.M., Haselton, M.G., Gildersleeve, K., & Pillsworth, E.P. (2012). Changes in women’s feelings about their romantic relationships across the ovulatory cycle. *Hormones and Behavior, 63,* 128 – 135.

Gildersleeve, K., Haselton, M.G., Larson, C.M., & Pillsworth, E.G. (2012). Body odor attractiveness as a cue of impending ovulation in women: Evidence from a study using hormone-confirmed ovulation. *Hormones and Behavior, 61,* 157 – 166.

Haselton, M.G. & Gildersleeve, K. (2011). Can men detect ovulation? *Current Directions in Psychological Science, 20,* 87 – 92.

Frederick, D.A., Pillsworth, E.G., Galperin, A., Gildersleeve, K., Filossof, Y., Fales, M.R., Lopez, C., Lurye, F., Phuphanich, M.E., Snider, J.B. (2009). Analyzing evolutionary social science and its popularizations – A review of The Caveman Mystique: Pop-Darwinism and the Debates Over Sex, Violence and Science. *Evolution and Human Behavior, 30,* 301 – 304.

Frederick, D.A., Galperin, A., Gildersleeve, K., Filossof, Y., Snider, J.B., Nguyen, D.M., Lurye, F., Fales, M.R., & Cantu, S.M. (2010). Sexuality: An evolutionary perspective. In T. Cooke (Ed.), *Sex and Society.* New York, NY: Marshall Cavendish.

Frederick, D.A., Filossof, Y.R., Gildersleeve, K., Maisel, N.C., Phuphanich, M.E., Poore, J.C., Preciado, M.A., Saxbe, D.E., & Usahacharoenporn, P. (2010). Fidelity. In T. Cooke (Ed.), *Sex and Society*. New York, NY: Marshall Cavendish.

Chapter 1: Introduction to Ovulatory Cycle Shifts in Human Sexuality

For most mammals, the high-fertility period just preceding and including the day of ovulation is the only time when sex can result in conception. This simple fact of biology, set in place hundreds of millions of years ago, has had profound implications for the evolution of mammalian sexuality. Throughout evolution, female reproductive success has depended largely on mating decisions at high fertility. In turn, male reproductive success has depended largely on mating efforts directed toward high-fertility females. Accordingly, two phenomena that are nearly universal among modern mammals are that a) females are more selective or differently selective in their choice of sex partners at high fertility, and b) males are particularly sexually attracted to females exhibiting cues of high fertility (Beach, 1976). That these phenomena characterize diverse species reinforces the notion that modern mammals represent the latest link in an unbroken chain of reproduction, and as such, they possess psychological adaptations for mating that have stood the test of time.

But where do humans fit in? Do women exhibit similar changes in their mate preferences and attractiveness across the ovulatory cycle? For many decades, the conventional scientific answer to this question—based largely on comparisons with some of our closest primate relatives—was simply *no*. For example, in chimpanzees, female sexual activity is confined almost exclusively to the high-fertility period of the ovulatory cycle. Within this phase, female mate preferences shift systematically as ovulation approaches and fertility increases (e.g., see Matsumoto-Oda, 1999). Female chimpanzees also display bold genital swellings in the high-fertility period, and males exhibit strong sexual interest in swollen as compared with nonswollen females (Deschner, Heistermann, Hodges, & Boesch, 2004). Clearly, humans do not exhibit changes of this magnitude, but are such changes completely absent?

Over the past two decades, scientific interest in this question has increased dramatically, leading to a proliferation of research on “cycle shifts” in human sexuality. This research has produced a large body of findings indicating that, whereas women's general sexual desire and sexual behavior do not show clear patterns of change across the ovulatory cycle, women's mate preferences *do* change systematically as they near ovulation within the cycle (reviewed by Thornhill & Gangestad, 2008; Gangestad, Thornhill, & Garver-Apgar, 2005b; and meta-analyzed by Gildersleeve, Haselton, & Fales, *in press*). This research has also produced a number of findings indicating that women’s sexual attractiveness increases at high fertility. Furthermore, emerging evidence suggests that exposing men to stimuli collected from women at high fertility or the high-fertility women themselves might alter their hormone levels and mating-related cognitions and behaviors (see Haselton & Gildersleeve, 2011). This suite of responses supports predictions derived from theories in evolutionary biology and parallels those previously observed in nonhuman mammals (reviewed in Thornhill & Gangestad, 2008; Gangestad, Thornhill, & Garver-Apgar, 2005b).

Research on cycle shifts in human sexuality has received attention from broad audiences due to its potentially important implications for understanding a wide range of phenomena, including initial attraction, mate choice, sexual behavior, romantic relationship dynamics, and hormone-behavior links. However, this work has also received criticism. Several published studies have not found support for key findings in this literature, raising doubts about the robustness of certain cycle shift effects and questions as to whether additional nonreplications remain unpublished. Furthermore, methods used to examine cycle shifts have varied extensively, making it difficult to coherently summarize and evaluate the evidence in the form of a traditional narrative review. Thus, a rigorous systematic review is clearly needed.

The primary goal of this dissertation was to quantitatively synthesize and evaluate the evidence for the two most widely studied cycle shift effects in this literature. Specifically, Chapter 2 presents a meta-analysis of the evidence for ovulatory cycle shifts in women's mate preferences, and Chapter 3 presents a meta-analysis of the evidence for cycle shifts in women's attractiveness. Through conducting these analyses, several gaps in the current literature became evident. For example, although women have occasionally been included along with men as raters in studies examining cycle shifts in women's attractiveness, no study to date has separately examined women's perceptions of the attractiveness of other women at high versus low fertility. Therefore, a secondary goal of this dissertation was to take a first step toward addressing this gap. Specifically, Chapter 4 presents a study examining women's attractiveness ratings of other women's high- and low-fertility body scents. Taken together, the results of this dissertation work suggest that the ovulatory cycle plays an important and nuanced role in human sexuality, with important implications for understanding the role of evolved psychological mechanisms and cycling reproductive hormones in human social motivations and behavior.

Chapter 2: Do Women's Mate Preferences Change across the Ovulatory Cycle? A Meta-analytic Review

The question of whether women experience systematic changes across the ovulatory cycle in mating-related motivations, preferences, cognitions, and behaviors has become a target of increasing empirical, theoretical, and popular attention over the past two decades. In particular, research examining ovulation-related “cycle shifts” in women’s mate preferences has reached landmark status in the evolutionary social sciences. Dozens of published studies have found evidence for cycle shifts in women’s mate preferences, and several lines of work have documented related effects (e.g., cycle shifts in women’s mating motivations, attraction to current relationship partners and other men, relationship satisfaction, and partner jealousy; reviewed by Gangestad & Thornhill, 2008; see also Larson, Haselton, Gildersleeve, & Pillsworth, 2013). Scientists and laypeople alike have increasingly cited these findings as evidence of the footprints of evolution in modern human sexuality and as revealing a potentially important, yet often overlooked, role of the ovulatory cycle in attraction, sexual behavior, and relationship dynamics.

However, there are ongoing debates as to whether current findings provide compelling evidence for ovulation-related cycle shifts in women’s mate preferences. Several recently published nonreplications have cast doubt on the robustness of these cycle shifts (e.g., Koehler, Rhodes, & Simmons, 2002), and some researchers have questioned whether the abundance of positive findings in the published literature reflects publication bias or other sources of bias.

Given the important implications of the existence of ovulation-related cycle shifts in women’s mate preferences for scientific and popular understandings of human sexuality, a rigorous evaluation of the extant empirical literature is clearly needed. However, published cycle shift studies have used a wide variety of methods and have examined preferences for a wide variety of characteristics in men. Furthermore, many cycle shift studies remain unpublished, possibly due to barriers to publishing null effects. Thus, even an exceptionally thorough narrative review of the published literature would be inadequate to compel firm conclusions about the existence and robustness of cycle shifts in women's mate preferences.

To address these issues, we conducted a meta-analysis on a large sample of 134 effects from 38 published and 12 unpublished studies. The goals of this meta-analysis were to use quantitative methods to assess the magnitude and robustness of predicted cycle shifts across the published and unpublished literatures, identify specific preferences for which cycle shifts are or are not robust and identify areas still in need of more research, and assess and adjust for bias that could have contributed to the observed pattern of cycle shifts.

Theoretical Background

For nearly all female mammals, the brief high-fertility window that precedes and includes the day of ovulation is the only time when sex can result in conception. Research on mating patterns in nonhuman mammals suggests that females of many mammalian species are more selective or differently selective at high fertility as compared with low fertility, possibly reflecting adaptive cycle shifts in their underlying mate preferences (e.g., for evidence in orangutans, chimpanzees, capuchins, and vervet monkeys, see Knott, Thompson, & Stumpf, 2007; Stumpf & Boesch, 2005; Pieta, 2008; for an early review, see Keddy-Hector, 1992). For example, one study found that female chimpanzees in the sexually active phase of their ovulatory cycle were more likely to mate repeatedly with high-ranking males on days of this phase when their fertility was maximally high than on days when their fertility was still relatively low. In contrast, the rate at which female mated repeatedly with low-ranking males did not increase with their fertility (Matsumoto-Oda, 1999).

The Ovulatory Shift Hypothesis

Observations such as these raise the question of whether women might also experience ovulation-related cycle shifts in their mate preferences. The ovulatory shift hypothesis, first discussed by Gangestad and Thornhill (1998) and later named as such in a review by Gangestad, Thornhill, and Garver-Apgar (2005b), proposes that women experience a nuanced pattern of relationship context-dependent cycle shifts in their preferences for certain characteristics in men. Specifically, the ovulatory shift hypothesis makes three key predictions that dozens of studies have aimed to test (reviewed in DeBruine et al., 2010; Gangestad & Thornhill, 2008; Thornhill & Gangestad, 2008).

Prediction 1

The first prediction of the ovulatory shift hypothesis is that women are more sexually attracted to characteristics in men that reflected relatively high *genetic quality* in ancestral males[1]—for example, the presence of genes with beneficial effects, absence of genes with harmful effects, or a low overall number of mutated genes—on high-fertility days of the ovulatory cycle as compared with low-fertility days of the cycle. This cycle shift in women's preference for cues of (ancestral) genetic quality is proposed to reflect psychological mechanisms that initially evolved because they increased ancestral females' likelihood of passing on certain genetic benefits to their offspring, thereby increasing their own reproductive success (roughly, their number of surviving descendants).

Cycling reproductive hormones, which underlie changes in female fertility across the

---

[1] We use the terms *genetic quality* and *reproductive success* as they are used in the field of biology. These terms do not imply that, because of their genetic constitution, some individuals are (or were ancestrally) superior to others in any way not outlined above. In addition, the ovulatory shift hypothesis makes no predictions regarding cycle shifts in women's preferences for female partners. Accordingly, most studies in this literature limit samples to women who identify as heterosexual, and our discussion of this literature is likewise limited to this group of women.

ovulatory cycle, could potentially exert a wide range of effects on female sexual motivations and attractions. According to the ovulatory shift hypothesis, ancestral females who experienced a shift in their attractions across the ovulatory cycle such that they experienced greater sexual attraction to males exhibiting cues of relatively high genetic quality at high fertility than at low fertility would have been more likely to have conceptive sex with such males and produce offspring who were also relatively high in genetic quality. Consequently, these females would have had higher reproductive success, on average, than would females whose attractions did not shift across the cycle in this way. Also, importantly, their descendants would have been more likely to possess any heritable aspects of the psychological mechanisms that produced the cycle shift in their mate preferences, making female descendants more likely to experience this cycle shift themselves. As long as conditions remained relatively stable, a cycle shift in preferences for males displaying cues of genetic quality would thereby have become increasingly common in females over evolutionary time.

Importantly, the ovulatory shift hypothesis predicts that the proposed cycle shift in women's attraction to men with characteristics that reflected genetic quality ancestrally will be present specifically when women evaluate men's immediate desirability as sex partners. Only if ancestral females' heightened preferences at high fertility for males displaying cues of genetic quality at least occasionally translated into higher rates of sex with such males during the fleeting high-fertility window would the posited cycle shift have been associated with higher reproductive success on average. Thus, it follows that the predicted cycle shift will be present specifically in the context of evaluating prospective partners for a short-term sexual affair or other types of relationships in which ancestral females' preferences would have been relatively likely to influence their immediate sexual behavior.

Prediction 2

The second prediction of the ovulatory shift hypothesis is that the proposed cycle shift in women's attraction to characteristics in men that reflected genetic quality in ancestral males will be absent or only weakly present when they evaluate men's desirability as a social partner in the long run. If ancestral females' heightened preferences at high fertility for males displaying cues of genetic quality did not translate into higher rates of sex with such males during the high-fertility window or translated instead into higher rates of nonsexual behaviors with such males during the high-fertility window (e.g., courtship behaviors that might lead to the formation of a long-term pair bond), the posited cycle shift would not have been associated with higher reproductive success on average. Thus, it follows that the predicted cycle shift in women's preferences for cues of ancestral genetic quality will be absent or only weakly present in the context of evaluating prospective partners for a long-term relationship (e.g., marriage) or other types of relationships in which ancestral females' preferences would have been relatively less likely to influence their immediate sexual behavior.

Many studies aiming to test the ovulatory shift hypothesis have asked women to evaluate men as potential partners for a "short-term relationship" or a "long-term relationship." To the extent that these terms imply a sexual affair and a long-term social partnership (such as marriage), respectively, it follows from Predictions 1 and 2 that the cycle shift in women's attraction to cues of genetic quality will be present and relatively pronounced in the former context but absent or only weakly present in the latter. Notably, however, the ovulatory shift hypothesis does not predict that the magnitude of the cycle shift will depend on how long women expect a relationship to last per se but rather on whether they expect the relationship to involve having sex in the near future.

In addition, many studies in this literature have asked women to evaluate men's attractiveness, physical attractiveness, sexual attractiveness, or sexiness or to evaluate the importance or desirability of a specific characteristic in a prospective partner without specifying any particular relationship context. The majority of these studies have assessed ratings of attractiveness, physical attractiveness, sexual attractiveness, or sexiness, whereas ratings of importance or desirability are very rare. Given previous research showing that women value physical attractiveness more when evaluating short-term sex partners than when evaluating long-term relationship partners (e.g., Li & Kenrick, 2006; Regan, 1998), it follows from the ovulatory shift hypothesis that women in these unspecified-context studies will generally exhibit a pattern of cycle shifts more similar to the pattern observed in a short-term context than to the pattern observed in a long-term context.

Predictions 1 and 2 highlight an implicit claim of the ovulatory shift hypothesis—namely, that certain potentially observable phenotypes in men constituted reliable "cues" to genetic quality in ancestral males. This claim rests on the following logic: differences between ancestral males in heritable genetic factors likely contributed to differences between males in immune function, vulnerability to environmental stressors, ability to compete with other males to attract mates, and other qualities that affected their reproductive success. Some of these genetic differences between males likely also contributed directly or indirectly (e.g., via effects on health) to detectable differences between males in physical appearance, body scents, vocal properties, and other phenotypes. For example, physical symmetry and masculinity are widely thought to have served as indicators of genetic quality in ancestral males (discussed in more detail below). In turn, selection could have acted on females to be sensitive to this phenotypic variation in males and, possibly, experience enhanced attraction to indicators of genetic quality under certain conditions.

Prediction 3

The third prediction of the ovulatory shift hypothesis is that, regardless of relationship context, women are *not* more sexually attracted to characteristics in men that reflected relatively high suitability as a long-term social partner and co-parent in ancestral males on high-fertility days of the ovulatory cycle as compared with low-fertility days of the cycle. The ovulatory shift hypothesis posits that females could have reproductively benefitted by mating with such males regardless of their current fertility (and in a variety of relationship contexts). For example, regardless of their fertility when they initiated the relationship, ancestral females who entered in to long-term pair bonds with males who were cooperative, caring, and highly investing partners and coparents would plausibly have had higher reproductive success, on average, than would females who entered in to long-term pair bonds with males who were uncooperative, negligent, or in other ways less suitable as a long-term partner and coparent.

Given the hypothesized reproductive benefits of mating with males relatively high in genetic quality, the ovulatory shift hypothesis raises the question of why females did not evolve to prefer males exhibiting cues of genetic quality at all times in the ovulatory cycle. One possible answer to this question is that cycle shifts in mate preferences initially evolved in an ancestral species (predating humans) that did not engage in high rates of pair bonding. In that context, females whose preferences shifted across the cycle in such a way that they were more likely to have sex with males displaying cues of genetic quality at high fertility but more likely to have sex with males offering other nongenetic reproductive benefits (e.g., material investment or protection) in the remainder of the cycle might have had greater reproductive success, on average, than did females whose preferences did not shift across the cycle in this way. In humans, for whom rates of pair bonding are high, these cycle shifts could simply be vestigial, reflecting remnants of psychological adaptations that now have a negligible impact on women's reproductive success or have a negative impact on women's reproductive success but have not yet been fully removed by selection (Gangestad & Garver-Apgar, 2013).

The Dual Mating Hypothesis

The *dual mating hypothesis* (Pillsworth & Haselton, 2006b) presents another possible answer to the question of why females did not evolve to prefer males with characteristics associated with relatively high genetic quality throughout the cycle. Like the ovulatory shift hypothesis, the dual mating hypothesis does not stipulate whether cycle shifts in mate preferences initially evolved in humans or in an ancestral species. However, unlike the ovulatory shift hypothesis (which is agnostic on this point), the dual mating hypothesis proposes that cycle shifts in mate preferences were associated with greater reproductive success among ancestral women and therefore are not merely vestigial.

According to the dual mating hypothesis, ancestral women would generally have maximized reproductive benefits by forming long-term pairbonds with men who were both high in genetic quality and highly suitable as a long-term social partner and co-parent. However, these characteristics were distributed across the population of men, and therefore, not all women could have formed long-term pair bonds with men who were high in both types of characteristics. The dual mating hypothesis proposes that women who formed long-term pair bonds with men who were relatively high in suitability as long-term partners but relatively low in genetic quality would have had higher reproductive success, on average, than women who formed long-term pair bonds with men who were relatively high in genetic quality but relatively low in suitability as long-term partners. This claim rests on the notion that high-quality biparental care and investment were critical for children's survival in ancestral environments (Geary, 2000; but see Sear & Mace, 2008). This claim is further reinforced by the notion that ancestral men who were relatively high in genetic quality might have been relatively less suitable and less available as long-term mates. Briefly, if men displaying cues of genetic quality were generally relatively desirable as sex partners, they might have tended to pursue short-term sexual relationships instead of pair bonds or outside of established pair bonds (thus diverting resources away from their long-term mate and children; see Gangestad & Simpson, 2000).

Following this line of reasoning, the dual mating hypothesis proposes that, among women who formed long-term pairbonds with men who were relatively high in suitability as long-term partners but relatively low in genetic quality, women who maintained their primary pairbond but also occasionally engaged in extra-pair sex with men of high genetic quality at high fertility (and when their sexual infidelity was unlikely to be discovered) would have had greater reproductive success, on average, than women who did not pursue this "dual mating" strategy. Evidence from nonhuman species in which females sometimes pursue this reproductive strategy suggests that behavioral adaptations that facilitate dual mating could have evolved even if rates of extra-pair sex were quite low (e.g., as low as 1%-5% in some bird species; see Thornhill & Gangestad, 2008).

Although many writings in this literature have suggested that cycle shifts in women's mate preferences reflect a long evolutionary history of dual mating in humans, the ovulatory shift hypothesis does not require that ancestral women engaged in extra-pair sex. For example, cycle shifts could be vestigial, as noted above. Alternatively, it is possible that cycle shifts have been maintained by selection in humans because they were historically associated with certain reproductive benefits in the context of sexually monogamous pair bonds, although this idea is not well developed in the current literature. In sum, if women experience the posited ovulation-related cycle shifts in their mate preferences, many interesting questions remain about the precise evolutionary pathways giving rise to them.

Cues of Genetic Quality in Ancestral Males

Research on cycle shifts in mate preferences has focused primarily on symmetry and masculinity as candidates for potentially observable characteristics that are likely to have been reliably associated with genetic quality in ancestral males[2]. Here, we briefly summarize the rationales typically given in support of claims that symmetry and masculinity were cues of genetic quality in ancestral males.

Symmetry

In biology, *developmental stability* is defined as "the ability of an organism to withstand genetic and environmental disturbances encountered during development so as to produce a predetermined optimum phenotype" (Clark, 1993, p. 15). Developmental stability is thought to reflect genetic quality as defined earlier (see, e.g., Thornhill & Gangestad, 2008; Van Dongen & Gangestad, 2011). Because researchers cannot directly measure developmental stability, they typically measure fluctuating asymmetry as a proxy (e.g., Klingenberg, 2003; Van Dongen, 2006). Fluctuating asymmetry is the extent to which the right and left sides of the body deviate

---

[2] A related hypothesis is that women will experience elevated preferences at high fertility for characteristics in men that reflect the presence of genes that would have been compatible with their own genes in the ancestral past. For example, it has been hypothesized that, all else equal, individuals who inherit different major histocompatibility complex (MHC) alleles from each of their parents have better pathogen defense than do individuals who receive the same alleles from both of their parents (e.g., Chen & Parham, 1989; Hughes & Nei, 1988; 1989; Penn, Damjanovich, & Potts, 2002). It follows that women might experience elevated attraction at high fertility to men with different MHC alleles than their own (men with whom they are, according to this view, genetically compatible). Our search discovered only two studies examining cycle shifts related to MHC-compatibility. One study found that women who shared a greater number of MHC alleles with their romantic partner (less compatible) experienced a greater increase at high fertility in their attraction to other men (Garver-Apgar, Gangestad, Thornhill, Miller, & Olp, 2006). A second study did not find evidence for a cycle shift in women's attraction to the scent of MHC-compatible men (Thornhill et al., 2003). Although the latter of these two studies was eligible for inclusion in this meta-analysis, we were unable to obtain the data needed to compute an effect size for it.

randomly from perfect bilateral symmetry (mirror images). To the extent that fluctuating asymmetry represents a departure from a genetic "blueprint" for a symmetrical body, it could indicate lower developmental stability and thus lower genetic quality. Consistent with this view, lower symmetry[3] (higher fluctuating asymmetry) has been linked to inbreeding, homozygosity, and deleterious recessive genes in nonhuman animals (see Rhodes, 2006; Thornhill & Gangestad, 1994; and see Carter, Weier, & Houle, 2009 for experimental evidence) and to negative health outcomes in humans (see Thornhill & Moller, 1997; Van Dongen & Gangestad, 2011).

In addition, fluctuating asymmetry appears to influence male success in attracting mates. Studies of many nonhuman animal species have found that more symmetrical individuals (lower in fluctuating asymmetry) have a significantly greater number of mates than do less symmetrical individuals (meta-analyzed by Moller & Thornhill, 1998). Several findings support parallel associations in humans. For example, more symmetrical men report having had a greater number of sex partners and having had sex at a younger age than do less symmetrical men (Thornhill & Gangestad, 1994). And women rate more facially symmetrical men as more attractive than less facially symmetrical men (meta-analyzed by Rhodes, 2006, and Van Dongen & Gangestad, 2011).

Masculinity

In biology, *masculine* characteristics refer to a number of physical and behavioral secondary sex characteristics that develop in males around the time of sexual maturity. Masculine characteristics are costly to produce and maintain; therefore, pronounced masculine

---

[3] Although it is typical in this literature to discuss effects of fluctuating asymmetry, for ease of interpretation, in the balance of this article we discuss effects of symmetry, by which we mean the inverse of fluctuating asymmetry. For example, we note that the ovulatory shift hypothesis predicts that women will demonstrate a stronger preference for more symmetrical men (men who are low in fluctuating asymmetry) at high fertility compared to low fertility.

characteristics could reflect good overall condition. Consistent with this view, studies of nonhuman animals have shown that food shortages bring about substantial reductions in the size of masculine characteristics, suggesting that masculine characteristics entail energetic costs that only individuals in good condition can afford (e.g., Wilson, Rogler, & Erb, 1979)[4]. Good condition is, in turn, partially tied to genetic quality (Rowe & Houle, 1996).

Like symmetry, masculine characteristics have been linked to male success in attracting mates. A meta-analysis of nonhuman lekking species, in which males engage in highly visible competitions against other males to attract females, found that males with larger masculine characteristics (e.g., antlers) attract a larger number of mates than do males with smaller masculine characteristics (Fiske, Rintamaki, & Karvonen, 1998). Relatedly, many studies support the idea that masculine characteristics have historically contributed to men's success in attracting mates, perhaps especially by increasing their success in competitive interactions with other men. For example, studies have found that experimentally increasing men's vocal, facial, and body masculinity increases others' perceptions of their dominance even more than perceptions of their attractiveness (see Puts, 2010). Studies also support a direct effect of masculinity on men's sexual attractiveness to women. For example, women in one study reported greater attraction to hypothetical men with more masculine faces, bodies, and voices when evaluating them as short-term sex partners than as long-term relationship partners (Little, Connely, Feinberg, Jones, & Roberts, 2011). Likewise, women in another study reported greater attraction to men whose photos they rated as more masculine and who had higher measured

---

[4] Some evidence suggests that testosterone, which is typically required to produce and often required to maintain masculine characteristics, also suppresses immune function. If correct, this implies that masculine characteristics entail immune costs (in addition to energetic costs) that only individuals in good condition—owing in part to their relatively high underlying genetic quality—can afford (see the immunocompetence handicap hypothesis, as discussed by Folstad & Karter, 1992; reviewed in Thornhill & Moller, 1997; meta-analyzed in Roberts, Buchanan, & Evans, 2004). Whether this is a likely mechanism through which masculinity was ancestrally associated with genetic quality has been contested. For a critique and alternative hypothesis, see Braude, Tang-Martinez, and Taylor (1999).

circulating testosterone when they evaluated those men's desirability for a brief affair than when they evaluated those men's desirability for a long-term relationship (Roney, Hasnon, Durante, & Maestripieri, 2006).

In sum, although research in this area has examined cycle shifts in women's preferences for a broad range of characteristics (discussed in detail in the Inclusion Criteria section), to date most studies have examined cycle shifts in women's preferences for symmetrical and masculine characteristics because these characteristics are widely thought to have served as cues of genetic quality in ancestral males. In addition, a smaller number of studies have examined cycle shifts in women's preferences for warmth and kindness, parenting ability, faithfulness, trustworthiness, material resources, and related characteristics because these characteristics are widely thought to have served as cues of "long-term partner quality" in ancestral males (a term which, for brevity, we use henceforward to refer to suitability as a long-term social partner and coparent). Importantly, we note that claims that certain characteristics were cues of genetic quality or long-term partner quality in ancestral males are conjectural and that the goal of this meta-analysis is not to directly test the accuracy of such claims. Rather, the goal is to determine whether predicted patterns of cycle shifts are robust for the characteristics most studied to date.

Methods

Search Strategy

As shown in Tables 1 and 2, we identified a large number of studies that collected data relevant to examining ovulation-related cycle shifts in women's preferences for various characteristics in men. We located studies through several channels, including reference sections of published articles, online databases and search engines, conference proceedings, listserv postings, and personal correspondence with researchers in this area. We chose several of these strategies with the specific goal of locating unpublished data and manuscripts not identified through other search methods. For example, we searched through the annual conference programs of the Society for Personality and Social Psychology (SPSP) (2005-2012) and of the Human Behavior and Evolution Society (HBES) (2000-2012) to identify researchers who had given talks or presented posters on research related to mating and the ovulatory cycle. We emailed all of these researchers a request for relevant unpublished data, including student projects. We also sent similar solicitations via listservs operated by the Society for Personality and Social Psychology, Society for the Psychological Study of Social Issues, and Society of Experimental Social Psychology and printed a solicitation in the summer 2010 Human Behavior and Evolution Society newsletter. Lastly, we emailed colleagues known to have conducted research on mating and the ovulatory cycle and requested that they alert us to any unpublished data that might be eligible for inclusion in the meta-analysis.

We used the following databases and search engines to locate published journal articles and unpublished manuscripts (e.g., master's theses and dissertations): PsycINFO, PubMed Central, Web of Science, BIOSIS, Dissertation Abstracts Online, ProQuest Dissertations & Theses, and Google Scholar. All searches utilized Boolean logic to search for entries that included a term related to ovulation, the menstrual cycle, fertility, or cycling hormones in conjunction with a term related to mate preferences – for example, "ovulat*" *or* "mid-cycle" *or* "menstrual cycle" *or* "cycl*" *or* "fertil*" *or* "high-fertility" *or* "low-fertility" *or* "conception risk" *or* "hormon*" *or* "luteal" *or* "follicular" *or* "estrogen" *or* "estradiol" *and* "mate" *or* "mating" *or* "attractive" *or* "partner" *or* "mate preference*" *or* "good genes" *or* "genetic quality" *or* "genetic benefits" *or* "fitness" *or* "symmet*" *or* "masculin*" *or* "dominan*" *or* "dimorph*" *or* "father" *or* "parent*." We also identified articles from the reference lists of empirical articles and earlier reviews of cycle shifts in women's sexual motivations and mate preferences (e.g., Gangestad, Thornhill, Garver-Apgar, 2005b; Gangestad & Thornhill, 2008; Jones et al., 2008). We discontinued our literature search in December 2012[5].

Inclusion Criteria

Studies have assessed ovulation-related cycle shifts in women's preferences for a variety of male characteristics using a variety of measures and have reported results and effect sizes in a variety of formats. We designed inclusion criteria that would retain a large and diverse sample of effects, while also limiting the sample to those effects that would facilitate a coherent evaluation of the evidence for the ovulatory shift hypothesis. In the following, we outline each of the specific inclusion criteria. For thoroughness, Tables 1 and 2 present all studies (and effects within studies) that met basic inclusion criteria (Criteria 1, 2, and 3), regardless of whether they were ultimately included in the meta-analysis. If a study assessed women's preferences for a variety of different characteristics, we included in the meta-analysis whichever effects were relevant to testing the ovulatory shift hypothesis and excluded those that were not.

Criterion 1: Naturally-cycling Women. The effect must have come from a study that included only naturally-cycling women—by which we mean reproductive-aged women not using hormonal contraception—or collected information about hormonal contraception use so that it was possible to examine naturally-cycling women's data separately[6].

Criterion 2: Assessed Ovulatory Cycle Position. The effect must have come from a study that collected information that could be used to estimate participants' position in the ovulatory

---

[5] Some researchers sent us unpublished data that have since been published (e.g., Thornhill, Chapman, & Gangestad, 2013). Thus, although the references of some studies included in the meta-analysis indicate a later date, we had in fact collected all data by December 2012.

[6] Most studies in this meta-analysis also reported having excluded women who were pregnant (or suspected pregnancy), breastfeeding, menopausal or postmenopausal, or reported a highly irregular cycle or other cycle abnormalities. However, we did not eliminate studies that did not report having collected and excluded women on the basis of this information.

cycle (e.g., date of last menstrual onset; see Appendix A for a more detailed description of cycle position estimation methods).

Criterion 3: Assessed Women's Preference for a Specific Characteristic in Men. The effect must have assessed a cycle shift in women's preference for a specific characteristic in men. For example, "facial masculinity" refers to a single, specific characteristic. In contrast, a man's relationship status or feelings about a current relationship partner could reflect a number of specific characteristics, as well as circumstances unrelated to those characteristics. It is unclear which specific characteristics women infer based on a man's relationship status. Therefore, we excluded effects that assessed women's preferences for men depicted as single, in love, having a girlfriend, or married (Bressan & Stranieri, 2008). In addition, we excluded effects that assessed women's attraction to real men whose characteristics were unknown to the researcher (e.g., a current relationship partner or celebrity; Gangestad, Thornhill, & Garver, 2002; Laeng & Falkenberg, 2007).

Physical attractiveness reflects a number of more specific characteristics and their interactions. It is unclear which specific characteristics women infer in men described as "physically attractive." Furthermore, the ovulatory shift hypothesis posits that the characteristics women find physically attractive vary systematically across the ovulatory cycle. For example, a "physically attractive" face could be a face high in masculinity for a woman at high fertility within the cycle but average in masculinity for the same woman at low fertility within the cycle. In other words, the ovulatory shift hypothesis posits that women's standards for what is physically attractive themselves shift across the cycle, making predictions about cycle shifts in women's attraction to men described as physically attractive unclear. For these reasons, we excluded effects that assessed women's preferences for physical attractiveness and handsomeness (e.g., Beaulieu, 2007; Caryl et al., 2009; Gangestad, Garver-Apgar, Simpson, & Cousins, 2007; Gangestad et al., 2010a).

Criterion 4: Assessed Preferences Pertinent to the Ovulatory Shift Hypothesis. The effect must have assessed a cycle shift in women's preference for a specific characteristic for which the ovulatory shift hypothesis makes a clear prediction—namely, a characteristic for which the extant literature provides a clear and widely accepted rationale for why it is likely to have been reliably associated with either genetic quality or long-term partner quality in ancestral males. Along these lines, we excluded effects measuring women's preference for social status, social competence, social sensitivity, and other social status-related characteristics (e.g., Izbicki & Johnson, 2010; Miller, 2003; Teatero, 2009); intelligence, inventiveness, creativity, academic achievement, and other intelligence-related characteristics (e.g., Prokosch, Coss, Scheib, & Blozis, 2009; Caryl et al., 2009; Miller, 2003, June); and cues of good health (e.g., a healthy-looking appearance; Jones, Perrett, et al., 2005). Extant findings suggest that all of these characteristics were associated with both genetic quality and partner quality in ancestral males, making predictions unclear (e.g., see Miller, 2000; Prokosh, Coss, Scheib, & Blozis, 2009; von Rueden, Gurven, & Kaplan, 2011).

Furthermore, the leading hypothesis pertaining to cycle shifts in women's preferences for cues of good health predicts that women will experience an elevated preference to affiliate with individuals *in general* (not only mates) displaying cues of good health when progesterone levels are highest within the cycle (e.g., in the luteal phase—the portion of the cycle following ovulation and extending to next menstrual onset; Jones, Perrett, et al., 2005). Progesterone dampens immune function in preparation for possible pregnancy, enabling the implantation of an embryo that is only partially genetically related to the mother and could otherwise be attacked by her immune system. Because of immune suppression associated with high progesterone levels, women might prefer to avoid potentially contagious individuals and instead affiliate with healthy individuals during the luteal phase. Fertility levels are also low during the luteal phase. Therefore, women could experience stronger preferences for cues of good health at low than at high fertility. Nonetheless, these progesterone-related cycle shifts in women's general social preferences would reflect different psychological mechanisms from those posited by the ovulatory shift hypothesis to produce ovulation-related cycle shifts in women's mate preferences. A meta-analysis evaluating evidence for progesterone-related cycle shifts would test a different hypothesis and require a different analysis strategy from that of the present meta-analysis (e.g., it would require comparing high-progesterone to low-progesterone days of the cycle, rather than high-fertility to low-fertility days of the cycle; therefore, we did not include these health effects in Tables 1 and 2).

Finally, because the ovulatory shift hypothesis only makes predictions about cycle shifts in women's preferences for cues of genetic quality and cues of long-term partner quality, we excluded a number of effects measuring preferences for characteristics that are not thought to have been associated with genetic quality or long-term partner quality in ancestral males (e.g., a mature appearance, a threatening appearance, same-race versus other-race facial appearance, adaptability, etc.; Izbicki & Johnson, 2010; McDonald & Navarrete, 2012; Miller, 2003).

Criterion 5: Assessed Preference for More over Less of One Characteristic, Rather than for One Characteristic Over Another. The effect must have assessed women's preference for more of a characteristic over less of that same characteristic (e.g., wealthy men over poor men), rather than women's preference for one characteristic over another characteristic (e.g., wealthy men over creative men). The latter confounds preference for one characteristic with preference for another, rendering effects from such studies incomparable with the other effects in the meta-analysis sample. For this reason, we excluded two studies: one that used a forced-choice paradigm to examine women's relative preference for creativity versus wealth in a prospective partner (Haselton & Miller, 2006) and another that used a "mate dollars" paradigm (Li, Bailey, Kenrick, & Linsenmeier, 2002) to examine the extent to which women traded off certain characteristics to "purchase" more of other characteristics in a hypothetical prospective partner (e.g., intelligence, social status, fit body, compatible interests, etc.; Li, Haselton, & Pillsworth, 2006).

Criterion 6: Common Mate Preference Measure. The effect must have been provided by a study that used a relatively common measure of mate preferences. We excluded studies that used highly uncommon measures of mate preferences in order to ensure that there was sufficient conceptual overlap among the measures included in the meta-analysis to yield interpretable mean effect sizes. Specifically, we excluded one study that measured women's self-reported perceived romantic compatibility with stimulus men (Flowe, Swords, & Rockey, 2012) and one study that measured women's self-reported likelihood of having sex with stimulus men (Rupp et al., 2009)[7]. We would have also excluded one study that used women's pupil dilation as a measure of attraction to male stimuli, but we had already excluded it on the basis of *Criterion 3* (Laeng & Falkenberg, 2007).

Criterion 7: Provided Information to Compute Appropriate Hedges's g. The article, poster, or study author must have provided the information needed to compute an appropriate Hedges's *g* effect size, as described below (see "Computing Effect Sizes"). If a paper or poster did not report the needed information, we contacted study authors to request this information. If

---

[7] Reported likelihood of having sex is conceptually different from attraction because it also entails attitudes towards casual sex and constraints on sexual behavior (e.g., having a current partner, risks associated with sex, taboos against casual sex, etc.).

the information was unavailable, we excluded the effect from the meta-analysis. Of those effects that were otherwise eligible for inclusion, we were unable to obtain effect size information for 11 effects from three studies: face and body averageness (one effect), face and body masculinity (one effect), and face and body symmetry (one effect; Peters, Rhodes, & Simmons, 2008); facial masculinity (one effect), body masculinity (one effect), facial symmetry (one effect), and body symmetry (one effect; Peters, Simmons, & Rhodes, 2009); vocal cues associated with perceived physical dominance (two effects) and vocal cues associated with perceived social dominance (two effects; Puts, 2005).

Analyses Conducted on "Broad" versus "Narrow" Sets of Mate Preference Measures

Even after removing effects that assessed cycle shifts in women's preferences for male characteristics for which the ovulatory shift hypothesis does not make a clear prediction (*Criterion 4*) and effects assessed with highly uncommon measures (*Criterion 6*), the remaining sample of effects was still very heterogeneous. A benefit of including all of these effects in the meta-analysis is that weighted mean effect sizes would reflect diverse male characteristics and measures. However, a cost is that weighted mean effect sizes would reflect male characteristics for which predictions are relatively weak (e.g., characteristics that are not yet widely accepted in this area as cues of genetic quality or long-term partner quality) and measures that are likely to be relatively insensitive to the fleeting, relationship context-dependent cycle shifts predicted by the ovulatory shift hypothesis. To resolve these tradeoffs, we created two nested samples of effects and conducted separate analyses on each. The first sample included a relatively "broad" set of male characteristics and measures, whereas the second sample included the relatively "narrow" subset of male characteristics and measures that we reasoned would be provide the strongest test of the ovulatory shift hypothesis.

The first, broad sample included effects examining cycle shifts in women's preferences for the following characteristics hypothesized to have served as cues of genetic quality in ancestral males: facial symmetry, body symmetry, scents associated with body symmetry, structural facial masculinity, male-typical facial movements, facial darkness, structural body masculinity (including, in addition to general body masculinity, muscularity, height, male-typical shoulder-to-hip ratio, male-typical waist-to-hip ratio, and strength), male-typical body motion (walking stride), torso hair, vocal masculinity (lower vocal pitch), behavioral dominance (including, in addition to general dominance, social presence, social respect and influence, direct intrasexual competitiveness, confrontativeness with other men, aggressiveness, arrogance and self-centeredness, egotism, and conceitedness), scents associated with behavioral dominance (specifically, scents associated with narcissism as assessed using the California Personality Inventory; see Havlíček, Roberts, & Flegr, 2005), facial cues associated with circulating testosterone, scents associated with circulating testosterone, and facial averageness. Although we might have excluded "social respect and influence" from this analysis for the same reason that we excluded social status (see *Criterion 4*), we chose to include it because a factor analysis in that study showed that social respect and influence had a very high loading on an Intrasexual Competitiveness factor (in fact, it had the highest loading of all characteristics rated in that study) and only a modest loading on a Good Investing Mate Qualities factor (see Gangestad et al., 2007). Excluding this characteristic had a negligible impact on the weighted mean effect sizes we report below and did not impact the statistical significance of any effects. The broad sample also included effects examining cycle shifts in women's preferences for the following characteristics hypothesized to have served as cues of partner quality in ancestral males: relationship skills, parenting skills, nurturance, sympathy, warmth, kindness, trustworthiness, faithfulness, financial success, and career success.

In terms of measures, the broad sample included studies in which women were asked to evaluate men or male stimuli as prospective short- or long-term relationship partners; to evaluate their attractiveness, physical attractiveness, sexual attractiveness, or sexiness without reference to a specific relationship context; or to evaluate a characteristic (e.g., “relationship skills”) on its importance or on how positive or negative they would feel about it in a prospective partner.

The second, narrow sample included the same studies and effects as the first sample, with three exceptions. First, we excluded effects measuring cycle shifts in women’s preferences for characteristics that are not yet widely accepted as cues of ancestral genetic quality: specifically, male-typical facial movements, male-typical walk, torso hair, skin darkness, and facial averageness (Frost, 1994; Izbicki & Johnson, 2010; Koehler, Rhodes, & Simmons, 2006; Morrison, Clark, Gralewski, Campbell, & Penton-Voak, 2010; Provost, Troje, & Quinsey, 2008; Rantala, Polkki, & Rantala, 2010). Some researchers in this area have suggested that the fact that a characteristic is more typical of men than of women suggests that that characteristic was linked to genetic quality in ancestral males. However, others have argued against this claim, noting an absence of strong theoretical or empirical reasons to posit that certain sex-differentiated characteristics were linked to male genetic quality ancestrally or that these characteristics play a role in male-male competition or in men’s sexual attractiveness to women. In addition, some researchers in this area have argued that, to the extent that averageness by definition indicates an absence of atypical features that might result from genetic mutations, rare alleles, homozygosity, or other potentially deleterious genetic factors, averageness might have served as a reliable indicator of genetic quality in ancestral males (e.g., see Thornhill & Gangestad, 1999a). However, recent evidence that extreme features are more attractive than average features for many dimensions of facial attractiveness poses a potential challenge to this view (Said & Todorov, 2011).

Second, we excluded studies that used measures of *stated preferences*. These measures involve women explicitly reporting how important or desirable a characteristic is in a prospective partner. Excluding these measures limited the sample to studies that used measures of *revealed preferences*. These measures involve women rating the attractiveness of (or choosing the most attractive among) male stimuli known by the researcher to vary on a characteristic. This allows the researcher to infer women's preferences based on their ratings (see Appendix A for more detail). We excluded studies using measures of stated preferences because we reasoned that such measures might tend to elicit women's reports of their general preferences, rather than in-the-moment preferences that might shift across the cycle. Thus, measures of stated preferences might be relatively insensitive to the temporally localized cycle shifts predicted by the ovulatory shift hypothesis. Furthermore, given that several studies have found that stated preferences are only weakly predictive of real-life dating behavior (see, e.g., Eastwick & Finkel, 2008; Eastwick, Luchies, Finkel, & Hunt, 2013; Todd, Penke, Fasolo, & Lenton, 2007), it remains an open question whether women have explicit knowledge of and can accurately report on the mate preferences that influence their real-life attractions. Finally, we reasoned that measures of stated preferences might not be as ecologically valid as measures of revealed preferences. That is, responding to a questionnaire about one's mate preferences might be less likely than directly evaluating male stimuli to bring online the evolved psychological mechanisms that are hypothesized to produce cycle shifts.

Third, we excluded studies that used stimuli that did not enable women to directly observe (see, hear, or smell) the characteristic of interest. For example, in one study, women viewed facial photos (no bodies) and rated the pictured men on attractiveness and physical strength (Izbicki & Johnson, 2010, June). Information relevant to judging men's physical strength is present to some extent in their facial appearance (Sell et al., 2009); therefore, the association between these two sets of ratings likely provides at least a rough measure of women's preference for strength. Nonetheless, women's ratings of body photos would likely have provided a more precise measure of their strength preferences. As a more extreme example, in another study, women read verbal descriptions of hypothetical men that varied only in the quality of their sense of humor and rated the men on attractiveness and body muscularity (Miller, 2003). In this case, the verbal descriptions contained little to no information relevant to judging body muscularity. To the extent that women envisioned more or less muscular men when rating the attractiveness of the hypothetical men, the association between their attractiveness ratings and body muscularity ratings could provide a rough measure of their preference for body muscularity. However, similar to strength ratings, body muscularity is a characteristic of the body; therefore, collecting women's ratings of body photos would likely have provided a more precise measure of their body muscularity preferences.

Computing Effect Sizes

The studies that we identified as potentially eligible for inclusion in this meta-analysis varied substantially in the type of data they produced and in the format in which they reported results. We used Hedges's *g* effect size metric for this meta-analysis because it could be computed for most of the studies in the sample, and its interpretation intuitively maps onto the predictions of the ovulatory shift hypothesis. In this meta-analysis, *g* represents the standardized mean difference between high and low fertility in women's preference for a characteristic (greater attraction to more versus less of the characteristic). A larger (more positive) *g* indicates that women's preference for a characteristic was stronger at high fertility than at low fertility. For example, a *g* of 0.2 would indicate that women's preference was, on average, two tenths of a standard deviation stronger at high fertility than at low fertility. Hedges's *g* is mathematically identical to Cohen's *d*, except that it includes an adjustment that reduces bias in small samples (Borenstein, Hedges, Higgins, & Rothstein, 2009). Hedges's *g* also has the same interpretation as Cohen's *d;* in psychology, effect sizes of 0.2, 0.5, and 0.8 are typically considered small, moderate, and large, respectively (Cohen, 1988).

Women's "preference" for a male characteristic was operationalized as one of the following: the proportion of forced-choice trials on which a woman chose stimuli with more of a characteristic over stimuli with less of that same characteristic, a woman's mean rating of the strength of her preference for stimuli with more of a characteristic over stimuli with less of that characteristic (in some studies, in each trial women completed a forced choice between two options and then rated the strength of their preference for the option they chose), the difference between a woman's mean rating of the attractiveness of stimuli with more of a characteristic and her mean attractiveness rating of stimuli with less of that characteristic, the correlation between a woman's attractiveness ratings of stimuli and the amount of a characteristic those stimuli possessed, the amount of a characteristic a woman perceived as most attractive (in some studies, women used a slider to manipulate a characteristic in a male stimulus until they had created what they perceived to be the most attractive version of the stimulus), or a woman's rating (or mean rating, if multiple items were used to assess a given preference) of the importance or desirability of a characteristic in a prospective partner.

If a study treated fertility as dichotomous (comparing high-fertility women to low-fertility women or the same women at high versus low fertility), computing Hedges's *g* to represent the difference between high and low fertility in women's preference for a male characteristic was straightforward. If a study treated fertility as continuous (assigning each woman a conception probability estimate based on her day in the cycle), computing Hedges's *g* entailed first computing the correlation between the continuous fertility variable and preference for the male characteristic across all women and then converting this correlation to *g*. We computed a Hedges's *g* for each preference assessed in each study; thus, studies that assessed multiple preferences contributed multiple effects (*g*s) to the meta-analysis.

Importantly, studies using measures of revealed preferences to assess women's mate preferences produce data that can be analyzed treating raters (women) or targets (men or male stimuli) as units of analysis. In this meta-analysis, all Hedges's *g*s were computed based on analyses that treated women as units of analysis. Thus, we can expect any statistically significant effects to generalize to new samples of women rating the stimuli that were included in this meta-analysis (rather than generalizing to new sets of male stimuli rated by the sample of women included in this meta-analysis). For example, for studies in which women rated the attractiveness of multiple male stimuli varying on a characteristic, we first computed for each woman the correlation between her attractiveness ratings of the male stimuli and the amount of the characteristic those stimuli possessed, then computed the mean correlation across all high-fertility women and the mean correlation across all low-fertility women, and finally computed a *g* representing the standardized difference between those two means. If available information could not be used to compute an effect size based on raters (women) as units of analysis but could be used to compute an effect size based on targets (men) as units of analysis, we report the latter effect size in Table 1 for thoroughness; however, we excluded effects based on targets as units of analysis from all analyses (Peters et al., 2009; Puts, 2005).

Several pieces of data identified as eligible for inclusion in this meta-analysis had not yet been analyzed to examine cycle shifts. In such cases, we asked the researcher to use the following guidelines to analyze the data or, if the researcher preferred, we used these guidelines to analyze their data. We developed these guidelines with the intent of retaining a large number of observations while providing a precise test of cycle shifts and giving researchers options to accommodate the format of their data while minimizing the potential for researchers to select among methods in order to obtain significant results (Simmons, Nelson, & Simonsohn, 2011).

First, we asked the researcher to exclude women who reported using hormonal contraception at the time of their participation. Next, if the researcher had collected this information, we asked the researcher to exclude women who, based on their self-reports, had irregular ovulatory cycles (typically operationalized as varying substantially in length from one cycle to the next), had used hormonal contraception at any time in the past three months (Nassaralla et al., 2011), were currently experiencing symptoms of or had experienced menopause, had an average cycle length shorter than 24 days or longer than 35 days (Harlow, 2000), suspected that they might be pregnant, or were over the age of 35 (and were therefore at an elevated likelihood of experiencing anovulatory cycles; Hale et al., 2007). Lastly, if a study included subsamples of women tested at both high and low fertility and women tested only at high or low fertility, we asked the researcher to limit the sample to women who had been tested at high and low fertility to enable within-participants comparisons.

If the researcher had already categorized women or observations as high- and low-fertility based on predetermined window definitions or had assigned each woman a conception probability estimate, we asked the researcher to retain their operationalization of fertility in effect size computations[8]. If the researcher had not yet defined high- and low-fertility windows or assigned conception probability estimates, we recommended that the researcher do so as follows: for studies using a between-participants design in which each woman completed a session at a single point in her cycle, we asked the researcher to assign each woman a conception probability estimate (Wilcox, Dunson, Weinberg, Trussell, & Day Baird, 2001) and to treat fertility as a continuous variable in effect size computations. If this was not possible, we asked researchers to instead categorize women who participated on forward cycle day 9-15 as high-fertility and women who participated on forward cycle day 21-35 as low-fertility and to exclude women falling outside of these windows. Likewise, for studies using a within-participants design in which each woman completed at least one session at high fertility and at least one session at low fertility, we asked researchers to categorize observations on forward cycle days 9-15 as "high-fertility" and observations on forward cycle days 21-35 as "low fertility" and to exclude observations falling outside of these windows. We chose these particular high- and low-fertility window definitions in order to maximize and minimize, respectively, the associated average conception probabilities (Wilcox et al., 2001), while still retaining a large number of observations in the analysis.

Coding Study Characteristics

Studies that have aimed to examine ovulation-related cycle shifts in women's mate preferences have varied in a number of ways—including, for example, characteristics of the sample of participants, researcher control over the research setting, methods for assessing women's fertility and mate preferences, and the specific characteristics for which preferences were assessed. Some of these methods have permitted greater researcher control and internal

---

[8] One study (Harris, 2010) reported multiple sets of results based on different high-fertility windows. For that study, we computed an effect size using the results based on the high-fertility window with the highest estimated average conception probability according to the values reported by Wilcox and colleagues (2001).

validity but limited sample size and external validity, whereas others have limited researcher control and internal validity but permitted a larger sample size and greater external validity. See Appendix A for a detailed discussion of the many sources of variation in this literature.

As shown in Table 2, we coded each study for a variety of characteristics. This included a) relationship context (short-term, long-term, or unspecified), b) country from which the sample of participants was drawn, c) sample type (college/university women, community women, or both), d) study setting (lab vs. “field,” which included online studies and one magazine survey with a mail-in response), e) study design (within-participants vs. between-participants), f) estimated average conception probability associated with the high- and low-fertility scheduling windows, g) cycle position estimation method (forward counting method vs. reverse counting method vs. average from forward and reverse counting methods vs. luteinizing hormone tests to verify impending ovulation vs. salivary ferning method to verify impending ovulation, noting for studies that used counting methods whether the benchmark date of menstrual onset was verified), h) type of stimuli (e.g., self-reported preferences vs. facial photos vs. body photos vs. average across face and body photos vs. vocal recordings vs. videotaped behavior vs. scent samples vs. face avatars vs. full body avatars vs. moving facial outlines, vs. body outline drawings, vs. verbal descriptions of hypothetical men vs. point-light walkers), i) method of determining the amount of the characteristic of interest possessed by the male stimuli (direct manipulations by the researcher vs. measured or coded by the researcher vs. rated by the participants in the cycle shift study vs. rated by a separate sample of participants), j) type of preference measure (stated preference vs. revealed preference), k) rating task (ratings of individual stimuli vs. two-option forced choice vs. multiple-option (three or more) forced choice vs. used a slider to manipulate the characteristic of interest), l) number of trials, and m) study publication status. Two researchers independently coded each study and then cross-checked their codes. In the case of discrepancies (which were rare), the researchers referred back to the paper or contacted the authors to verify the correct code. Thus, all codes were verified as correct.

Coding study characteristics was generally straightforward. As an exception, coding relationship context required additional considerations. In many studies examining cycle shifts, women were asked to complete two sets of ratings—one in which they evaluated men or male stimuli as potential “short-term” partners (typically defined as someone with whom they would consider having a brief sexual affair) and another in which they evaluated men or male stimuli as potential “long-term” partners (typically defined as someone with whom they would consider having a long-term dating or marriage/marriage-like relationship; e.g., Little, Jones, & Burriss, 2007)—or, less commonly, just one or the other. When studies explicitly specified a short-term and/or long-term relationship context, we coded them as such. Notably, however, in many studies, women were asked to evaluate men or male stimuli on “attractiveness” (e.g., Rupp, Librach, et al., 2009), "physical attractiveness" (e.g., Roney & Simmons, 2008), “sexual attractiveness” (e.g., Rantala et al., 2010), or “sexiness” (e.g., Thornhill & Gangestad, 1999b) or less commonly, to evaluate the importance or desirability of a characteristic in a potential partner (e.g., Caryl et al., 2009) without reference to a specific relationship context. When studies did not explicitly specify a short- or long-term relationship context, we coded them as “unspecified.”

Analyses

We used multilevel modeling for all analyses. Meta-analysis can be viewed as a special case of a multilevel model, involving effects nested within studies (Raudenbush & Bryk, 1985). The multilevel modeling approach offers a range of benefits over traditional meta-analytic methods, including the ability to properly include multiple, non-independent effects from the same sample within a single analysis and to test effect-level and study-level predictors of effect size and their cross-level interactions. As is conventional in meta-analysis, we weighted each effect by its inverse variance in order to give more precisely measured effects—often, those from larger studies—more "pull" on weighted mean effect sizes and regression coefficients (Raudenbush & Bryk, 2002). We estimated fixed effects and variance components using restricted maximum likelihood estimation procedures, which tend to reduce downward bias in variance components as compared with full maximum likelihood estimation procedures (O'Connell & McCoach, 2008). We conducted all analyses in HLM 7.0 and used the weighting and known variance options to weight effects by their inverse variances.

As indicated in the "Inclusion in Analyses" section of Table 1, we conducted two separate sets of analyses: one to examine cycle shifts in women's preferences for hypothesized cues of genetic quality in ancestral males and another to examine cycle shifts in women's preferences for hypothesized cues of long-term partner quality in ancestral males. Within each set of analyses, we first conducted analyses on the broad sample of effects and then conducted analyses on the narrow subset of effects described above (see Analyses Conducted on Broad versus Narrow Sets of Mate Preference Measures).

As described below, focal analyses revealed robust cycle shifts in women's preferences for all hypothesized cues of genetic quality. These analyses included a large but heterogeneous sample of effects. Therefore, they were sufficiently powered to provide clear results regarding the robustness of cycle shifts across all hypothesized cues of genetic quality but could not provide insight into how the magnitude and robustness of these cycle shifts differed across different kinds of studies (e.g., using different methods) or across different specific male characteristics (e.g., facial versus body masculinity). To address this issue, we conducted two additional sets of analyses. First, in both the broad and narrow samples of effects, we ran a series of moderation analyses. These analyses examined associations between specific study characteristics and the magnitude of cycle shifts across all hypothesized cues of ancestral genetic quality. Second, in the narrow sample of effects, we examined cycle shifts separately for each specific hypothesized cue of genetic quality for which the sample contained at least three effects. These analyses included small but relatively homogeneous samples of effects. Consequently, they were often underpowered and sometimes contained effects in only one or two relationship contexts. Nonetheless, their results provide insight into the specific male characteristics for which cycle shifts in women's preferences are or are not robust and highlight areas still in need of more research.

In the following, we describe the models used in these analyses in more detail. Results from the key analyses are presented in Figure 1.

Step 1. In each sample of effects, we first specified an unconditional random-effects model to compute the weighted mean *g* as an estimate of the "true" (population) mean standardized mean difference between high and low fertility in women's preference for a characteristic across all relationship contexts.

Level-1 Model (Effects)

$g_{ij} = \delta_{0j} + e_{ij}.$

Level-2 Model (Studies)

$\delta_{0j} = \gamma_{00} + u_{0j}.$

In the model above, g_{ij} is the observed standardized mean difference *i* for study *j*, δ_{0j} is the corresponding "true" mean *g* in the population of effects, e_{ij} is the sampling error associated with g_{ij} as an estimate of δ_{0j}, γ_{00} is the observed weighted mean *g* in the sample of effects, and u_{0j} is a study-level random error. Specifying δ_{0j} as random entails conceiving of *g* as varying randomly over the population of studies, thus allowing *g* to vary both as a function of sampling error and as a function of true between-studies variance (whereas specifying δ_{0j} as fixed would allow *g* to vary as a function of sampling error alone; see Raudenbush & Bryk, 2002). This approach is appropriate given that the studies included in this set of analyses are diverse in terms of sample characteristics, methods, and measures.

Step 2. In each sample of effects, we next specified three models to compute the weighted mean *g* in a short-term context (where the cycle shift is predicted to be largest), unspecified context (where the cycle shift is predicted to be intermediate between a short-term and long-term context), and long-term context (where the cycle shift is predicted to be smallest or absent), respectively, and to compare the weighted mean *g* across these three contexts. We created several variables to represent relationship context. These included *short*, a dummy-coded dichotomous variable taking on a value of 1 for effects measured in a short-term context and 0 for effects measured in a long-term or unspecified context; *long*, a dummy-coded dichotomous variable taking on a value of 1 for effects measured in a long-term context and 0 for effects measured in a short-term or unspecified context; and *unspecified*, a dummy-coded dichotomous variable taking on a value of 1 for effects measured in an unspecified context and 0 for effects measured in a short-term or long-term context. Starting with the unconditional model described in *Step 1,* we added dummy-coded relationship context variables, two at a time, as effect-level predictors. For example, in the following model, we have added *long* and *unspecified*. This establishes a short-term context as the comparison group, thereby enabling us to compute the weighted mean *g* in a short-term context and to estimate the magnitude of the difference between *g* in a short-term versus long-term context and between the weighted mean *g* in a short-term versus unspecified context. Although we report the results from all three models (with each of the three contexts as a comparison group) in the text of the Results section, for brevity, we present the complete results from only the models in which a short-term context was the comparison group in Tables 3 through 13.

Level-1 Model (Effects)

$g_{ij} = \delta_{0j} + \delta_{1j}(\text{long}) + \delta_{2j}(\text{unspecified}) + e_{ij}.$

Level-2 Model (Studies)

$\delta_{0j} = \gamma_{00} + u_{0j}$

$\delta_{1j} = \gamma_{10}$

$\delta_{2j} = \gamma_{20}.$

In the model above, g_{ij} is the observed standardized mean difference i for study j, δ_{0j} is the "true" mean g in a short-term relationship context, δ_{1j} is the "true" difference between g in a long-term versus short-term context, δ_{2j} is the "true" difference between g in an unspecified versus short-term context, e_{ij} is the residual sampling error associated with g_{ij} as an estimate of δ_{0j} unexplained by relationship context, γ_{00} is the observed weighted mean g in a short-term relationship context, u_{0j} is a study-level random error, γ_{10} is the regression coefficient representing the expected difference between g in a long-term versus short-term context (a negative value indicates that g is larger in a short-term context than in a long-term context), and γ_{20} is the regression coefficient representing the expected difference between g in an unspecified versus short-term context (a negative value indicates that g is larger in a short-term context than in an unspecified context).

We specified δ_{1j} and δ_{2j} as fixed in the above model. This assumes that any effect of relationship context on g varies across studies as a function of sampling error alone (and not as a function of true between-studies variance). Studies differed in how they defined short-term and long-term relationships and, if no relationship context was specified, in whether they asked women to evaluate male stimuli on physical attractiveness, attractiveness, sexual attractiveness, sexiness, or another variable. Thus, any effect of relationship context on *g* could vary as a function of true between-studies variance in addition to sampling error. For many of the analyses, we were working with relatively small samples of effects and therefore had insufficient power to specify relationship context effects as random. However, when possible, we tested these effects as both fixed and random and found that this did not change the pattern of results. For consistency, in the text and tables, we report results based on models in which relationship context effects were fixed.

Step 3. In the genetic quality analysis only, we then ran numerous analyses to test whether specific study characteristics were associated with between-studies variance in effect size (*g*) after controlling for relationship context and, if there was sufficient power, to test whether specific study characteristics were associated with between-studies variance in the effect of relationship context (short-term versus unspecified and short-term versus long-term) on effect size. When possible, we ran moderation analyses for each of the study characteristics displayed in Table 2, with the exception of sample country. This included study publication status, sample composition, setting, design, estimated difference in the average conception probability of the high- versus low-fertility windows, counting method used to estimating ovulatory cycle position, whether the benchmark date of menstrual onset had been verified, type of stimuli, method used to determine the amount of a characteristic male stimuli possessed, whether the study used a stated or revealed preference measure to assess mate preferences, type of rating task, and number of trials.

Notably, moderation analyses were limited in several ways. Because power was often low, we tested one study characteristic at a time. Thus, if analyses revealed an association between a study characteristic and effect size, other correlated study features could account for this association. Indeed, many study characteristics were highly intercorrelated. For example, nearly all studies using a between-participants design also used the forward counting method to estimate women's position in the ovulatory cycle. In addition, very few studies used rigorous methods to determine women's position in the ovulatory cycle (e.g., few studies verified ovulation using luteinizing hormone tests). Thus, analyses examining associations between the use of these methods and effect size were underpowered. As in all research literatures, many factors influence the extent to which studies provide precise measures of effects. Even if this meta-analysis cannot examine all of the many sources of variation in cycle shifts, it can still examine key sources of variation and determine whether robust patterns of cycle shifts emerge despite this variation.

Results

As explained in detail above, the ovulatory shift hypothesis posits that women experience a relationship context-dependent cycle shift in their preferences for characteristics that reliably indicated genetic quality in ancestral males. Specifically, the ovulatory shift hypothesis predicts that women's preferences for these characteristics are stronger at high fertility than at low fertility and that this shift will be most pronounced when women evaluate prospective partners in a short-term relationship context and least pronounced when they evaluate prospective partners in a long-term relationship context. Most studies categorized as "unspecified" in this meta-analysis asked women to evaluate men or male stimuli on *attractiveness*. As noted above, previous research has shown that women value physical attractiveness more in short-term sex partners than in long-term relationship partners (e.g., Regan, 1998; Li & Kenrick, 2006); therefore, we further predict that women will exhibit a pattern of cycle shifts in an unspecified relationship context that more closely resembles the pattern of cycle shifts in a short-term context than in a long-term context. Although a cycle shift in women's preferences for cues of genetic quality could emerge across the three relationship contexts, this is not a requirement of the ovulatory shift hypothesis. Rather, the more precise prediction is that any such cycle shift will be strongly moderated by relationship context.

Lastly, the ovulatory shift hypothesis posits that, regardless of relationship context, women do not experience a cycle shift in their preferences for characteristics that reliably indicated suitability as a long-term social partner and coparent in ancestral males.

Preference for All Hypothesized Cues of Ancestral Genetic Quality: Broad Set of Measures

The first analysis examined cycle shifts in preferences for all cues of genetic quality in the sample of effects that included a broad set of mate preference measures. This analysis included 96 effects from 50 studies (total $N = 5{,}471$). As shown in Table 3, *Step 1* revealed that the weighted mean g estimating the true population mean standardized mean difference between high and low fertility in women's preference for hypothesized cues of ancestral genetic quality across short-term, long-term, and unspecified relationship contexts was small ($g = 0.15$, $SE = 0.04$) but statistically significant ($p < .001$). Thus, in this set of effects, women's preference for these characteristics was approximately 0.15 of a standard deviation stronger at high fertility than at low fertility.

Step 2 revealed that the weighted mean g in a short-term context was small ($g = 0.21$, $SE = 0.06$) but statistically significant ($p = .001$); the weighted mean g in an unspecified relationship context was small ($g = 0.16$, $SE = 0.05$) but statistically significant ($p = .003$); and the weighted mean g in a long-term context was near zero ($g = 0.06$, $SE = 0.06$) and not statistically significant ($p = .32$). Comparing the three contexts revealed that the weighted mean g was larger in a short-term context than in a long-term context, and this difference was statistically significant ($p =$

.002). The weighted mean *g* did not significantly differ between a short-term context and an unspecified context or between an unspecified context and a long-term context (*p* = .54 and .19, respectively).

Step 3 revealed several moderation effects. All of the following study characteristics were associated with a larger cycle shift after controlling for the effect of relationship context: using scent stimuli, rather than any other type of stimuli (*p* = .01); direct measurement to determine, rather than any other method to determine the amount of a characteristic possessed by male stimuli (*p* = .06); and the study being published (*p* = .03). In contrast, all of the following study characteristics were associated with a smaller cycle shift after controlling for the effect of relationship context: having participants in the cycle study rate a characteristic in male stimuli, rather than using any other method to determine the amount of a characteristic possessed by male stimuli (*p* = .06).

All of the following were associated with a larger difference between the magnitude of the cycle shift in a short-term and long-term relationship context (short-term > long-term): a field (usually, online) setting, rather than a lab setting (*p* = .02); a forward counting method, rather than a backward counting method or an average of forward and backward counting methods, to estimate women's cycle position (*p* = .01); and the study being published (*p* =.003). In contrast, the following was associated with a smaller difference between the magnitude of the cycle shift in a short-term and long-term relationship context: using facial photos as stimuli, rather than any other kind of stimuli (*p* = .03).

All of the following were associated with a larger difference between the magnitude of the cycle shift in a short-term and unspecified relationship context (short-term > unspecified): using body photos as stimuli, rather than any other type of stimuli (*p* = .001); directly manipulating the male characteristic, rather than using any other method to determine the amount of the male characteristic possessed by the stimuli ($p < .001$); using a two-option forced choice task, rather than any other task to assess mate preferences ($p = .02$). In contrast, all of the following were associated with a smaller difference between the magnitude of the cycle shift in a short-term and unspecified relationship context: using a rating task, rather than any other task to assess mate preferences ($p = .04$).

Preference for All Hypothesized Cues of Ancestral Genetic Quality: Narrow Set of Measures

The next analysis examined cycle shifts in preferences for all cues of genetic quality in the sample of effects that included a narrow set of mate preference measures. This analysis included 68 effects from 42 studies (total $N = 4,884$). As shown in Table 4, *Step 1* revealed that the weighted mean g estimating the true mean standardized mean difference between high and low fertility in women's preference for hypothesized cues of ancestral genetic quality across short-term, long-term, and unspecified relationship contexts was small ($g = 0.17$, $SE = 0.04$) but statistically significant ($p < .001$). Thus, in this set of effects, women's preference for these characteristics was generally approximately 0.17 of a standard deviation stronger at high fertility than at low fertility.

Step 2 revealed that the weighted mean g in a short-term context was small to moderate ($g = 0.26$, $SE = 0.07$) and statistically significant ($p < .001$); the weighted mean g for attractiveness ratings made in an unspecified relationship context was small ($g = 0.20$, $SE = 0.05$) but statistically significant ($p = 0.001$); and the weighted mean g in a long-term context was near zero ($g = 0.02$, $SE = 0.06$) and not statistically significant ($p = .75$). Comparing the three contexts revealed that the weighted mean g was larger in a short-term context than in a long-term context, and this difference was statistically significant ($p < .001$). The weighted mean g did not differ between a short-term context and an unspecified context (*p* = .42). The weighted mean *g* was significantly larger in an unspecified context than in a long-term context (*p* = .04).

Step 3 revealed several moderation effects. All of the following study characteristics were associated with a significantly or marginally significantly larger effect after controlling for relationship context: a sample composed of women from the community or a combination of undergraduate and community women, rather than only undergraduate women (*p* = .08); a field (usually, online) setting, rather than a lab setting (*p* = .08); a between-participants design, rather than a within-participants design (*p* = .08); using scent stimuli, rather than any other type of stimuli (*p* = .02); direct measurement, rather than any other method to determine the amount of a characteristic possessed by male stimuli (*p* = .09); and the study being published (*p* = .03). In contrast, the following study characteristic was associated with a significantly smaller effect after controlling for relationship context: having participants in the cycle study rate a characteristic in male stimuli, rather than any other method to determine the amount of a characteristic possessed by male stimuli (*p* = .02). Lastly, the following characteristic was associated with a larger difference between the effect size in a short-term and long-term relationship context: a field (usually, online) setting, rather than a lab setting (*p* = .09).

Preference for Facial Symmetry

The next few analyses examined cycle shifts in women's preferences for specific hypothesized cues of genetic quality in the sample of effects that included a narrow set of mate preference measures. The first of these analyses examined cycle shifts in women's preference for facial symmetry and included eight effects from seven studies (total *N* = 870). As shown in Table 5, *Step 1* revealed that the weighted mean *g* estimating the true mean standardized mean difference between high and low fertility in women's preference for symmetry across short-term, long-term, and unspecified relationship contexts was near zero (*g* = 0.07, *SE* = 0.10) and not statistically significant (*p* = .48). Thus, in this set of effects, women's preference for symmetry was not generally stronger at high fertility than at low fertility.

Step 2 revealed that the weighted mean *g* in a short-term context was small to moderate (*g* = 0.30, *SE* = 0.20) and not statistically significant (*p* = .19); the weighted mean *g* in an unspecified context was near zero (*g* = -0.02, *SE* = 0.16) and not statistically significant (*p* = .90); and the weighted mean *g* in a long-term context was small and negative (*g* = -0.16, *SE* = 0.25) and not statistically significant (*p* = .54). Comparing the three contexts revealed that the weighted mean *g* was larger in a short-term context than in a long-term context, and this difference was marginally statistically significant (*p* = .08). The weighted mean *g* was somewhat larger in a short-term context than in an unspecified context, but this difference was not statistically significant (*p* = .27). Likewise, the weighted mean *g* was somewhat less negative in an unspecified context than in a long-term context, but this difference was not statistically significant (*p* = .66).

Preference for Scents associated with Face and Body Symmetry

The next analysis examined cycle shifts in women's preference for scents associated with face and body symmetry and included a small sample of three effects from three studies (total *N* = 141). As shown in Table 6, *Step 1* revealed that the weighted mean *g* estimating the true mean standardized mean difference between high and low fertility in women's preference for scent cues of symmetry was large (*g* = 0.83, *SE* = 0.20) but not statistically significant (*p* = .14). We could not perform *Step 2* because all of the effects in this sample were measured in an unspecified relationship context. Thus, in this set of effects, women's preference for scents associated with symmetry was approximately 0.83 of a standard deviation stronger at high fertility than at low fertility, but more data are needed to confidently determine the robustness of this cycle effect and to examine differences across relationship contexts.

Preference for Structural Facial Masculinity

The next analysis examined cycle shifts in women's preference for structural facial masculinity and included 23 effects from 19 studies (total $N = 3{,}335$). As shown in Table 7, *Step 1* revealed that the weighted mean g estimating the true mean standardized mean difference between high and low fertility in women's preference for structural facial masculinity across short-term, long-term, and unspecified relationship contexts was small ($g = 0.13$, $SE = 0.06$) but statistically significant ($p = .05$). Thus, in this set of effects, women's preference for structural facial masculinity was generally approximately 0.13 of a standard deviation stronger at high fertility than at low fertility.

Step 2 revealed that the weighted mean g in a short-term context was near zero ($g = -0.02$, $SE = 0.14$) and not statistically significant ($p = .91$); the weighted mean g in an unspecified context was small ($g = 0.17$, $SE = 0.07$) and statistically significant ($p = .02$); and the weighted mean g in a long-term context was near zero ($g = -0.01$, $SE = 0.13$) and not statistically significant ($p = .95$)[9]. Comparing the three contexts revealed that the weighed mean g was somewhat larger in an unspecified context than in a short-term or long-term context, but these differences were not significant ($p = .24$ and .23, respectively). The weighted mean g did not differ between a short-term and long-term context ($p = .96$).

---

[9] Luevano and Zebrowitz (2006) and Izbicki and Johnson (2010) both presented participants with facial photographs and asked them to rate the pictured men for "masculinity," as well as certain personality characteristics (e.g., dominance, warmth, maturity, etc.). Because participants were asked to evaluate the pictured men for personality characteristics, it is possible that participants evaluated the men on inferred personality masculinity rather than on structural facial masculinity. Excluding the four effects (two measured in a short-term context, two measured in a long-term context) from these two studies changed the results as follows: overall weighted mean $g = 0.18$ ($SE = 0.05$, $p < .01$), short-term weighted mean $g = 0.28$ ($SE = 0.20$, $p = .19$), unspecified weighted mean $g = 0.18$ ($SE = 0.06$, $p = .01$), long-term weighted mean $g = 0.17$ ($SE = 0.19$, $p = .38$), and there were no statistically significant differences between relationship contexts.

Preference for Structural Body Masculinity

The next analysis examined women's preference for structural body masculinity and included 12 effects from five studies (total $N = 589$). As shown in Table 8, *Step 1* revealed that the weighted mean g estimating the true mean standardized mean difference between high and low fertility in women's preference for structural body masculinity across short-term and long-term relationship contexts was small ($g = 0.21$, $SE = 0.08$) and marginally statistically significant ($p = .07$). Thus, in this set of effects, women's preference for structural body masculinity was generally 0.21 of a standard deviation stronger at high fertility than at low fertility, but more data are needed to determine the robustness of this cycle shift.

Step 2 revealed that the weighted mean g in a short-term context was small to moderate ($g = 0.35$, $SE = 0.10$) and statistically significant ($p = .04$); and the weighted mean g in a long-term context was near zero ($g = 0.09$, $SE = 0.09$) and not statistically significant ($p = .40$). The weighted mean g was larger in a short-term context than in a long-term context, and this difference was statistically significant ($p = .03$). None of the effects in this sample were measured in an unspecified context.

Preference for Vocal Masculinity (Lower Vocal Pitch)

The next analysis examined cycle shifts in women's preference for vocal masculinity and included a small sample of four effects from two studies (total $N = 159$). As shown in Table 9, *Step 1* revealed that the weighted mean g estimating the true mean standardized mean difference between high and low fertility in women's preference for vocal masculinity (lower vocal pitch) across short-term and long-term relationship contexts was small ($g = 0.28$, $SE = .18$), but power was insufficient to test the statistical significance of this effect. Thus, in this set of effects, women's preference for vocal masculinity appeared to be 0.28 of a standard deviation stronger at high fertility than at low fertility, but more data are needed to determine the robustness of this cycle shift.

Step 2 revealed that the weighted mean *g* in a short-term context was small to moderate (*g* = 0.40, *SE* = 0.20), and the weighted mean *g* in a long-term context was small (*g* = 0.18, *SE* = 0.20). Power was insufficient to test the statistical significance of either effect. The weighted mean *g* was somewhat larger in a short-term context than in a long-term context, but this difference was not statistically significant (*p* = .39). None of the effects in this sample were measured in an unspecified context.

Preference for Behavioral Dominance or Felt Superiority over Other Men

The next analysis examined cycle shifts in women's preference for behavioral dominance and included 12 effects from three studies (total *N* = 255). As shown in Table 10, *Step 1* revealed that the weighted mean *g* estimating the true mean standardized mean difference between high and low fertility in women's preference for behavioral dominance across short-term and long-term relationship contexts was near zero (*g* = 0.04, *SE* = 0.06) and not statistically significant (*p* = .55). Thus, in this set of effects, women's preference for behavioral dominance was not generally stronger at high fertility than at low fertility.

Step 2 revealed that the weighted mean *g* in a short-term context was small (*g* = 0.19, *SE* = 0.07) and marginally statistically significant (*p* = .09); and the weighted mean *g* in a long-term context was small and negative (*g* = -0.11, *SE* = 0.07) and not statistically significant (*p* = .28). Comparing the two contexts revealed that the weighted mean *g* was larger in short-term context than in a long-term context, and this difference was statistically significant, (*p* = .01). None of the effects in this sample were measured in an unspecified relationship context.

Preference for Facial Cues of Testosterone

The next analysis examined cycle shifts in women's preference for a facial appearance associated with higher levels of circulating testosterone and included a small sample of three effects from three studies (total *N* = 135). As shown in Table 11, *Step 1* revealed that the weighted mean *g* estimating the true mean standardized mean difference between high and low fertility in women's preference for facial cues of testosterone was small (*g* = 0.20, *SE* = 0.22) and not statistically significant (*p* = .46). Thus, in this set of effects, women's preference for facial cues of circulating testosterone was not generally stronger at high fertility than at low fertility. All of the effects in this sample were measured in an unspecified relationship context. More data are needed to determine whether there is any cycle effect on women's preference for facial cues of circulating testosterone and to examine possible differences across relationship contexts.

Preference for All Hypothesized Cues of Ancestral Partner Quality: Broad Set of Measures

The next analysis examined cycle shifts in women's preferences for cues of long-term partner quality in the sample of effects that included a broad set of mate preference measures. This analysis included 38 effects from eight studies (total *N* = 622). As shown in Table 12, *Step 1* revealed that the weighted mean *g* estimating the true mean standardized mean difference between high and low fertility in women's preference for hypothesized cues of long-term partner quality across short-term, unspecified, and long-term relationship contexts was near zero (*g* = -0.004, *SE* = 0.04) and not statistically significant (*p* = .91). Thus, in this set of effects, women's preferences for these characteristics did not generally shift across the cycle.

Step 2 revealed that the weighted mean *g* was near zero and not statistically significant in a short-term (*g* = -0.06, *SE* = 0.05, *p* = .30), unspecified (*g* = -0.04, *SE* = 0.17, *p* = .83); or long-term relationship context (*g* = 0.05, *SE* = 0.05, *p* = .31). The weighted mean *g* was somewhat more negative in a short-term context than in a long-term context (suggesting that women's preferences for these characteristics are somewhat *weaker* at high fertility as compared with low fertility when they evaluate men as short-term partners), and this difference was marginally statistically significant (*p* = .09). The weighted mean *g* was somewhat more negative in an unspecified context than in a long-term context, but this difference was not statistically significant (*p* = .61). The weighted mean *g* did not significantly differ between an unspecified and short-term context (*p* = .92).

Preference for All Hypothesized Cues of Ancestral Partner Quality: Narrow Set of Measures

The next analysis examined cycle shifts in women's preferences for cues of long-term partner quality in the sample of effects that included a narrow set of mate preference measures. This analysis included eight effects from a single study (total N = 243). Because all effects were from the same study, we used least squares estimation procedures.

As shown in Table 13, *Step 1* revealed that the weighted mean *g* estimating the true mean standardized mean difference between high and low fertility in women's preference for cues of long-term partner quality across short-term and long-term relationship contexts was near zero (*g* = -0.05, *SE* = 0.05) and not statistically significant (*p* = .28). Thus, this preliminary analysis did not reveal any evidence that women's preferences for these characteristics shift across the cycle.

Step 2 revealed that the weighted mean *g* in a short-term context was small and negative (*g* = -0.12, *SE* = 0.07) and marginally significant (*p* = .11), and the weighted mean *g* in a long-term context was near-zero (*g* = 0.01, *SE* = 0.07) and not statistically significant (*p* = .83). The weighted mean *g* was somewhat more negative in a short-term than in a long-term context, but this difference was not statistically significant (*p* = .19). None of the effects in this sample were measured in an unspecified relationship context. Ultimately, more data from a larger number of studies are needed to determine with confidence whether women experience relationship context-dependent cycle shifts in their preferences for these characteristics.

Can Bias Account for the Observed Patterns of Cycle Shifts?

Underrepresentation of Small Effects. When a meta-analysis reveals robust, nonzero mean effects, and perhaps particularly when those effects are consistent with predictions from a theory or previously published findings, an important question is whether these mean effects have been inflated by an underrepresentation of small effects in the meta-analysis sample. Larger effects are more likely to reach statistical significance, and statistically significant findings are more likely to make their way into the published literature (e.g., due to pressure on researchers and journals not to publish null effects). In turn, published findings are typically easier for meta-analysts to locate. In addition, if researchers are more confident in or keep better track of unpublished data showing significant effects, they might be more likely to share these data with meta-analysts. Therefore, larger effects might be more likely to make their way into a meta-analysis sample, whereas smaller effects are more likely to be overlooked.

A common method for assessing whether it is likely that small effects are underrepresented in a meta-analysis sample is to examine funnel plots. In funnel plots, effect sizes are plotted against their standard errors, with larger effects on the right and smaller standard errors—indicating more precise estimates (often, from larger studies)—at the top. If small effects are sufficiently well represented, effects will be distributed symmetrically about the mean effect size from the top to the bottom of the funnel. This is because sampling error is equally likely to result in an overestimation as an underestimation of the true effect size. If, however, small effects are underrepresented, more precise effects (top of the funnel) will be symmetrically distributed about the mean effect size, but less precise effects (bottom of the funnel) will skew to the right.

At low precision, only large effects will reach statistical significance. Therefore, the gap that forms in the lower left quadrant of the funnel suggests small effects that are missing due to publication bias or some other source of bias favoring the inclusion of significant effects.

We used funnel plots to assess whether it was likely that small effects were underrepresented among those effects for which the ovulatory shift hypothesis predicts a relationship context-dependent cycle shift—namely, effects measuring cycle shifts in women's preferences for hypothesized cues of genetic quality. We predicted based on the ovulatory shift hypothesis that women would exhibit cycle shifts in these preferences in a short-term and unspecified relationship context but not in a long-term relationship context, and indeed this is the pattern we observed in the focal analyses examining cycle shifts in preferences for all hypothesized cues of ancestral genetic quality. Therefore, we plotted effects in a short-term or unspecified context separately from effects in a long-term context. We created these plots for both the broad and narrow samples of effects.

As shown in Figure 2A, the funnel plots did not reveal any evidence of bias in the sample of effects that included a broad set of mate preference measures. Observed effect sizes are roughly evenly distributed about the mean from the top to the bottom of the funnel in the long-term context and in the combined short-term and unspecified context. Furthermore, Duval and Tweedie's (2000) "trim and fill" procedure, performed with Comprehensive Meta-analysis software, did not indicate an absence of any putative missing effects in either plot.

As shown in Figure 2B, the funnel plots revealed evidence of slight bias in the sample of effects that included a narrow set of mate preference measures. Whereas observed effect sizes are roughly evenly distributed about the mean from the top to the bottom of the funnel in the long-term context, observed effect sizes skew slightly to the right moving from the top to the bottom of the funnel in the combined short-term and unspecified context. Accordingly, the Trim and Fill procedure indicated that one effect was missing from the short-term and unspecified plot. Imputing the putative missing effect resulted in a negligible reduction in the weighted mean effect size in the combined short-term and unspecified context (from $g = 0.21$ to 0.20, with no change in the 95% confidence interval). Therefore, overall, the funnel plots and trim and fill procedures did not reveal compelling evidence that the pattern of cycle shifts observed in this meta-analysis is accounted for by an underrepresentation of small effects in the sample.

Researcher Degrees of Freedom in Defining High- and Low-fertility Windows

"Researcher degrees of freedom" refers to ambiguity or flexibility in data collection and analysis practices that enables researchers to try out several different methods and, possibly, choose whichever method or analysis produces significant results (therefore dramatically increasing the Type 1 error rate; Simmons et al., 2011). Most aspects of study design are determined in advance of data collection, eliminating concerns about researcher degrees of freedom therein. However, one aspect of study design that is relatively unique to cycle shift research and is not always determined in advance of data collection is how to define high- and low-fertility windows. This leaves open the possibility that researchers could select, post hoc, high- and low-fertility windows that happen to produce predicted cycle shifts.

We initially attempted to address this potential concern by conducting a moderation analysis on the sample of effects examining cycle shifts in women's preferences for hypothesized cues of genetic quality. Specifically, we examined the association between effect size and the difference between the estimated average conception probability of the high-fertility window and the estimated average conception probability of the low-fertility window. We reasoned that if true cycle shifts were present, effects would be larger among studies that used a stronger fertility “manipulation” (a larger difference between the estimated average conception probability of the high- versus low-fertility windows). We did not observe any such association. However, notably, our method of estimating the average conception probability of high- and low-fertility windows had several potential shortcomings (see Appendix A).

Given the uninformative nature of this null finding, we next attempted to address the issue by visually examining associations between effect size and high- and low-fertility window definitions. Figure 3 presents the high- and low-fertility window used to measure each effect that was predicted to be positive—namely, each effect assessing cycle shifts in women’s preferences for hypothesized cues of (ancestral) genetic quality in a short-term or unspecified context. Effects are presented in ascending order by effect size. We reasoned that, if true cycle shifts are absent, and the (spurious) cycle shifts observed resulted from researchers selecting whichever high- and low-fertility windows produced significant findings, larger effects would be associated with a) more variable high- and low-fertility window definitions, b) more poorly placed high- and low-fertility windows (high-fertility windows that included true low-fertility days of the cycle and/or low-fertility windows that included true high-fertility days of the cycle) , and c) less frequent use of a continuous fertility variable, which circumvents the problem of window definition flexibility because all cycle days are included in the analysis. Although a visual analysis cannot replace rigorous statistical tests of associations between effect size and high- and low-fertility window definitions, it is noteworthy that Table 14 does not reveal obvious evidence of the pattern just described; smaller and larger effects do not appear to differ in a, b, or c.

Finally, we conducted an analysis examining cycle shifts in women’s preferences for all hypothesized cues of genetic quality but limited the analysis to those studies that used a continuous fertility variable. As noted above, we reasoned that, if cycle shifts observed in the full sample resulted from researcher degrees of freedom in high- and low-fertility window definitions, these cycle shifts would not be robust in the subsample of effects that is less vulnerable to this problem (although we cannot definitively rule out the possibility that researchers chose *post hoc* to use a continuous fertility variable because this yielded positive findings). We conducted this analysis, first, in the sample of effects that included a broad set of mate preference measures and, then, in the sample that included a narrow set of measures.

The first, broad sample included 31 effects from 12 studies. The weighted mean *g* across contexts was small to moderate ($g = 0.26$, $SE = 0.12$) and borderline statistically significant ($p = .05$). The weighted mean *g* in a short-term context was small ($g = 0.17$, $SE = 0.11$) and fell short of statistical significance ($p = .14$); the weighted mean *g* in an unspecified relationship context was moderate to large ($g = 0.62$, $SE = 0.17$) and statistically significant ($p = .004$); and the weighted mean *g* in a long-term context was near-zero ($g = -0.03$, $SE = 0.11$) and not statistically significant ($p = .77$). Comparing the three contexts revealed that the weighted mean *g* was significantly larger in a short-term context than in a long-term context and in an unspecified context than in a long-term context ($p = .005$ and $.003$, respectively). The weighted mean *g* was also significantly larger in an unspecified context than in a short-term context ($p = .01$). This difference is likely due to the influence of several particularly large positive effects included in the unspecified subsample of effects (e.g., studies examining women's preferences for scents associated with symmetry) and one large *negative* effect included in the short-term subsample (Morrison et al., 2010).

The second, narrow sample included 20 effects from nine studies. The weighted mean *g* across contexts was small to moderate ($g = 0.38$, $SE = 0.13$) and statistically significant ($p = .02$). The weighted mean *g* in a short-term context was small to moderate ($g = 0.29$, $SE = 0.12$) and statistically significant (*p* = .04); the weighted mean *g* in an unspecified relationship context was moderate to large (*g* = 0.62, *SE* = 0.16) and statistically significant (*p* = .005); and the weighted mean *g* in a long-term context was near zero (*g* = 0.03, *SE* = 0.11) and not statistically significant (*p* = .81). Comparing the three contexts revealed that the weighted mean *g* was significantly larger in a short-term context than in a long-term context and in an unspecified context than in a long-term context (*p* = .002 and .009, respectively). The weighted mean *g* did not differ between a short-term and an unspecified context (*p* = .12). Thus, results were largely consistent with those observed in the full samples of effects.

In sum, we used multiple procedures to assess and adjust for various forms of potential bias. The results of these procedures do not suggest that these sources of bias account for the robust cycle shifts observed in this meta-analysis.

Discussion

Summary of Meta-analysis Findings

We evaluated evidence for the ovulatory shift hypothesis in a large sample of published and unpublished effects and found clear support for the predicted pattern of relationship context-dependent cycle shifts in women's mate preferences. Women exhibited a stronger preference for characteristics widely thought to have reflected genetic quality in ancestral males on high-fertility days of the cycle as compared with low-fertility days of the cycle. However, this cycle shift depended on the type of relationship for which women evaluated a prospective partner. Women exhibited a robust cycle shift in their preferences for hypothesized cues of ancestral genetic quality when they evaluated men or male stimuli as prospective partners for a short-term relationship (e.g., a one-night stand) or evaluated the attractiveness of male stimuli or desirability of male characteristics without reference to a specific relationship context. In contrast, women exhibited no such cycle shift when they evaluated men or male stimuli as prospective partners for a long-term relationship (e.g., marriage). Likewise, women did not exhibit a cycle shift in their preferences for characteristics widely thought to have reflected suitability as a long-term social partner and coparent in ancestral males in any relationship context. This pattern of cycle shifts was robust across both a broad sample of effects that included a diverse set of male characteristics and measures of mate preferences and a narrow sample of effects that included only those characteristics and measures that we reasoned would provide a particularly strong test of the predicted cycle shifts. Furthermore, importantly, the observed cycle shifts do not appear to be accounted for by an underrepresentation of small effects in the meta-analysis sample (as could result from publication bias) or by researcher degrees of freedom in definitions of high- and low-fertility cycle phases.

We conducted more focused analyses to examine cycle shifts in women's preferences for specific characteristics hypothesized to have indicated genetic quality in ancestral men. Many of these analyses were conducted on small samples of effects, and in such cases, results should be considered preliminary. Among the specific characteristics we examined, body masculinity and behavioral dominance showed the strongest support for the pattern of cycle shifts predicted by the ovulatory shift hypothesis. Analyses revealed a significant and marginally significant cycle shift in women's preference for body masculinity and behavioral dominance, respectively, in a short-term relationship context, no cycle shift in a long-term relationship context, and a significant difference in the magnitude of this cycle shift comparing a short-term to a long-term context. Analyses examining cycle shifts in preferences for facial symmetry and vocal masculinity hinted at a similar pattern, but the predicted cycle shifts fell short of statistical significance. However, given that these analyses were underpowered, more data are needed to make any confident claims about the presence or absence of cycle shifts in women’s facial symmetry and vocal masculinity preferences.

Analyses examining cycle shifts in women’s preference for facial masculinity revealed partial support for the ovulatory shift hypothesis. Analyses revealed a significant cycle shift in attractiveness ratings made without reference to a specific type of relationship and no cycle shift in a long-term context. However, analyses did not reveal a cycle shift in a short-term context (where a cycle shift was predicted). Removing two studies that used potentially problematic measures of women’s facial masculinity preferences revealed a small, though still not statistically significant, cycle shift in a short-term context. Ultimately, more data are needed to determine whether this unexpected pattern of results is robust and in need of explanation or reflects the influence of idiosyncratic features of the particular studies included in this analysis.

Lastly, analyses examining cycle shifts in women’s preferences for scents associated with symmetry and facial cues associated with circulating testosterone both hinted at a cycle shift in attractiveness ratings made without reference to a specific type of relationship, but these cycle shifts fell short of statistical significance. However, these analyses were underpowered, so again, more data are needed to make any confident claims about the presence or absence of cycle shifts in these preferences.

Interpreting Differences in Statistical Significance across Contexts and Characteristics

This meta-analysis revealed differences in the magnitude of cycle shifts across relationship contexts and across specific male characteristics, raising the question of how to properly interpret these differences. Interpreting a single statistically significant cycle shift—for example, the high-fertility increase in short-term body masculinity preferences—is straightforward: although possible, the probability that a cycle shift of this magnitude and level of statistical significance is accounted for by chance alone is very low, and thus it is conventional to infer that the cycle shift is probably real. Likewise, given a statistically significant difference between relationship contexts in the magnitude of a given cycle shift—for example, the difference between a short-term and long-term relationship context in the magnitude of the cycle shift in body masculinity preferences—we can also straightforwardly conclude that the probability that this apparent context effect is accounted for by chance alone is very low.

In contrast, it is less clear how to properly interpret null effects and comparisons between null and statistically significant effects. For example, analyses revealed a nonsignificant cycle shift in women's short-term vocal masculinity preferences that was, nonetheless, comparable in magnitude to the statistically significant cycle shift in women's short-term body masculinity preferences. One possible interpretation of this pattern of statistical significance is that women's preferences for body masculinity shift across the cycle, whereas their preferences for vocal masculinity do not. If the ovulatory shift hypothesis is correct, this could indicate that body masculinity reflected genetic quality ancestrally, whereas vocal masculinity did not. However, importantly, several other possibilities are equally consistent with this pattern of statistical significance. For example, it is possible that the body masculinity analysis was sufficiently powered to detect a cycle shift, whereas the vocal masculinity analysis was not (and, in fact, the body masculinity analysis included three times as many effects as did the vocal masculinity analysis). It is also possible that researchers manipulated or measured body masculinity with greater precision than they manipulated or measured vocal masculinity, that participants were able to perceive variation in body masculinity in male body photos or drawings with greater acuity than they were able to perceive variation in vocal masculinity in vocal recordings, or that studies examining preferences for body masculinity incidentally used more rigorous methods (e.g., for determining women's position in the ovulatory cycle) than did studies examining preferences for vocal masculinity. Ultimately, in the case of null effects, especially those produced by analyses that are likely to have been underpowered, additional studies are needed to test for the presence and magnitude of cycle shifts. In summary, whereas statistically significant effects indicate the likely presence of real phenomena deserving of explanation, null effects based on small numbers of effects indicate a need for more evidence.

Limitations

The focal analyses examining cycle shifts in women's preferences for all characteristics hypothesized to have reflected genetic quality in ancestral males contained many effects and produced a clear pattern of results supporting the ovulatory shift hypothesis. However, a common limitation of the more focused analyses examining cycle shifts in preferences for specific male characteristics—for example, vocal masculinity, scents associated with symmetry, and facial cues of testosterone—was a lack of sufficient statistical power. Therefore, although the overall pattern of results was typically consistent with the ovulatory shift hypothesis, the meta-analysis findings do not compel firm conclusions regarding the robustness of cycle shifts in preferences for these or other specific characteristics.

In addition, although many analyses revealed significant unexplained between-studies variation in the magnitude of cycle shifts, the moderation analyses revealed few and somewhat inconsistent associations between study characteristics and effect size. A possible explanation is that studies in this meta-analysis varied in so many ways that there was simply too much noise to observe true moderation effects. In addition, despite substantial methodological heterogeneity in the sample as a whole, there often was not enough variation on specific moderators to obtain a precise estimate of their effect. For example, only three of the 50 studies that contributed effects to the analysis examining cycle shifts in preferences for all hypothesized cues of ancestral genetic quality (broad sample of effects) used luteinizing hormone tests to verify the timing of ovulation, though this method is widely regarded in this area as one of the most rigorous for assessing cycle position. Therefore, moderation analyses examining associations between the use of this particular method and the magnitude of cycle shifts (or between the use of this method and the moderating effect of relationship context on cycle shifts) were underpowered. We emphasize that these null findings do not indicate that methodological rigor has no association with effect size; rather, there currently is an absence of evidence for such associations.

Several moderators did emerge across both the broad and narrow samples of effects as being significantly or marginally significantly associated with the pattern of cycle shifts predicted by the ovulatory shift hypothesis. Studies that used scent stimuli, used direct measurement to determine the amount of the characteristic of interest possessed by the male stimuli, or were published generally showed larger predicted cycle shifts after controlling for the effect of relationship context. In addition, studies conducted outside of the lab (usually online) generally showed larger predicted cycle shifts in a short-term relationship context relative to a long-term relationship context. In contrast (contrary to the predictions of the ovulatory shift hypothesis), studies in which participant ratings were used to determine the amount of the characteristic of interest possessed by the male stimuli generally showed smaller predicted cycle shifts after controlling for the effect of relationship context. Importantly, the moderation analyses tested for associations between study characteristics and effect size, rather than for causal relationships. Nonetheless, the results provide preliminary insight into the kinds of studies that might be better at capturing true context-dependent cycle shifts in mate preferences if they are present.

Also important, the finding that predicted cycle shifts were generally larger in published studies than in unpublished studies is consistent with several possible, non-mutually exclusive interpretations. One possibility is that the mean effect size within the published literature overestimates the true magnitude of cycle shifts. Upward bias in effect size among published studies could reflect a tendency among reviewers, journal editors, or researchers themselves to evaluate papers that report positive findings as more worthy of publication than papers that report null or negative findings simply by virtue of the fact that they provide support for the hypothesis in question. It is important to note that any such tendency did not result in a detectable underrepresentation of small effects in the meta-analysis sample as a whole (see funnel plots above). Another possibility is that the mean effect size within the unpublished literature underestimates the true magnitude of cycle shifts. Downward bias in effect size among unpublished studies could reflect a tendency among reviewers, journal editors, or researchers to evaluate papers that report positive findings as more worthy of publication than papers that report null or negative findings, not because they provide support for the ovulatory shift hypothesis but rather because these studies actually used more rigorous methods or otherwise provided more precise tests of predicted cycle shifts. In sum, publication status appears to be an additional source of between-studies variation in cycle shift magnitude, but this finding should be interpreted with due caution.

An additional limitation of this meta-analysis is that the results cannot provide insight into whether women find high levels of a given characteristic particularly attractive at high fertility, find low levels of a given characteristic particularly aversive at high fertility, or both. This limitation is in fact not unique to this meta-analysis but, rather, is a limitation of many of the studies in the meta-analysis sample—including, for example, all studies that used a forced-choice or slider task to assess women's preference for a characteristic. In order to accommodate the large number of these studies in the meta-analysis sample, we selected an effect size that does not differentiate between the above possibilities.

Lastly, in general, meta-analyses evaluate the strength and robustness of effects in an empirical literature, rather than provide a direct test of the hypothesis of interest. Thus, this meta-analysis provides a test of the ovulatory shift hypothesis only to the extent that the set of empirical findings it synthesized provided a test of that hypothesis. Given the challenges of estimating and verifying women's position in the ovulatory cycle (see Appendix A), it is likely that some studies included in this meta-analysis provided a relatively weak test of the ovulatory shift hypothesis. Therefore, the weighted mean effect sizes we report here could be conservative estimates of the true effect sizes. Despite these issues and other limitations, the findings of the focal analyses examining cycle shifts in women's preferences for all hypothesized cues of ancestral genetic quality offer clear support in the extant empirical literature for the pattern of cycle shifts predicted by the ovulatory shift hypothesis.

Strengths

The focal analyses examining cycle shifts in preferences for hypothesized cues of ancestral genetic quality included large numbers of effects from unpublished studies (e.g., 34 of the 96 effects in the analysis that included a relatively broad set of mate preference measures) obtained through a variety of methods (e.g., listserv posts). Although unpublished studies have often yielded null results, the key analyses revealed cycle shifts that were robust across the entire sample of published and unpublished effects. Furthermore, funnel plots and trim and fill procedures did not provide compelling evidence that the statistically significant cycle shifts observed in this meta-analysis could be accounted for by an underrepresentation of small effects.

In addition, we used several procedures to assess whether the statistically significant cycle shifts observed in this analysis appeared to result from bias in researchers' definitions of high- and low-fertility cycle phases but did not find evidence of such bias. Thus, publication bias and researcher degrees of freedom in high- and low-fertility definitions do not appear to account for the cycle shifts observed in this meta-analysis.

Another strength of this meta-analysis is that we used multilevel meta-analytic methods. This enabled us to include multiple effects from the same study in a single analysis, while properly accounting for the non-independence of these nested effects. It also enabled us to test cross-level interactions among effect- and study-level predictors, for example, to identify study characteristics that moderated relationship context effects.

Lastly, we used carefully designed inclusion criteria to create two samples of effects: a relatively heterogeneous, "broad" sample of effects that we reasoned would capture the diversity of mate preference measures used in this literature and a relatively homogeneous, "narrow" sample of effects that we reasoned would provide a relatively strong test of the ovulatory shift hypothesis. In fact, in an earlier version of this manuscript, we had reported results based only on the narrow sample. However, in response to suggestions from reviewers, we subsequently relaxed the inclusion criteria twice to create two broader samples. We report the broader of these two samples here. Although the pattern of cycle shifts predicted by the ovulatory shift hypothesis was somewhat stronger in the narrow sample, it remained robust in both of the broader samples. This indicates that the pattern of cycle shifts observed in this meta-analysis is not a mere artifact of the particular inclusion criteria that we used to select the initial, narrow sample of effects.

Convergent Evidence for Cycle Shifts in Mating Motivations

The key findings of this meta-analysis are consistent with a growing body of research supporting the overarching idea that women's mating-related motivations, preferences, cognitions, and behaviors shift near ovulation, leading to systematic changes across the ovulatory cycle. For example, other lines of work have documented cycle shifts in women's attractions to their relationship partners and other individuals (e.g., Larson, Pillsworth, & Haselton, 2012), opportunistic orientation toward sex (Gangestad et al., 2010a)**,** evaluations of their relationship partner's flaws and virtues and feelings of closeness and satisfaction with their partners (Larson, Haselton, Gildersleeve, & Pillsworth, 2013), preferences for attractive and revealing clothing (Haselton, Mortezaie, Pillsworth, Bleske-Rechek, & Frederick, 2007; Durante, Li, & Haselton, 2008), interest in attending events where they might meet potential partners (Haselton & Gangestad, 2006), and receptiveness to others' attempts to initiate romantic involvements with them (Guéguen, 2009a, 2009b).

The body of research examining cycle shifts in women's attractions to men other than their primary partners is particularly relevant to the idea that women's mate preferences shift across the cycle. This line of research aims to test the prediction that, if women's primary partner is relatively lacking in the characteristics they particularly prefer at high fertility—namely, characteristics thought to have reflected genetic quality in ancestral males—they will experience an increase at high fertility relative to low fertility in their attraction to other men (presumably, men who possess higher levels of these characteristics). Consistent with this idea, across five studies, the extent to which women reported experiencing greater extra-pair attraction (attraction to men other than their primary partner) at high fertility relative to low fertility depended on their partner's sexual attractiveness or on the extent to which their partner possessed specific characteristics thought to have reflected genetic quality in ancestral men (e.g., partner sexual attractiveness, Pillsworth & Haselton, 2006a; partner sexual attractiveness relative to investment attractiveness, Haselton & Gangestad, 2006; partner facial masculinity, Gangestad et al., 2010b; facial masculinity and partner facial attractiveness [marginally significant], Gangestad et al., 2010b; composite partner face and body attractiveness, Larson, Pillsworth, & Haselton, 2012). Furthermore, in several studies, women's reports of their partner's mate retention behavior (e.g., jealousy, possessiveness, and attentiveness) increased at high relative to low fertility, (Gangestad et al., 2002; Haselton & Gangestad, 2006), and this effect appeared to depend on the extent to which their partner possessed characteristics that women are thought to particularly prefer at high fertility (Haselton & Gangestad, 2006; Pillsworth & Haselton, 2006a). These findings are consistent with the notion that, as ancestral females evolved psychological mechanisms that produced cycle shifts in mate preferences, males coevolved psychological mechanisms that facilitated behaviors that mitigated the risk of a mate engaging in extra-pair sex at high fertility.

Suggested Directions for Future Research

The existence of robust ovulation-related changes in women's mate preferences across the ovulatory cycle highlights a number of interesting and potentially illuminating avenues for future research and theory in this area. First, it is not yet known whether cycle shifts in women's mate preferences represent the output of psychological mechanisms that have been favored by selection in human evolutionary history or psychological mechanisms that were favored by selection in an ancestral species but are vestigial in humans. Therefore, the specific conditions that initially gave rise to and have maintained or modified the psychological mechanisms posited to produce cycle shifts in women's mate preferences are not yet well-understood. A phylogenetic analysis could help to shed light on the precise evolutionary pathways that gave rise to the posited psychological adaptations. In addition, if these psychological mechanisms initially evolved in an ancestral species, theoretical and empirical work could help to clarify how these mechanisms have since been modified in the context of high rates of pairbonding among humans (Gangestad & Garver-Apgar, 2013).

Second, future research should seek to identify the hormonal mechanisms underlying cycle shifts in women's mate preferences. Previous research has suggested several possible candidates for hormonal mediators of such cycle shifts. For example, two studies have found a positive association between women's measured estradiol levels within the ovulatory cycle and their preferences for facial cues of testosterone in men (Roney & Simmons, 2008; Roney, Simmons, & Gray, 2011). In addition, several studies have used women's position within the ovulatory cycle to estimate their hormone levels and have found a negative association between women's estimated progesterone levels and preferences for scents associated with symmetry and vocal masculinity (Garver-Apgar, Gangestad, & Thornhill, 2008; Puts, 2005), a positive association between women's estimated luteinizing hormone and follicle stimulating hormone levels and preference for dominance in a short-term sex partner (Lukaszewski & Roney, 2009), and a positive association between women's estimated levels of testosterone and preference for facial masculinity (Welling et al., 2007). It is possible that all of these hormones play a role in shifts in women's mate preferences across the cycle or that a particular hormone, such as estradiol, is the primary hormone driving cycle shifts. Ultimately, research directly measuring each of these potential hormonal mediators is needed to better address the question of which hormonal mechanisms underlie cycle shifts.

Third, future research should examine the impact of cycle shifts in women's mate preferences on long-term relationship functioning and longevity. As noted above, several lines of work suggest that women whose long-term partners possess relatively low levels of the characteristics women find most attractive at high relative to low fertility might be particularly likely to experience a cycle shift in their attraction to other men (e.g., Haselton & Gangestad, 2006), in their satisfaction with their current partner (Larson et al., 2012), and in their partner's mate retention behaviors toward them (e.g., Haselton & Gangestad, 2006; Pillsworth & Haselton, 2006a), potentially leading them to experience increased conflict with their partner or other changes in their relationship in the fertile period of the cycle. What remains unknown is whether such changes completely resolve, allowing relationships to return to their prior state after each fertile period, or have a cumulative effect on relationship functioning and longevity. Furthermore, it remains unknown how hormonal contraceptive use, pregnancy, menopause, and other factors that dramatically alter or eliminate cyclic variation in women's hormones impact relationship functioning and longevity. Given the important and far-reaching implications of these questions, rigorous research is needed to examine the long-term impacts of cycle shifts on long-term relationships.

Fourth, research in this area has primarily involved Western samples of educated young women. Overreliance on such samples is common throughout psychology and not unique to this research area (Henrich, Heine, & Norenzayan, 2010). Nonetheless, future research should examine variation in the robustness and magnitude of cycle shifts in mate preferences in other ecologies and cultural contexts. For example, as a result of having more frequent pregnancies and breastfeeding for longer periods of time, women in traditional, "natural-fertility" populations experience far fewer ovulatory cycles than do women in Western populations (see Lancaster & Alvarado, 2010). Among the Dogon of Mali, for example, women have about 100 ovulatory cycles in their lifetime, compared with an estimated 400 lifetime ovulatory cycles among

American women (see Strassman, 1997). This raises the question of whether women who have relatively few ovulatory cycles in their lifetime experience cycle shifts in mate preferences similar to those experienced by women who have relatively many ovulatory cycles, such as the women included in this meta-analysis. Furthermore, it remains unknown whether the behavioral effects of these cycle shifts vary across different populations. Are women who experience relatively few ovulatory cycles in their lifetime more or less likely to act on their shifting desires?

Lastly, as noted above, there is not yet an established set of conventions for how to best design studies to measure ovulatory cycle shifts. At present, there is considerable variation in the methods researchers use to examine cycle shifts (see Appendix A), including in whether researchers (a) use a between- versus within-participants design, (b) obtain hormonal confirmation of women's ovulatory cycle position versus estimate women's cycle position based on a "counting method," (c) estimate women's cycle position based on a forward versus reverse counting method, (d) base estimates of cycle position solely on participants' retrospectively recalled or predicted dates of menstrual onset versus dates of menstrual verified during the course of the study, (e) treat fertility as continuous by assigning each woman a conception probability estimate from actuarial tables versus treat fertility as dichotomous by defining discrete high- and low-fertility cycle phases, and so on. An important task for future research is to empirically evaluate these methods and their relative strengths. For example, it is reasonable to argue that studies that track women over time, obtain verified dates of menstrual onset, and use hormone tests to confirm ovulation within the fertile period provide some of the most precise tests of ovulatory cycle shifts. However, using such methods is very costly. A key question, therefore, is how simpler methods—for example, a between-participants design, requiring only women's retrospectively recalled date of menstrual onset—compare with more rigorous methods.

Notably, the majority of the studies included in this meta-analysis used counting methods that rely on women's reports of retrospectively recalled or predicted dates of menstrual onset to estimate their position in the ovulatory cycle. Given the ease with which these methods can be used, they are likely to continue to be popular. As noted above, among studies using counting methods to estimate women's position within the ovulatory cycle, there is considerable variation in the cycle days researchers have defined as high and low fertility (see Figure 3). Ideally, researchers will work to establish a convention about the best days to include in these windows. However, a straightforward alternative, which we recommend, is to treat fertility as continuous by assigning each woman a conception probability estimate based on actuarial tables (Wilcox et al., 2001). By eliminating the opportunity to select among different high- and low-fertility windows that produce somewhat different results, this method helps to alleviate concerns that any observed statistically significant cycle shifts reflect researcher degrees of freedom.

Conclusions

Over the past two decades, there has been a surge of interest in examining systematic shifts in women's mate preferences across the ovulatory cycle, with dozens of empirical papers examining these and related effects and many more referencing the work. This meta-analysis shows that there is robust support in the extant published and unpublished empirical literatures for the pattern of relationship context-dependent cycle shifts in women's preferences predicted by the ovulatory shift hypothesis. Although this meta-analysis answers the important empirical question of whether these cycle shifts are robust, it also highlights a number of unresolved issues to be addressed by future theory and research, as noted above. Nonetheless, the findings of this meta-analysis have important implications for understanding the ultimate evolutionary and proximate causes of systematic day-to-day variation in women’s attractions, motivations, and social relationships.

Chapter 3: Are Women More Attractive at High Fertility? A Meta-analytic Review

In nonhuman mammals, the high-fertility period of the ovulatory cycle is often marked by changes in females' social interactions with males. For example, a common and prominent change at high fertility is that females receive increased sexual interest from males at this time (Beach, 1976). What triggers these changes? Ovulation is brought about by a series of hormonal events. These hormonal events could contribute to outward physical and behavioral changes in females that others could potentially detect (briefly reviewed below). Therefore, a plausible explanation for the social changes that often accompany the high-fertility period is that males (and possibly also other females) detect and respond to these changes in females (e.g., see Beach 1976; Huchard & Cowlishaw, 2011).

But what about women? The past two decades have witnessed a surge of scientific interest in the question of whether the high-fertility phase of the ovulatory cycle is accompanied by changes in women that others could potentially detect—so-called "fertility cues." Following from research on nonhuman animals, much of this research has focused specifically on investigating whether there are attractive physical and behavioral changes in women at high fertility relative to low fertility and on whether men's responses to women at high fertility relative to low fertility are consistent with increased mating effort toward those women (e.g., increased testosterone levels; Miller & Maner 2010a, b). In total, over 50 published studies to date have examined a variety of changes in women that could serve as fertility cues or others' behavioral and hormonal responses to these cues.

Despite the empirical attention they have received, questions regarding whether women's high-fertility phase is accompanied by detectable cues, what those cues are, and how others respond to those cues have not yet received clear answers. Effect size estimates in this literature have ranged from very small to quite large (see Haselton & Gildersleeve, 2011), and several published studies using rigorous methods have not replicated key effects (e.g., cycle shifts in women's facial attractiveness; Bleske-Recheck, 2011). Moreover, many additional pertinent studies remain unpublished. Studies in this area have also used a wide variety of methods and measures (similar to the literature on cycle shifts in mate preferences, see Chapter 2). This heterogeneity has impeded attempts to evaluate and summarize the evidence in the form of a traditional narrative review.

This work has potentially important implications for understanding the role of evolved psychological mechanisms and cycling reproductive hormones in human attraction, fertility, sexual behavior, relationship dynamics, and other important social phenomena. Therefore, we conducted a meta-analysis to quantitatively synthesize and evaluate current evidence for a number of possible fertility cues in women, as well as evidence that others detect and respond discernibly differently to cues of high versus low fertility. Our meta-analysis focused on the effects that have been most studied to date—namely, differences between the high- and low-fertility phases of the ovulatory cycle in women's other-rated attractiveness, self-rated attractiveness, engagement in attractive or attractiveness-enhancing behaviors (including flirtatiousness and receptivity), and men's hormonal and behavioral responses to the cycling women. In the following we briefly review evidence for fertility cues in nonhuman mammals before presenting the meta-analysis methods and results.

Cycle Shifts in Female Attractiveness in Nonhuman Mammals

In nonhuman mammals, the high-fertility period of the ovulatory cycle is often accompanied by changes in females across one or more channels. For example, some females experience changes in their physical appearance at high fertility, including the development of prominent genital swellings or subtle changes in facial coloration (e.g., Dixson, 1983, 1998; Higham et al., 2011). Changes in female scents—usually inferred on the basis of male responses to these scents—are also common (see Gildersleeve, Haselton, Larson, & Pillsworth, 2012). Research has also documented changes in females' voices at high fertility. For example, in some primate species, female vocalizations following copulation ("copulation calls") at high fertility differ in acoustic structure from calls at low fertility (e.g., Buesching, Heistermann, Hodges, & Zimmerman, 1998; Maestripieri & Roney, 2005). Research has also identified a variety of behavioral changes in females at high fertility, including increased receptivity to males, active solicitation of males for sex, increased grooming behavior, scent marking, and "trill" calling (a vocal utterance that is unique to the high-fertility period for some mammals; e.g., see Beach, 1976; Buesching, Heistermann, Hodges, & Zimmerman, 1998; Ferkin, Lee, & Leonard, 2004).

The magnitude of changes accompanying the high-fertility period is highly variable across species. In some species, these changes are very bold. This includes, for example, the genital swellings of female chimpanzees. In other species, these changes are very subtle. For example, female rhesus macaques experience systematic changes in facial coloration across the ovulatory cycle; however, these changes are sufficiently subtle that only males who are well acquainted with the females can detect them (Higham et al., 2011). Whether these cues were shaped via evolutionary processes to *signal* female fertility or are mere byproducts of the hormonal events that precede ovulation but nonetheless provide reliable information relevant to assessing a female's fertility is still being debated (e.g., Domb & Pagel, 2001), and the correct answer likely differs across species and possible even across cues within a given species.

What is clear is that males are generally attracted to changes in females that accompany the high-fertility period, often responding with enhanced sexual interest, increased hormone levels (e.g., testosterone and luteinizing hormone), and stereotypic mating behaviors (e.g., erection and attempts at copulation; see Gildersleeve, Haselton, Larson, & Pillsworth, 2012). Therefore, setting aside the question of what the function of these cues is for females, these findings suggest that males possess adaptations that enable them to detect fertility cues and respond in ways that could facilitate mating with females precisely when conception is most probable (see Reichert, Heistermann, Hodges, Boesch, & Hohmann, 2002). The empirical literature now contains a large number of studies relevant to evaluating whether, in humans, findings are also consistent with this idea.

Methods Used in Past Research

Studies examining cycle shifts in women's attractiveness have used a wide variety of methods that largely overlap with those that have been used to examine cycle shifts in women's mate preferences. Therefore, for a detailed exposition of methods used to study cycle shifts, see Gildersleeve, Haselton, & Fales (2014; Chapter 2 of this dissertation). Here, we briefly describe typical methods that have been used to study fertility cues in women.

Most studies examining cycle shifts in women's attractiveness use a within-participants design, wherein each woman provides responses or stimuli (e.g., samples of their natural body scent) at both high and low fertility within the cycle. In such studies, researchers typically estimate women's position in the cycle in advance of their participation (e.g., in an intake interview) by using the forward or reverse counting method. They then schedule sessions to fall within predicted high- and low-fertility windows. Less commonly, studies use a between-participants design, wherein each woman provides responses or stimuli at only high or low fertility. In such studies, researchers typically estimate women's cycle position after their participation by using the forward counting method. They then typically categorize each woman as high- or low-fertility on the basis of whether her session appeared to have fallen within a designated high- or low-fertility window, and they exclude from the analysis women whose sessions fell outside of both windows. In both types of studies, women sometimes also complete tests of luteinizing hormone or other hormones to verify their fertility status.

Many studies involve men or both men and women rating stimuli collected from women at high and low fertility (e.g., facial or body photographs, body odor samples, vocal recordings, videotaped behavior, etc.). Typically, raters evaluate stimuli one at a time for attractiveness or other qualities. Importantly, analyses comparing ratings of high- and low-fertility stimuli can treat stimulus donors or raters as the "unit of analysis," and the results of these two types of analyses afford different interpretations and generalizations. As explained in greater detail below, all effects included in this meta-analysis were based on analyses that treated stimulus donors as the unit of analysis.

Finally, much less commonly, studies involve raters completing a series of two-option forced-choice trials, in which they indicate which stimulus from a pair of high- and low-fertility stimuli collected from the same woman is more attractive or possesses more of some other quality of interest. Analyses performed on this type of data yield effect sizes in the form of mean proportions—for example, the mean proportion of trials on which women's high-fertility scent samples were chosen as more attractive than their low-fertility samples. Mean proportions cannot be combined with the standardized difference effect size metrics used for studies that collected individual ratings of high- and low-fertility stimuli. Therefore, such effects were excluded from the focal analyses, though a subset were analyzed separately (explained below).

Meta-analysis Methods

Search Strategy

As shown in Table 14, we identified a large number of studies that examined differences between the high- and low-fertility phases of the ovulatory cycle in women's physical features, behaviors, and in men's responses to the cycling women. Following the methods of Gildersleeve, Haselton, and Fales (2014), we located published studies and unpublished manuscripts and data through a variety of channels, including reference sections of published articles, online databases and search engines (e.g., PsycINFO, PubMed Central, Web of Science, BIOSIS, Dissertation Abstracts Online, ProQuest Dissertations & Theses, and Google Scholar), conference proceedings, listserv postings, personal correspondence with researchers in this area, and the annual conference programs of the Society for Personality and Social Psychology (2005 to 2012) and of the Human Behavior and Evolution Society (2000 to 2012). We also sent solicitations for manuscripts and data via the Society Psychology Network listserv and printed a solicitation in the Summer 2010 Human Behavior and Evolution Society newsletter.

Database searches used Boolean logic to search for entries including a term related to ovulation, the menstrual cycle, fertility, or cycling hormones, as well as a term related to attractiveness or some channel through which detectable cues of high fertility would plausibly be emitted – for example, "ovulat*" *or* "mid-cycle" *or* "menstrual cycle" *or* "cycl*" *or* "fertil*" *or* "high-fertility" *or* "conception risk" *or* "hormon*" *or* "follicular" *or* "estrogen" *or* "estradiol" *and* "attract*" *or* "appearance" *or* "fac*" *or* "body" *or* "scent" or "odor" *or* "voice" *or* "vocal" *or* "walk" *or* "gait" *or* "dance" *or* "body motion" *or* "flirt*" *or* "receptive*" *or* "dress" *or* "cloth*" *or* "red" *or* "makeup" *or* "cosmetic" *or* "hair" *or* "groom*" (asterisk allows the search to retrieve multiple possible word endings; e.g., "cloth*" would retrieve both "clothing" and “clothes”). We discontinued our literature search in May 2014.

Inclusion Criteria

Studies were deemed eligible for inclusion in the meta-analysis if they met the following inclusion criteria:

Criterion 1: Naturally-cycling Women. The study must have included only naturally-cycling women—by which we mean reproductive-aged women not using hormonal contraception—or collected information about hormonal contraception use so that it was possible to examine naturally-cycling women’s data separately from other women’s data.

Criterion 2: Assessed Ovulatory Cycle Position. The study must have collected information that could be used to estimate participants’ position in the ovulatory cycle (e.g., date of last menstrual onset or a positive or luteinizing hormone test results).

Criterion 3. Assessed One or More Variables of Interest. The study must have used women’s self-reports, others’ ratings, or direct measurement or observation to assess one or more of the following: attractive physical features in women, broadly defined (e.g., facial attractiveness or vocal attractiveness); an established correlate of women’s attractiveness (e.g., higher voice pitch is an established correlate of women’s vocal attractiveness); attractive behaviors, attractiveness-enhancing behavior, or apparent attractiveness-enhancement motivations (e.g., dancing attractively, using cosmetics, flirting, or desiring to go shopping); or others’ responses to women that suggest that they perceive those women as attractive, are directing increased mating effort toward those women, or are otherwise responding in ways consistent with fertility cue detection (e.g., giving larger tips to women exotic dancers or behaving jealously toward a female romantic partner).

Our data search sometimes retrieved studies that assessed variables that were clearly of interest per the explanation above, as well as additional variables that were not initially of interest but could nonetheless potentially serve as fertility cues. For example, in addition to assessing mean vocal pitch (of interest because it is a known correlate of vocal attractiveness), Bryant and Haselton (2009) assessed vocal pitch variability, which has not yet been linked to attractiveness. For thoroughness and in order to make the data available to others who might wish to examine cycle shifts in these other variables, we included all such effects in Table 14.

Criterion 4: Provided Information to Compute Appropriate Effect Size. The study must have provided information to compute an effect size in the form of Hedges's *g*. In this meta-analysis, *g* represents the standardized mean difference between high and low fertility in women's attractiveness or another variable of interest. For example, a *g* of 0.2 would indicate that women's attractiveness is, on average, two tenths of a standard deviation higher at high fertility than at low fertility. Hedges's *g* differs from Cohen's *d* only in that it includes an adjustment to reduce bias in small samples. Therefore, it shares the same interpretation (Borenstein, Hedges, Higgins, & Rothstein, 2009). In psychology, it is conventional to interpret effect sizes of 0.2, 0.5, and 0.8 as small, moderate, and large, respectively (Cohen, 1988).

Two researchers coded all effect sizes independently and then resolved discrepancies by checking the original article or data file. If reported information was insufficient to compute Hedges's *g*, we requested the needed information from the study authors. If the author did not respond to our request or could not provided the needed data, we listed those data as missing and excluded them from analysis (see Table 14).

As noted above, four studies produced effect sizes of a form that could not be converted to Hedges's *g.* These studies used a two-option forced-choice task to assess differences in raters' perceptions of stimuli collected from women at high versus low fertility and therefore produced effect sizes in the form of mean proportions (Bleske-Recheck, 2011; Bobst & Lobmaier, 2012; Haselton, Mortezaie, Pillsworth, Bleske-Recheck, & Frederick, 2007; Roberts et al., 2004). Mean proportions cannot be combined with standardized mean differences. Given that these effects comprised a small minority of the meta-analysis sample, we excluded them from the focal analyses. However, as explained below, we conducted a separate analysis to compute the weighted mean proportion across effects assessing cycle shifts in women's facial attractiveness.

Finally, several articles reported analyses treating raters as the unit of analysis, rather than stimulus donors as the unit of analysis. Statistically significant results from analyses treating raters as the unit of analysis can be expected to generalize to other possible samples of raters. Therefore, these analyses address the question of whether new raters are also likely to show a preference for high-fertility over low-fertility stimuli *in the same or a very similar set of stimuli*. In contrast, statistically significant results from analyses treating stimulus donors as the unit of analysis can be expected to generalize to other possible samples of stimulus donors. Therefore, these analyses address the question of whether new stimulus donors are also likely to produce high-fertility stimuli that would be preferred over their low-fertility stimuli so long as they are rated *by the same or a very similar sample of raters.*

To illustrate why this distinction is important, one could imagine a study designed to examine men's perceptions of women's high- versus low-fertility body scents that included a large sample of men as raters but only a handful of women as scent stimulus donors. If one stimulus donor smoked cigarettes at low fertility but did not inform the researcher (scent studies typically forbid participants from smoking during odor collection), this could have a major impact on the study's results. However, this would depend on whether analyses treated raters or stimulus donors as units of analysis.

Assuming that raters generally perceive cigarette smoke as unpleasant, an analysis treating raters as units of analysis could easily reveal a statistically significant tendency for men to evaluate women's high-fertility scent samples as more attractive than their low-fertility samples. This significant finding would not, however, reflect a true fertility effect but rather the undue influence of the contaminated sample (which was a low-fertility sample purely by coincidence). In contrast, an analysis treating stimulus donors as units of analysis would (correctly) be unlikely to reveal a statistically significant tendency for women to receive higher attractiveness ratings at high than at low fertility. Thus, analyses treating raters as units of analysis are vulnerable to problems introduced by extraneous (non-fertility-linked) variation in the attractiveness of high- versus low-fertility stimuli, particularly when studies include small samples of stimulus donors.

In light of these issues, and given that the primary aim of this meta-analysis was to determine whether naturally-cycling women, in general, are more attractive at high fertility (or differ in other ways or in the responses that they elicit from others), all Hedges's *g*s were computed based on analyses that treated stimulus donors (women) as units of analysis. If an article only reported analyses treating raters as units of analysis, we asked the authors to perform new analyses treating stimulus donors as units of analysis or to allow us to reanalyze their data. If this was not possible, we excluded those studies (Maner & McNulty, 2013).

Analyses

We used a multilevel meta-analytic approach in order to properly include multiple non-independent effects from the same study within a single analysis. As is conventional, we weighted each effect by its inverse variance, giving more precise effects more "pull" on weighted mean effect sizes and regression coefficients (Raudenbush & Bryk, 2002). We estimated fixed effects and variance components using restricted maximum likelihood estimation procedures, which tend to reduce downward bias in variance components as compared with full maximum likelihood estimation procedures (O'Connell & McCoach, 2008). We conducted all analyses in HLM 7.0 using the weighting and known variance options.

As shown in Table 14, we conducted analyses on relatively broad, heterogeneous samples of effects, as well as on narrower, more homogeneous samples of effects. For all analyses, we specified an unconditional random-effects model to compute the weighted mean *g* as an estimate of the "true" mean *g* in the corresponding population of effects (as a reminder, *g* represents the standardized mean difference between high and low fertility in women's attractiveness or another variable of interest).

Level-1 model (effects):

$g_{ij} = \delta_{0j} + e_{ij}$

Level-2 model (studies):

$\delta_{0j} = \gamma_{00} + u_{0\mathrm{j}}$

In the model above, g_{ij} is the observed standardized mean difference *i* for study *j*, δ_{0j} is the "true" mean *g* in the population of effects, e_{ij} is the sampling error associated with g_{ij} as an estimate of δ_{0j}, γ_{00} is the observed weighted mean *g* in the sample of effects, and $u_{0\mathrm{j}}$ is a study-level random error. Specifying δ_{0j} as random entails conceiving of *g* as varying randomly over the population of studies. This allows *g* to vary both as a function of sampling error and as a function of true between-studies variance (see Raudenbush & Bryk, 2002). This approach is generally recommended and is appropriate here, given that the studies in this meta-analysis differed from each other in a number of ways. Finally, some studies included multiple measures of interest or multiple experimental conditions. For such studies, when possible, we used a composite of the pertinent measures and collapsed across conditions, respectively.

Results

Other-rated Attractiveness

As shown in Table 14, the first analysis examined cycle shifts in women's other-rated physical attractiveness, broadly defined—specifically, other's ratings of women's general attractiveness or the pleasantness, sexiness, or sexual attractiveness of their natural body scents. We limited the analysis to attractive features that are unlikely to be under women's volitional control—for example, we included measures of natural body scent attractiveness but excluded measures of clothing sexiness—in order to separate out attractive physical changes from attractive behavioral changes. There were 10 such effects in 10 studies (total $N = 234$). The weighted mean g estimating the "true" mean standardized mean difference between high and low fertility in women's other-rated attractiveness was small ($g = 0.20$, $SE = 0.06$) but statistically significant ($p = .007$). Thus, women's other-rated attractiveness was 0.20 of a standard deviation higher at high fertility than at low fertility, on average.

Some studies that did not collect attractiveness evaluations nonetheless examined cycle shifts in features that are known to contribute to women's physical attractiveness. Therefore, we ran a follow-up analysis including all effects assessing other-rated attractiveness as described above or an established correlate of attractiveness—namely, lower waist-to-hip ratio (e.g., see Singh, 1993a; Furnham, Moutafi, & Baguma, 2002) or higher voice pitch (e.g., see Feinberg et al., 2008b). Both studies examining changes in women's voice pitch across the cycle measured women's voice pitch from a social utterance (e.g., a sentence introducing themselves or free speech) and from a non-social utterance (vowel sounds; see Bryant & Haselton, 2009; Fischer et al., 2011). In order to focus on physical changes in women's voice that are unlikely to be under women's volitional control, we limited this analysis to measures of vocal pitch from non-social utterances but examined social utterances in a later analysis. In total, this analysis included 14 effects in 13 studies (total *N* = 343). The weighted mean *g* in this sample was small (*g* = 0.13, *SE* = 0.06) but borderline significant (*p* = .05).

The next few analyses examined cycle shifts in specific measures of other-rated attractiveness or known correlates of attractiveness. These analyses included more homogeneous sets of effects but were often underpowered. Therefore, they provide preliminary insight into which information channels might carry particularly strong fertility cues but also point to a general need for more research on these specific cues.

The first of these analyses examined cycle shifts in women's other-rated body scent attractiveness—specifically, scent pleasantness, sexiness, sexual attractiveness, or a composite of these measures. If a study reported measures of scent pleasantness and sexiness, we used the effect examining scent sexiness given that past research has shown that high-fertility female scents are often sexually attractive (also, the two tend to be very highly correlated; see Gildersleeve et al., 2012). For one within-participants study, authors were unable to provide the correlations between high- and low-fertility scent ratings that we needed to compute Hedges's *g* effect sizes that account for the nonindependence of high- and low-fertility observations nested within women (Doty, Ford, Preti, & Huggins, 1975). Therefore, to compute effect sizes for this study, we imputed correlations from a similar scent study included in this meta-analysis (*r* = .67 for pleasantness, *r* = .65 for intensity; Havlíček, Dvořáková, Bartoš, & Flegr, 2005).

Across eight effects in eight studies, there was a small but significant cycle shift, indicating that women's natural body scents are 0.25 of a standard deviation more attractive at high than at low fertility, on average (*g* = 0.25, *SE* = 0.08, *p* = .02, total *N* = 183). For exploratory purposes, we also examined cycle shifts in women's scent intensity. The weighted mean *g* was -0.12, indicating that women's body scents were generally rated as somewhat *less* intense at high fertility as compared with low fertility; however, this effect did not reach statistical significance (*SE* = 0.09, *p* = .24, seven effects in seven studies, total *N* = 162).

The next analysis examined cycle shifts in women's other-rated facial attractiveness. Because all such effects were in the form of mean proportions (proportion of high-fertility facial photos chosen over low-fertility facial photos of the same women), we conducted a random-effects analysis using Comprehensive Meta-analysis software. Across three effects from three studies, the weighted mean proportion was somewhat higher than chance (0.56 vs. chance of 0.50, *SE* = 0.05, total *N* = 67), and this difference was marginally significant (*p* = .08).

Lastly, we examined cycle shifts in women's voice pitch as measured from non-social utterances (vowel sounds). Across three effects in two studies, analyses did not reveal any cycle shift in voice pitch (*g* = -0.06, *SE* = 0.06, *p* = 0.31, total *N* = 92). In general, however, more research is needed to compel any firm conclusions about the presence or absence of cycle shifts in these specific components and correlates of women's other-rated physical attractiveness.

Self-rated Attractiveness

The next analysis examined cycle shifts across a variety of measures of women's self-rated physical attractiveness and sexual desirability—specifically, general attractiveness, facial attractiveness, body attractiveness, sexiness, and desirability as a "short-term mate" (e.g., a sexual affair partner or one-night stand), and composites of these variables (see Table 14). We did not include measures of women's self-rated desirability as a "long-term mate" (e.g., a long-term relationship or marriage partner) in this analysis because such measures commonly include items assessing self-perceived kindness, ambition, intelligence, and other qualities that are likely desirable in a long-term partner but are not clearly linked to physical attractiveness or sexual desirability. Therefore, there are not clear predictions about why women's self-perceived long-term mate value would shift across the cycle.

In total, the analysis examining cycle shifts in women's self-rated attractiveness included 15 effects from 14 studies (total $N = 638$). The weighted mean g estimating the true population mean g was small ($g = 0.14$, $SE = 0.05$) but statistically significant ($p = .01$). Thus, women rated themselves as 0.14 of a standard deviation more attractive at high fertility than at low fertility, on average.

Next, we examined cycle shifts in more specific measures of self-rated attractiveness. For women's self-rated facial attractiveness, the weighted mean g was 0.18 ($SE = 0.11$, $p = .18$, five effects in five studies, total $N = 339$). For body attractiveness, the weighted mean g was 0.02 ($SE = 0.10$, $p = .81$, five effects in five studies, total $N = 340$). For sexiness, the weighted mean g was 0.09 ($SE = 0.03$, $p = 0.03$, seven effects in seven studies, total $N = 394$). And for short-term mate value, the weighted mean g was -0.01 ($SE = 0.06$, $p = .83$, three effects in three studies, total $N = 185$). Therefore, cycle shifts were generally weak, with the exception of the small but robust cycle shift in women's self-rated sexiness. However, ultimately, more research is needed in order to make any confident claims about the presence or absence of cycle shifts across these specific measures of self-rated attractiveness.

Attractive and Attractiveness-Enhancing Behavior

Whereas the first set of analyses focused on attractive possible fertility cues that are unlikely to be under women's volitional control (e.g., changes in vocal pitch in non-social utterances, which presumably result from the direct actions of hormones on the vocal chords), the next set of analyses focused on attractive behavioral changes in women that presumably reflect shifts in women's underlying mating motivations, as well as more direct measures of such motivations. Specifically, this very broad analysis included self-rated, other-rated, or direct observations of women's use of red clothing and accessories (thought to signal mating motivation; see Beall & Tracy, 2013); clothing provocativeness or attractiveness; provocativeness or attractiveness of clothing that women drew on a female figure or that women selected for a female figure on a computerized clothing selection task; women's preference to purchase sexier clothing, sexier accessories, and attractiveness-enhancing products on a computerized shopping task; the amount of, attractiveness of, and time women spent applying cosmetics; the amount of time women spent to style their hair, take care of their hygiene, decide what to wear, and make themselves beautiful; women's use of diet and exercise to enhance their appearance, desire to go to a hairdresser, and desire to shop; women's walk sexiness; women's dance sexiness; and women's voice pitch as measured from social utterances. Across 28 effects in 31 studies (total $N = 2{,}349$), the weighted mean g was small ($g = 0.28$, $SE = 0.05$) but statistically significant ($p < .001$). Thus, women engaged in attractive behavior, attractiveness-enhancing behavior, or exhibited an apparent motivation to engage in such behaviors 0.28 of a standard deviation more at high fertility than at low fertility, on average.

The next few analyses examined cycle shifts in specific attractive behaviors and attractiveness-enhancing behaviors and apparent motivations. First, we examined cycle shifts among studies in which women "purchased" (in a simulated shopping task), selected, or drew an illustration of clothing or accessories that they would hypothetically like to wear. The weighted mean g was small but marginally significant, suggesting a stronger preference for sexy clothes at high relative to low fertility ($g = 0.11$, SE = 0.05, $p = 0.09$, nine effects in nine studies, total $N = 594$). Next, we examined cycle shifts in the sexiness, revealingness, or provocativeness of the clothes woman actually wore on the day of their session. This effect was slightly larger and also marginally significant (g = 0.31, SE = 0.12, *p* = .08, four effects in four studies, total *N* = 155). Thus, both analyses suggest an increase in women's preference for sexy clothes at high fertility.

Next, we examined cycle shifts in women's body motion and vocal pitch. Across two effects examining cycle shifts in women's walk attractiveness, the weighted mean *g* was 0.84 (*SE* = 0.14, two effects in two studies, total *N* = 129). Notably, this was one of the largest mean effect size observed in this meta-analysis; however, because it was based on only two effects, there was not enough power to determine its statistical significance. Across two effects examining cycle shifts in women's vocal pitch as measured from a social utterance, the weighted mean g was 0.17 (*SE* = 0.05, two effects in two studies, total *N* = 92). Again, power as insufficient to test the statistical significance of this effect.

The next analysis examined cycle shifts in women's flirtatiousness and receptivity to men's attempts to initiate romantic involvements with them. This sample included measures of women's observed acceptance of an attractive male confederate's request to dance (at a dance club) or of an attractive male confederate's request to give him her phone number (on a public street), women's behavioral engagement (response speed) with the male host of a simulated game show, women's other-rated verbal and nonverbal flirtation in a videotaped speed-date-like interaction with a male confederate, women's self-rated flirtatiousness, and women's self-reports of their recent flirting behavior with men. In total, the analysis included 8 effects in 8 studies (total *N* = 522). The weighted mean g was small (*g* = 0.22, *SE* = 0.06) but statistically significant (*p* = .01), indicating that women were generally 0.22 of a standard deviation more flirtatious or receptive at high fertility as compared with low fertility, on average.

Men's Hormonal and Behavioral Responses to Women

Lastly, we examined cycle shifts in men's responses to women. This analysis included women's reports of their male romantic partners' mate retention behaviors (see Gangestad et al., 2002), men's measured testosterone responses to possible scent cues of fertility, and the amount men tipped female lap dancers. Across five effects in five studies (total $N = 74$), analyses revealed a moderate cycle shift of $g = 0.52$ ($SE = 0.17$), and this effect was statistically significant ($p = 0.04$). This suggests that men respond to high-fertility women more than to low-fertility women with an increase in testosterone and these mating-related behaviors.

Next, we examined more specific responses within this category. The first analysis included just two effects examining cycle shifts in men's jealousy as reported by their female romantic partner (total $N = 50$) and should be considered very preliminary. The weighted mean g was 0.30 ($SE = 0.25$), but a reliable p value could not be computed. The second analysis also included just two effects examining differences between men's testosterone response to women's high- versus low-fertility body scents (total $N = 13$). The weighted mean g was 1.05 ($SE = 0.72$), but a reliable p value could not be computed. More research is needed to make confident claims about the robustness of either effect.

Can Bias Account for the Observed Cycle Shifts?

An important question is whether the statistically significant weighted mean cycle shifts reported above have been artificially inflated by publication bias, selective reporting, or other sources of bias that might produce an underrepresentation of small effects in the meta-analysis sample. As is conventional in meta-analysis, we used funnel plots to assess and adjust for any such bias for each of the key overall analyses presented above. In funnel plots, each effect size is plotted against its standard error, with effect size increasing from left to right and precision increasing from bottom to top. This typically creates a funnel shape, with effects measured at greater precision zeroing in on the true mean effect size. An absence of effects in the lower lefthand portion of the funnel indicates that small effects measured at low precision—precisely those effects that are least likely to have reached statistical significance—are still missing from the meta-analysis sample.

The trim and fill procedure can be used to correct for this bias. This procedure uses nonparametric methods to estimate the number of missing effects and how the inclusion of these putative missing effects would change the overall mean effect size. The trim and fill procedure has been shown to substantially improve point estimates of mean effect size, as well as confidence intervals around those estimates (see Duval & Tweedie, 2000).

As shown in Figures 4 - 6, we created funnel plots for each of the three key analyses reported above—namely, the analyses examining cycle shifts in women's other-rated attractiveness; self-rated attractiveness; and attractive behavior, attractiveness-enhancing behavior, and apparent motivations to engage in such behaviors. As shown in Figure 4, the funnel plot revealed evidence of a slight underrepresentation of small effects in the analysis examining cycle shifts in women's other-rated attractiveness. Imputing two adjusted effects (filled circles), as recommended by the trim and fill procedure, resulted in a small reduction in the weighted mean effect size (from g = .20 to .16). Nonetheless, the effect remained statistically significant (95% CI [0.07, 0.25]).

As shown in Figure 5, the funnel plot did not reveal any bias in the analysis examining cycle shifts in women's self-rated attractiveness. Thus the small but statistically significant weighted mean effect we observed is unlikely to be accounted for by an underrepresentation of small effects in our sample.

As shown in Figure 6, the funnel plot revealed considerable bias in the analysis examining cycle shifts in women's attractive behavior and attractiveness-enhancement. Imputing seven adjusted effects (filled circles), as recommended by the trim and fill procedure, reduced the weighted mean effect size from *g* = 0.29 to 0.21. Thus, the weighted mean effect we reported above was upwardly biased by an underrepresentation of small effects in our sample, though notably, the effect remained statistically significant even after adjusting for this bias (95% CI [0.11, 0.30]).

Discussion

This meta-analysis set out to synthesize and evaluate the evidence for changes in women's attractiveness across the ovulatory cycle. The focal analysis revealed strong support in the extant empirical literature for a small but robust increase in women's attractiveness at high relative to low fertility. Across diverse measures of women's other-rated attractiveness and features known to contribute to attractiveness (e.g., higher vocal pitch), analyses revealed a small but statistically significant cycle shift. Although funnel plot and trim and fill procedures detected evidence that this cycle shift had been slightly inflated by an underrepresentation of small effects in our sample, the effect remained significant after adjusting for this bias.

Analyses also revealed support for an increase in women's perceptions of their own attractiveness at high relative to low fertility. Across diverse measures of self-rated attractiveness, analyses revealed a small but statistically significant cycle shift. Importantly, funnel plot and trim and fill procedures revealed no evidence that this effect had been inflated by an underrepresentation of small effects in our sample.

We also conducted a number of analyses to examine cycle shifts in specific measures of women's other- and self-rated attractiveness; however, the results of these analyses were largely inconclusive due to small samples of studies using any given measure. As an exception, a large subsample of studies examined cycle shifts in women's natural body scent attractiveness and revealed a small but significant increase at high relative to low fertility. A large subsample of studies also examined cycle shifts in women's self-rated sexiness and revealed a small but significant increase at high relative to low fertility; however, analyses did not reveal similar cycle shifts in women's self-rated facial attractiveness, body attractiveness, or short-term mate value. Thus, more research is needed to clarify which components of other- and self-rated attractiveness drove the robust overall cycle shifts that we observed.

Although it was not the primary focus of this meta-analysis, we evaluated the evidence for additional cycle shifts implicated by previous findings in this literature. An analysis across diverse measures of women's attractive behavior and attractiveness-enhancing behavior and apparent motivations revealed a small but statistically significant cycle shift, indicating a robust increase at high relative to low fertility. Three large subsamples of studies within this category examined women's preference to "purchase" sexy clothes and accessories and attractiveness-enhancing products on a simulated shopping task, examined the sexiness or revealingness of women's actual clothing worn on the day of their session, or examined women's flirtatiousness and receptivity. These analyses all revealed marginally significant or significant increases at high relative to low fertility and might, therefore, have played a large role in driving the overall effect. More research is needed to examine cycle shifts in the other measures included in this sample.

Notably, although the cycle shift remained statistically significant even after adjusting for this bias, funnel plot and trim and fill procedures suggested that the observed overall cycle shift in the attractiveness-enhancement analysis had been substantially inflated by an underrepresentation of small effects in our sample. As noted above, we made special efforts to retrieve all pertinent published and unpublished data (e.g., using listserv posts to request unpublished data and student projects). It is unclear why, then, small effects were systematically missing from our sample. One possibility is that researchers feel uncertain about or are simply less likely to keep track of data showing null or negative findings and therefore are less likely to share these data with meta-analysts. Whatever the explanation, the findings of all analyses related to attractiveness-enhancement should be interpreted with due caution.

We conducted one final set of analyses to examine cycle shifts in men's hormonal and behavior responses to women. Analyses revealed an overall significant increase at high relative to low fertility in responses that might facilitate mating-related behaviors (e.g., an increase in testosterone). Follow-up analyses of specific measures within this category suggested possible increases at high relative to low fertility in men's testosterone in response to scent cues of fertility and in men's possessive and jealous behavior toward their romantic partner, but in both cases, power was insufficient to test the statistical significance of those effects. Therefore, these findings are provocative, but more research is needed in this area.

In sum, the findings of this meta-analysis indicate that the high-fertility period of the human ovulatory cycle is accompanied by attractive physical and behavioral changes in women and, possibly, in the hormonal and behavioral responses that women elicit from others. Thus, these findings support the notion that the ovulatory cycle plays an important—yet still often overlooked—role in human social behavior. These findings could have far-reaching implications for understanding social motivations and behavior in a variety of domains but are particularly relevant to understanding mating behavior. For example, the existence of a robust cycle shift in other-rated attractiveness suggests that women's romantic partners (and perhaps also other individuals with whom women have frequent contact) might experience systematic day-to-day variation in their attractions to their partner that—unbeknownst to them—is tied to their partner's cycling reproductive hormones and fertility. Furthermore, this variation in their feelings of attraction could be mediated in part by changes in women's body scent attractiveness, a component of attractiveness that has often been overlooked in research on human mating.

Evidence for the existence of robust cycle shifts in women's attractive behavior, attractiveness-enhancement motivations, flirtatiousness, and related behaviors reinforces the notion that the ovulatory cycle could influence women's interactions with current or prospective mates, with possible consequences for their sexual behavior and romantic relationships. Furthermore, if women's romantic partners experience systematic shifts in their jealousy and possessiveness (an effect that is suggested though not conclusively supported by this meta-analysis), this could help to explain patterns of conflict and dissatisfaction in romantic relationships.

This meta-analysis raises a number of interesting and important questions for future theoretical and empirical work. For example, what is the adaptive function, if any, of cycle shifts in women's attractiveness? As mentioned above, cues of high fertility vary widely in their magnitude across nonhuman mammal species, and there are ongoing debates regarding the adaptive function—if any—of these cues (see Thornhill & Gangestad, 2008; Gangestad & Thornhill, 2008; Nunn, 1999).

Some researchers have argued that bold changes accompanying the high-fertility period—such as the dramatic sexual swellings of female chimpanzees—are *signals* of impending ovulation. According to this argument, selection has favored bold advertisements of high fertility in females because these advertisements entice males to mate with them when their chances of conception are highest, thus increasing their chances of fertilization and reproductive success (reviewed in Nunn, 1999). However, others have criticized this argument, noting that it is unlikely that selection would have favored the evolution of bold advertisements of high fertility in females, given that these advertisements are costly to produce and that selection will have favored adaptations in males to detect even subtle cues of high fertility (Thornhill & Gangestad, 2008).

Other researchers have argued that bold changes accompanying the high-fertility period are indeed signals but of good overall condition (e.g., health and reproductive potential), rather than impending ovulation (Domb & Pagel, 2001). According to this argument, selection has acted on females to consistently advertise good overall condition to the extent that they are able. This increases their likelihood of attracting high-quality mates. This argument typically assumes that overall condition is tied to estrogen levels. Therefore, because estrogen levels rise at high fertility, females are (incidentally) best able to advertise good overall condition at this time.

As for subtle changes accompanying the high-fertility period, such as those that appear to characterize women, some researchers have argued that these changes are indeed signals of impending ovulation. However, these signals have been designed by selection to be *narrowcast* to particular others, rather than broadcast to all others (Fales, Aktipis, & Haselton, June 2012). According to this argument, selection has favored signals of high fertility in females that are just bold enough to be detected by others permitted close contact (e.g., males preferred as potential mates) but are too weak to be detected by others denied such access (e.g., non-preferred males and female rivals).

And finally, some researchers have argued that subtle changes in females at high fertility are not signals at all. Rather, they are mere byproducts of hormonal and other changes surrounding ovulation. These “leaky cues” have not been favored by selection to serve female reproductive interests by communicating their fertile status to others (Thornhill & Gangestad, 2008; Haselton & Gildersleeve, 2011). Nonetheless, they have persisted due to biological constraints preventing their complete suppression – for example, because complete suppression might require suppression of reproductive hormones and therefore interfere with fertility. Or they have persisted simply because they have historically had little impact on women's reproductive success, and therefore selection has neither acted for nor against them.

In sum, researchers in this area have proposed a variety of possible ultimate explanations for attractive cues of high fertility in females. Importantly, not all of these explanations are mutually exclusive, different explanations could (and likely do) account for different cues, different explanations could account for different cues in different species, signals and leaky cues of fertility could co-occur within a single species (or even within a single information channel), and so on. Future work is needed to determine which of these explanations best accounts for cycle shifts in different components of women's attractiveness.

Another important question for future research is *which* information channels carry bolder or weaker fertility cues. To date, too few studies in the extant literature have examined particular channels of communication to allow for reliable comparisons between them. The findings of this analysis suggest that changes in women's scent attractiveness and social behavior might provide particularly strong cues to fertility. However, whereas changes in women's scent might be relatively difficult to detect outside of the tightly controlled lab conditions in which they have been studied, changes in social behavior might be even more pronounced when freed from the constraints of the lab environment. Further research is needed to test this idea.

In sum, this meta-analysis builds on a growing body of research suggesting that ancient psychological mechanisms sensitive to women's current fertility within the ovulatory cycle contribute to modern human sexuality, linking us with near and distant mammalian relatives. As this literature continues to expand and diversify, we expect that meta-analysis and other data synthetic techniques will play a crucial role in establishing which findings are and are not robust and in identifying areas still in need of more research. Thus, such projects will continue to guide this field to produce work that affords compelling and illuminating conclusions about the role of the ovulatory cycle in human social behavior.

Chapter 4: Experimental Evidence that Women Perceive Women's High-fertility Scents as More Attractive than their Low-fertility Scents

A growing research literature indicates that the ovulatory cycle plays an important—yet often overlooked—role in human social motivations and behavior (reviewed in Gangestad & Thornhill, 2008; and Haselton & Gildersleeve, 2011; meta-analyzed in Gildersleeve, Haselton, & Fales, 2014 [Chapter 2]; and Chapter 3 of this dissertation). Within this literature, several studies have found that men perceive women's high-fertility body scents as more attractive than their low-fertility scents (meta-analyzed in Chapter 3), consistent with the notion that men possess psychological adaptations for detecting fertility in women.

In nonhuman mammals, detecting cues of high fertility in females leads to a variety of adaptive changes in males, including increases in attempts to initiate sex with those females, male-male competition, mate guarding, and certain hormones that facilitate sex and other mating-related behaviors (see Gildersleeve, Haselton, Larson, & Pillsworth, 2012). Emerging evidence suggest that detecting cues of high fertility in women leads to similar changes in men. For example, men appear to respond to scent cues of high fertility with increased sexual attraction, as suggested by the large literature on cycle shifts in women's attractiveness (meta-analyzed in Chapter 3). In addition, recent studies have found that, compared with men who smelled a T-shirt worn by a woman at low fertility, men who smelled a T-shirt worn by a woman at high fertility responded with significantly greater cognitive access to sexual concepts (Miller & Maner, 2010b, Study 1), greater perceptions of sexual arousal (but not other emotional states) on behalf of the woman whose T-shirt they had smelled (Miller & Maner, 2010b, Study 2), and greater testosterone levels (Miller & Maner, 2010a, Studies 1 & 2). Notably, these results were based on analyses that treated men, rather than women (T-shirt donors), as the "unit of analysis."

Therefore, these findings might not generalize to other possible samples of women (for a detailed discussion of this issue, see Chapter 3). Nonetheless, these preliminary findings are provocative.

In addition, several studies have produced evidence for cycle shifts in men's responses to women in naturalistic settings. For example, one study found that naturally-cycling lap dancers received more tips from male patrons when they were at high fertility than when they were at low fertility (Miller, Tybur, & Jordan, 2007). In addition, studies examining women's reports of their male romantic partners' behavior across the cycle suggest that men are more loving and attentive but also more jealous and possessive when their partners are at high fertility (e.g., Gangestad, Thornhill, & Garver-Apgar, 2002; Haselton & Gangestad, 2006; Pillsworth & Haselton, 2006a). Although these studies cannot directly link shifts in men's behavior to their detection of fertility cues in women, the findings are nonetheless consistent with this account.

Taken together, these findings point to a potentially important role for fertility detection in male-female interactions. But what about female-female interactions? Research on nonhuman mammals suggests that adaptations for detecting fertility and that facilitate responding differently toward females on the basis of their apparent fertility status are not unique to males. For example, a recent study of chacma baboons showed that females in the high-fertility period of the ovulatory cycle were significantly more likely than females in other reproductive states (e.g., cycling but in the low-fertility period, pregnant, or lactating) to be the target of aggression from other females (Huchard & Cowlishaw, 2011). This finding implies that fertility cue detection could facilitate adaptive responses among females, much as it does among males, though these responses would likely serve quite different functions.

This line of reasoning leads to the hypothesis that women, like men, will be able to detect cues of high fertility in other women. Specifically, given that such perceptions might historically have facilitated adaptive responses—such as increased competition or mate guarding—in the presence of high-fertility women, we predict that women will perceive cues of high fertility in other women as more attractive than cues of low fertility. To date, just a few studies have produced data relevant to evaluating this hypothesis, and these studies have been limited in several ways. For example, Roberts and colleagues (2004) found that women significantly preferred women's high-fertility photos over their low-fertility photos. However, this result was based on an analysis that treated raters, rather than stimulus donors (photographed women), as the unit of analysis and therefore does not necessarily generalize to other possible samples of stimulus donors.

More closely related to the present study, Kuukasjärvi and colleagues (2004) found that women rated scent samples collected from women at mid-cycle (near ovulation) as more attractive than scents collected from women at other cycle points. However, this study included a small sample of female raters (*N* = 12); compared ratings of high- and low-fertility scent samples provided by different women, rather than using a more powerful within-participants design; relied on potentially error-prone self-reported dates of menstrual onset to estimate stimulus donors' cycle position; and used an unconventional method to standardize each donor's cycle to 28 days. In a related study, Trouton and colleagues (2012) found no preference for the scent of women's high-fertility over their low-fertility T-shirts among heterosexual (*n* = 12) or non-heterosexual (*n* = 8) female raters. However, these analyses again treated raters as the unit of analysis and were likely underpowered due to small sample sizes.

Lastly, Maner and McNulty (2013) found that, compared with women who smelled a T-shirt worn by a woman at low fertility, women who smelled a T-shirt worn by a woman at high fertility responded with significantly greater testosterone levels. However, T-shirts were provided by just four donors, and again, analyses treated the women exposed to the T-shirts as units of analysis. It is noteworthy that none of the above studies provided hormonal confirmation of stimulus donors' fertility status. Therefore, the purpose of the present study was to use rigorous methods to address the question of whether women perceive other women's high- body scents as more attractive than their low-fertility scents.

Methods

Scent Sample Collection

Twenty young, naturally-cycling women provided high- and low-fertility samples of their natural body odor. This entailed wearing cotton gauze pads under their arms for 24 hours while following strict behavioral guidelines designed to minimize the contamination of their samples by non-natural scents (e.g., deodorant use was prohibited). Fourteen of these pairs of high- and low-fertility scent samples were used as stimuli in a previous study (see "Ideal Donors," Gildersleeve, Haselton, Larson, & Pillsworth, 2012). We collected the additional six pairs of high- and low-fertility scent samples for the present study in order to ensure sufficient power for statistical analyses treating scent sample donors as units of analysis. All women completed scent collection on a high- and a low-fertility day of the cycle (a within-participants design), as verified with urine tests of luteinizing hormone and prospectively collected dates of menstrual onset. Recruitment, scheduling, and scent collection methods were identical to those used in the previous study (Gildersleeve et al., 2012). Scent samples were frozen until the rating session.

Rating Session

Prior to the rating session, we transferred the scent samples into 2 oz plastic bottles, which we left open and allowed to thaw for approximately 2 h. We then set up scent rating "stations" around a large classroom. At each station, we placed two bottles, marked "A" and

"B," containing a woman's high- and low-fertility scent samples (bottles A and B randomly assigned). Scent rating stations were spaced apart by at least a few feet and were surrounded by cardboard carrels to create privacy and minimize distraction.

Ninety-one women acted as raters. They were informed only that they would be rating gauze pads worn by women and that we were interested in the role of scent in attraction. All women rated all of the scent samples. The order in which raters evaluated different donors' scent samples was pre-randomized.

At each station, raters completed two tasks: a forced-choice between the high- and the low-fertility sample and individual ratings of each sample. For the forced-choice, participants were instructed to hold Bottle A just under their nose, being careful not to touch it to their face, and take a hearty sniff; repeat this procedure for Bottle B; and then mark on their rating form which of the two bottles—A or B—smelled more attractive to them. For the rating task, participants were instructed to smell Bottle A again, rate its scent for pleasantness, sexiness, and intensity (1 = Very Unpleasant/Very Unsexy/Very Unintense [Very Weak]; 9 = Very Pleasant/Very Sexy/Very Intense), and then complete this procedure for Bottle B. Raters completed the rating tasks at their own pace but usually finished all ratings within 30 to 60 min.

Results

We conducted two sets of analyses: one treating stimulus donors as the unit of analysis and another treating raters as the unit of analysis. Statistically significant findings from these analyses can be expected to generalize to other possible samples of stimulus donors and to other possible samples of raters, respectively (see Chapter 3). First, we computed for each stimulus donor the percentage of raters who chose her high-fertility scent sample as more attractive than her low-fertility sample, as well as the average ratings of pleasantness, sexiness, and intensity received by her high-fertility sample and received by her low-fertility sample. Next, we computed for each rater the percentage of donors whose high-fertility scent sample she preferred over the low-fertility sample, as well as the average ratings of pleasantness, sexiness, and intensity that she gave to high-fertility samples and that she gave to low-fertility samples.

Stimulus Donors as the Unit of Analysis. On average, donors' high-fertility scent samples were chosen as more attractive than their low-fertility samples by 60.5% (*SD* = 22.2%) of the raters, just bordering on being statistically significantly greater than chance (50%), $p = .05$. Compared with their low-fertility scent samples, donors' high-fertility scent samples were rated as significantly more pleasant ($M_{High} = 4.87$, $SD = 1.02$; $M_{Low} = 4.16$, $SD = 1.32$), $t(19) = 2.37$, $d = 0.60$, $p = .03$; significantly sexier ($M_{High} = 3.87$, $SD = 0.79$; $M_{Low} = 3.34$, $SD = 1.0$), $t(19) = 2.28$, $d = 0.58$, $p = .03$; and marginally significantly *less* intense ($M_{High} = 3.89$, $SD = 1.15$; $M_{Low} = 4.59$, $SD = 1.59$), $t(19) = -2.02$, $d = -0.50$, $p = .06$.

Raters as the Unit of Analysis. On average, raters chose the high-fertility scent sample as more attractive than the low-fertility scent sample for 61.1% (*SD* = 10.3%) of the donors, a rate statistically significantly greater than chance, $p < .001$. Compared with low-fertility scent samples from the same women, raters rated high-fertility samples as significantly more pleasant ($M_{High} = 4.89$, $SD = 1.01$; $M_{Low} = 4.23$, $SD = .95$), $t(90) = 12.52$, $d = 0.67$, $p < .001$; significantly sexier ($M_{High} = 3.88$, $SD = 1.37$; $M_{Low} = 3.40$, $SD = 1.22$), $t(90) = 8.72$, $d = 0.36$, $p < 0.001$; and significantly *less* intense ($M_{High} = 3.86$, $SD = 1.22$; $M_{Low} = 4.56$, $SD = 1.17$), $t(90) = -12.79$, $d = -0.58$, $p = < .001$.

Discussion

A growing body of evidence indicates that men can detect cues of high fertility in women, that they generally perceive such cues as attractive, and that this could have far-reaching implications for their attractions, motivations, and behavior (see Chapter 3). Comparatively little is known about women's detection of fertility in other women. This study used rigorous methods to address this gap and produced evidence supporting the notion that women, too, can detect certain cues of fertility in other women. Specifically, the findings indicate that women perceive other women's high-fertility natural body scents as more attractive—more pleasant, more sexy, and perhaps less intense—than their low-fertility scents. Importantly, because these findings were robust across analyses treating stimulus donors and raters as units of analysis, we can conclude that these effects are likely to replicate in other possible samples of donors and raters and are unlikely to reflect idiosyncratic features of the particular set of scent stimuli used in this study.

That women perceive other women's high-fertility scents as more attractive than their low-fertility scents suggests that women—like men and males and females of a large number of nonhuman mammal species—possess psychological adaptations for fertility detection. This raises important questions for future research. First, how do women respond to fertility cues in other women? Does detecting cues of high fertility lead to an increase in competitive motivations and behavior, and are any such effects mediated by the increase in testosterone suggested by Maner and McNulty's recent study (2013)? Focusing instead on women as cue emitters, given that ancestral females who revealed their fertile status in the presence of other females might have suffered certain costs (e.g., becoming the target of competitive aggression), do women possess adaptations for suppressing cues of high fertility in the presence of other women?

By providing initial evidence for women's detection of fertility cues in other women, this study sets the stage for future research examining the many ways in which this could impact women's motivational, behavioral, and hormonal responses to other women. Therefore, this study provides an important first step toward developing a more nuanced understanding of the role of the ovulatory cycle—and psychological adaptations sensitive to the cycle—in human social behavior.

Appendix A

Studies examining cycle shifts in women's mate preferences have varied in numerous ways. Here, we describe typical study characteristics, noting common variations.

Sample Characteristics and Study Design Basics. In most cases, samples are comprised of undergraduate students at a college or university. Less commonly, samples are comprised of community members or a combination of undergraduates and community members. In most studies to date, participants have been recruited in the US or UK (although we note exceptions in Table 2). Typically, study sessions take place in the lab. Somewhat less commonly, sessions take place in the field, by which we mean online or, very rarely, via a magazine survey with mail-in response. Either the same women complete sessions at high-fertility and low-fertility points in the cycle (a within-participants design), or different women complete sessions at high-fertility and low-fertility points in the cycle (a between-participants design).

Cycle Position Estimation Methods. Researchers have used many methods to estimate and verify women's position in the ovulatory cycle, and for the most part, there currently is not strong consensus on which of these methods or combinations of methods is most accurate and precise. Given that assessments of cycle position are crucial in determining the extent to which a study provides a precise measure of cycle shifts, we discuss these methods in detail here.

Designing a study to examine cycle shifts entails unique challenges. Researchers cannot manipulate women's hormones to mimic the naturally occurring changes that typically accompany (and drive) changes in women's fertility across the ovulatory cycle. Therefore, researchers cannot conduct a true experiment to test whether shifts in women's cycling hormones cause shifts in their mate preferences. Instead, researchers must capitalize on naturally occurring variation in women's hormones across the cycle.

In most previous studies that have examined cycle shifts in women’s mate preferences, researchers have used one of the following methods to schedule women’s sessions. In many cases, researchers have allowed participants to complete a single session at their convenience, without making any effort to have women participate on specific days of the cycle. Following their participation, researchers then used high- and low-fertility windows (usually defined in advance of data collection) to categorize women as high- or low-fertility or to exclude from analysis women who participated outside of both windows (e.g., Harris, 2010). Alternatively, researchers have sometimes assigned each woman a conception probability (fertility) value based on actuarial estimates (e.g., Wilcox, Dunson, Weinberg, Trussell, & Day Baird, 2001), thus retaining all women in the sample. Less commonly, researchers have allowed participants to complete their first session at their convenience but then scheduled subsequent sessions at some pre-designated interval (e.g., scheduling participants to complete their second session approximately two weeks after their first session; e.g., Johnston, Hagel, Franklin, Fink, & Grammer, 2001). In this case, researchers then use high- and low-fertility windows (usually defined in advance of data collection) to retrospectively categorize sessions already completed as high- or low-fertility or to exclude from analysis sessions that fell outside of both windows. Alternatively, some researchers have scheduled each participant to complete at least one session at a predicted high-fertility point and at least one session at a predicted low-fertility point based on scheduling windows defined in advance of data collection (e.g., Izbicki & Johnson, 2010, June; Garver-Apgar & Gangestad, *unpublished*).

In order to schedule sessions in high- and low-fertility windows, retrospectively categorize women or sessions as high- and low-fertility, or retrospectively assign women conception probability values, researchers must estimate each participant’s position in the ovulatory cycle.

Typically, researchers have used a “counting” method to estimate cycle position. The forward counting method typically entails identifying each participant's date of *last* menstrual onset, counting forward from that date—usually, by 14 days—to predict the date of her next ovulation, and then determining her current cycle position relative to that date (e.g., Little, Jones, Burt, & Perrett, 2007). A less common version of the forward counting method entails identifying each participant's date of last menstrual onset, counting forward from that date to her current day in the cycle, and then assigning her a conception probability value based on the number of days since her last menstrual onset (e.g., Gangestad & Thornhill, 1998). The reverse counting method entails predicting each participant’s date of *next* menstrual onset, counting back from that date—usually, by 14 or 15 days—to predict the date of her next ovulation, and then determining her current cycle position relative to that date (e.g., Puts, 2005).

Large-scale studies have found that cycle length varies substantially both between and within women. However, studies have generally found that the follicular phase of the cycle (from last menstrual onset to ovulation) is more variable in length than is the luteal phase of the cycle (from ovulation to next menstrual onset; e.g., Baird et al., 1995; Fehring, Schneider, & Raviele, 2006). Because only the reverse counting method takes into account each participant’s unique cycle length, while holding constant the length of the luteal phase (and thus reducing the influence of variability in the length of the follicular phase on cycle position estimates), some researchers have asserted that the reverse counting method provides more accurate estimates of ovulatory cycle position than does the forward counting method. Nonetheless, regardless of which method is superior, it is noteworthy that the two counting methods often produce discrepant estimates of cycle position.

Furthermore, both forward and reverse counting methods of cycle position estimation rely on a “benchmark” date of menstrual onset, which itself can introduce error into cycle position estimates. The forward counting method uses the date of last menstrual onset (preceding participation) as a benchmark. Typically, researchers obtain this date by asking participants to retrospectively recall it. The reverse counting method uses the date of next menstrual onset (following participation) as a benchmark. Researchers typically predict this date by asking participants to retrospectively recall their date of last menstrual onset and then adding either participants’ estimate of their average cycle length or a standard estimate of cycle length (e.g., 28 days) to that date.

Notably, retrospectively recalled dates of menstrual onset appear to be error-prone. For example, in one study, researchers asked women to recall their date of last menstrual onset, which had also been prospectively collected within the past 30 days as part of a related study. Although 56% reported the correct date, 18% were off by 1 day, 7% were off by 2 days, and 19% were off by 3 or more days (Wegienka & Day Baird, 2005). We are not aware of any studies assessing the accuracy of women’s self-reports of their average cycle length, but it is possible that these self-reports are also error-prone. Given that women’s high-fertility window spans only about 6 days, and fertility decreases dramatically following this window, error such as that associated with participants’ retrospectively recalled dates of last menstrual onset and self-reported average cycle lengths could seriously compromise the precision with which researchers can estimate women’s cycle position and fertility.

To safeguard against potential error introduced by women’s self-reported cycle information, some researchers have taken additional steps to ensure the accuracy of cycle position estimates. In some cases, researchers using the forwarding counting method have waited for participants to start a menstrual period (and report the start date to them) before scheduling their session(s). Similarly, some researchers using the reverse counting method have followed up with participants after they have completed the study (e.g., via regular phone calls or emails) to verify the date that they start their next menstrual period, which they then use to retrospectively categorize sessions already completed as high- and low-fertility and exclude observations collected outside of those windows (e.g., Roney et al., 2011).

Occasionally, researchers have used rigorous methods to verify ovulation, rather than relying on counting methods alone to estimate cycle position. Although rare, researchers sometimes use the salivary ferning technique, which involves asking participants' to lick or smear saliva on a slide in their predicted high-fertility window. Researchers examine the saliva under a microscope for ferning (crystallization) patterns that indicate impending ovulation (Fehring & Gaska, 1998). More commonly, researchers use the luteinizing hormone method, which involves having participants complete a series of urine tests to measure concentrations of luteinizing hormone in their predicted high-fertility window. Luteinizing hormone rises approximately 24-48 hours prior to ovulation (Testart & Frydman, 1982), and tests of LH in urine are highly accurate in verifying ovulation as detected by ultrasound (e.g., ClearBlueEasy tests are 97% accurate; Guermandi et al., 2001). Researchers using either of these ovulation verification methods typically either schedule participants' sessions using the forward or reverse counting method and later exclude from analysis those participants who did not show evidence of impending ovulation in close proximity to their putative high-fertility session and/or who showed evidence of impending ovulation in close proximity to their putative low-fertility session (e.g., Gangestad, Thornhill, & Garver-Apgar, *unpublished*) or wait for participants to show evidence of impending ovulation prior to scheduling their sessions (e.g., Peters, Rhodes, & Simmons, 2009).

In addition to the sources of variation already discussed, researchers using high- and low-fertility windows to schedule sessions or to categorize women or observations as high- and low-fertility must decide how to define these cycle phases. Given that there is not yet an established convention, the cycle shift literature is characterized by substantial variation across studies in the breadth and placement of high- and low-fertility windows. Women's fertility begins to rise three to five days before ovulation. Fertility is at a maximum on the day before and the day of ovulation and decreases dramatically the day after ovulation (Wilcox, Dunson, Weinberg, Trussell, & Day Baird, 2001). Women are effectively nonfertile for the remainder of the cycle. Thus, how a researcher defines the low-fertility window might not be very influential in determining whether a study provides a precise measure of cycle shifts. As long as the low-fertility window is not so badly misplaced or overly broad that it includes high-fertility days, it is likely to capture women on true low-fertility days. In contrast, given that women are effectively fertile on only a handful of days in the ovulatory cycle, how a researcher defines the high-fertility window is crucial in determining whether a study provides a strong test—or *any* test—of cycle shifts. A misplaced or overly broad high-fertility window can result in women completing putative "high-fertility" sessions on true low-fertility days of the cycle. This is analogous to failing to expose participants in a test condition to the manipulation and, therefore, comparing controls to controls.

To attempt to assess the precision of the various high- and low-fertility windows that have been used by the studies included in this meta-analysis, we estimated the average conception probability of the high- and low-fertility window for each study using actuarial estimates of the probability of conceiving from a single act of unprotected sexual intercourse on a given day of the ovulatory cycle. For studies that verified ovulation using luteinizing hormone tests, we contacted study authors to obtain the data needed to estimate each woman's conception probability at her high-fertility session based on her proximity to ovulation (Weinberg and Wilcox, 1995). In order to convert days from the luteinizing hormone surge to days from ovulation, we added one day (Testart & Frydman, 1982). For example, if a woman completed her high-fertility session on the same day she showed a luteinizing hormone surge, we converted this to one day before ovulation and assigned her a conception probability estimate of 0.31 (see Weinberg & Wilcox, 1995). We then averaged across all women in that sample to obtain an estimate of the average conception probability of the high-fertility window. We estimated the average conception probability of the low-fertility window based on the forward cycle days included in that window (Wilcox, Dunson, Weinberg, Trussell, & Day Baird, 2001). Likewise, for studies that used a forward or reverse counting method (without verifying ovulation), we estimated the average conception probability of the high-fertility and low-fertility windows based on the forward cycle days included in those windows (Wilcox et al., 2001). In order to convert reverse cycle days (days until next menstrual onset) to forward cycle days (days since last menstrual onset), we assumed a cycle length of 28 days. For example, if a study defined the high-fertility window as including reverse cycle days 14-21, we first converted that to forward cycle days 7-14 and then took the average of the conception probability values (Wilcox et al., 2001) associated with those eight days: (.018 + .032 + .050 + .069 + .085 + .094 + .093 + .085) / 8 = 0.068. For this example, we would interpret the average conception probability estimate of 0.068 as indicating that women in the study's high-fertility window had an estimated 7% probability of conceiving from a single act of unprotected sexual intercourse.

Although the technique just described was useful to the extent that it enabled us to compare the precision of high-fertility and low-fertility windows across studies using very different cycle position estimation methods, our estimates of the average conception probability of high-fertility and low-fertility windows are based on several assumptions that are unlikely to be fully met and should therefore be regarded with due caution. First, our method of converting days from the LH surge to days from ovulation assumes that one day passes between the LH surge and ovulation, which is not the case for all women and all cycles. Second, our method of converting reverse cycle days to forward cycle days assumes an average cycle length of 28 days, which is likewise not the case for all women and all cycles. Lastly, our method of computing the average conception probability of a given high- or low-fertility window assumes that an equal number of women participated on each day of that window, when in fact that is unlikely to be the case, perhaps especially in small samples. Therefore, our average conception probability estimates are likely to include some error, and the methods we present here are in need of further honing and validation.

Determining the Amount of a Characteristic in Male Stimuli. Researchers have used several methods to vary or measure the amount of a characteristic possessed by male stimuli. In some studies, researchers have selected stimuli that naturally vary on the characteristic of interest and then either directly measured the characteristic of interest (e.g., used calipers to measure facial masculinity), obtained reliable codes of the characteristic (e.g., trained research assistants to reliably code specific behaviors in participant videos), collected ratings of the characteristic from participants in the cycle study (e.g., asked participants in the cycle study to rate how masculine the faces looked), or collected ratings of the characteristic from a separate sample of participants. In other studies, researchers have directly manipulated stimuli to vary the characteristic of interest (e.g., morphed faces to increase or decrease symmetry) or allowed participants to directly manipulate a characteristic in stimuli.

Mate Preference Measures. Researchers have assessed women's mate preferences using numerous methods. Typically, researchers have used measures of *revealed preferences* to assess women's attraction to a given characteristic. By our definition (see Wood & Brumbaugh, 2009, for a similar but somewhat narrower definition), this entails either a) asking women to rate or choose (e.g., in a two-option or multiple-option (3 or more) forced choice) among men or male stimuli known to vary on a characteristic and then examining the association between women's reported attraction to the stimuli and the amount of the characteristic possessed by those stimuli or b) asking women to directly manipulate a characteristic until they have achieved what they perceive to be the most attractive version of a male stimulus (e.g., a facial photograph) and then examining the amount of the characteristic deemed most attractive. Less commonly, researchers have used *stated preference* measures to assess women's attraction to a given characteristic. By our definition, this entails asking women to explicitly rate the importance or desirability of a characteristic in a prospective partner. Regardless of which of these measures researchers have used, they have sometimes also asked women to make their evaluations in reference to a specific relationship context (discussed in more detail in the Methods section).

References

* Indicates studies included in the meta-analysis presented in Chapter 2.

** Indicates studies included in the meta-analysis presented in Chapter 3.

**Adams, H. M. (2010). It's that time of the month: The effect of hormonal shifts on female mate value, depressive symptomology, and short-term mating orientation. Master's Thesis, Department of Psychology, California State University, Stanislaus.

Alliende, M. E. (2002). Mean versus individual hormonal profiles in the menstrual cycle. *Fertility and Sterility, 78*, 90 – 95.

Alvergne, A. & Lummaa, V. (2010). Does the contraceptive pill alter mate choice in humans? *Trends in Ecology and Evolution*, *25*, 171 – 179.

Anderson, J. L. & Crawford, C. B. (1992). Modeling costs and benefits of adolescent weight Control as a mechanism for reproductive suppression. *Human Nature, 3,* 299 – 334.

Baird, D. D., McConnaughey D. R., Weinberg C. R., Musey P. I., Collins D. C., Kesner J. S., Knecht, E. A., & Wilcox A. J. (1995). Application of a method for estimating day of ovulation using urinary estrogen and progesterone metabolites. *Epidemiology*, *6*, 547 – 550.

Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173 – 1182.

Beach, F. A. (1976). Sexual attractivity, proceptivity, and receptivity in female mammals. *Hormones and Behavior*, *7*, 105 – 138.

**Beall, A. T. & Tracy, J. L. (2013). Women are more likely to wear red or pink at peak fertility. *Psychological Science.*

Beaulieu, D. A. (2007). *Avoiding costly mating mistakes: The evolution of female reproductive safeguards*. (Unpublished doctoral dissertation). University of California Santa Barbara, CA.

**Beaulieu, D. A. (2007). Avoiding costly mating mistakes: Ovulatory shifts in personal mate value assessment. *Journal of Social and Personal Relationships, 24,* 441 – 455.

Benacerraf, B. (1981). Role of MHC Gene Products in Immune Regulation. *Science, 212,* 1229 – 1238.

Betzig, L. (1989). Causes of conjugal dissolution: A cross cultural study. *Current Anthropology, 30,* 654 – 676.

**Bleske-Recheck, A., Harris, H. D., Denkinger, K., Webb, Rose Mary, Erickson, Leah, & Nelson, Lyndsay A. (2011). Physical cues of ovulatory status: A failure to replicate enhanced facial attractiveness and reduced waist-to-hip ratio at high fertility. *Evolutionary Psychology, 9,* 336 – 353.

**Bobst, Cora & Lobmaier, J.S. (2012). Men's preference for the ovulating female is triggered by subtle face shape differences. *Hormones and Behavior, 62,* 413 – 417.

Bollen, K. A. & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology*, *20*, 115 – 140.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to Meta-Analysis*. Sussex, United Kingdom: John Wiley & Sons, Ltd.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2008). Comprehensive meta-analysis (Version 2.2) [Computer software]. Englewood, NJ: BioStat.

Booth, A., Shelley, G., Mazur, A., Tharp, G., Kittok, R. (1989). Testosterone, and winning and losing in human competition. *Hormones and Behavior, 23,* 556 – 571.

Braude, S., Tang-Martinez, Z., & Taylor, G. T. (1999). Stress, testosterone, and the immunoredistribution hypothesis. *Behavioral Ecology, 8,* 345 – 350.

*Bressan, P. & Stranieri, D. (2008). The best men are (not always) already taken: Female preference for single versus attached males depends on conception risk. *Psychological Science, 19*, 145-151.

Brewis, A. & Meyer, M. (2005). Demographic evidence that human ovulation is undetectable (at least in pair bonds). *Current Anthropology, 46,* 465 – 471.

Bröder, A. & Hohmann, N. (2003). Variations in risk taking behavior over the menstrual cycle: An improved replication. *Evolution and Human Behavior, 24,* 391 – 398.

**Bryant, G. A. & Haselton, M. G. (2009). Vocal cues of ovulation in human females. *Biology Letters*, *5*, 12 – 15.

Buesching, C. D., Heistermann, M., Hodges, J. K., & Zimmermann, E. (1998). Multimodal oestrus advertisement in a small nocturnal prosimian, *Microcebus murinus*. *Folia Primatologica,69,* 95 – 308.

Buffet, N. C., Djakoure, C., Maitre, S. C., & Bouchard, P. (1998). Regulation of the human menstrual cycle. *Frontiers in Neuroendocrinology, 19,* 151 – 186.

Bullivant, S.B., Sellergren, S.A., Stern, K., Spencer, N.A., Jacob, S., Menella, J.A., & McClintock, M.K. (2004). Women's sexual experience during the ovulatory cycle: Identification of the sexual phase by noninvasive measurement of luteinizing hormone. *Journal of Sex Research, 41,* 82 – 93.

*Bullock, H.L. (2000). *Multiple components of facial attractiveness in humans.* (Unpublished doctoral dissertation). Queen's University, Ontario, CA.

Buss, D. M. (1988). From vigilance to violence: Tactics of mate retention in American undergraduates. *Ethology and Sociobiology*, *9*, 291 – 317.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences, 12,* 1 – 49.

Buss, D. M. (1996). Paternity uncertainty and the complex repertoire of human mating strategies. *American Psychologist, 51,* 161 – 162.

Buss, D. M. & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review, 100,* 204 – 232.

Buss, D. M. & Shackelford, T. K. (1997). From vigilance to violence: Mate retention tactics in married couples. *Journal of Personality and Social Psychology, 72,* 346 – 361.

Buss, D. M. (2002). Human mate guarding. *Neuroendocrinology Letters Special Issue, 23*, 23 – 29.

**Cantú, S. M., Simpson, J. A., Griskevicius, V., Weisberg, Y. J., Durante, K. M., & Beal, D. J. (2013). Fertile and selectively flirty: Women's behavior toward men change across the ovulatory cycle. *Psychological Science.*

*Cárdenas, R.A. & Harris, L.J. (2007). Do women's preferences for symmetry change across the ovulatory cycle? *Evolution and Human Behavior, 28,* 96 – 105.

Carrington, M., Nelson, G. W., Martin, M. P., Kissner, T., Vlahov, D., Goedert, J. J., Kaslow, R., Buchbinder, S., Hoots, K., & O'Brien, S. J. (1999). HLA and HIV-1: Heterozygote advantage and B*35-Cw*04 disadvantage.

Carter, A. J. R., Weier, T. M., & Houle, D. (2009). The effect of inbreeding on fluctuating asymmetry of wing veins in two laboratory strains of Drosophila melanogaster. *Heredity*, *102*, 563-572.

*Caryl, P. G., Bean, J. E., Smallwood, E. B., Barron, J. C., Tully, L., & Allerhand, M. (2009).

Women's preference for male pupil size: Effects of conception risk, sociosexuality, and relationship status. *Personality and Individual Differences, 46*, 503-508.

Charlesworth, D. & Charlesworth, B. (1987). Inbreeding depression and its evolutionary Consequences. *Annual Review of Ecology and Systematics, 18,* 237 – 268.

Chen, B. P., & Parham, P. (1989). Direct binding of influenza peptides to class I HLA molecules. *Nature, 337*, 743-745.

Clark, G. M. (1993). The genetic basis of developmental stability. I. Relationships between stability, heterozygosity, and genomic coadaptation. *Genetica, 89,* 15 – 23.

**Cobey, K. D., Buunk, A. P., Pollet, T. V., Klipping, C., & Roberts, S. C. (2013). Men perceive their female partners, and themselves, as more attractive around ovulation. *Biological Psychology, 94,* 513 – 516.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Second Edition. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Russell Sage Foundation.

Dabbs, J. M. Jr. & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences, 27,* 801 – 804.

Daniels, D. (1983). The evolution of concealed ovulation and self-deception. *Ethology and Sociobiology*, *4*, 69 – 87.

Dausset, J. & Colombani, J. (Eds.). (1973). Proceedings from the Colloque de l'institut national de la santé et de la recherche médicale Fifth International Histocompatibility Workshop and Conference, Evian, France, May 1972: *Histocompatibility testing 1972: Report of an international workshop and conference*. Copenhagen, Denmark: Munksgaard.

DeBruine, L. M. (2002). Facial resemblance enhances trust. *Proceedings of the Royal Society of London, B, 269,* 1307 – 1312.

DeBruine, L., Jones, B. C., Frederick, D. A., Haselton, M. G., Penton-Voak, I. S., & Perrett, D. I. (2010). Evidence for menstrual cycle shifts in women’s preferences for masculinity: A response to Harris (2011) Menstrual cycle and facial preferences reconsidered. *Evolutionary Psychology, 8,* 768 – 775.

DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2005). Women’s attractiveness judgments of Self-resembling faces change across the menstrual cycle. *Hormones and Behavior, 47,* 379 – 383.

Deschner, T., Heistermann, M., Hodges, K., & Boesch, C. (2004). Female sexual swelling size, timing of ovulation, and male behavior in wild West African chimpanzees. *Hormones and Behavior*, *46*, 204 – 215.

Diamond, L. M. & Wallen, K. (2011). Sexual-minority women's sexual motivation around the time of ovulation. *Archives of Sexual Behavior*, *40*, 237-246

Dixson, A. F. (1983). Observations on the evolution and behavioral significance of "sexual skin" in female primates. *Advances in the Study of Behavior*, *13*, 63 – 106.

Dixson, A. F. (1998). *Primate sexuality: Comparative studies of the prosimians, monkeys, apes, and human beings.* Oxford, United Kingdom: Oxford University Press.

Domb, L. G. & Pagel, M. (2001). Sexual swellings advertise female quality in wild baboons. *Nature*, *410*, 204 – 206.

Donnay, F. (2000). Maternal survival in developing countries: What has been done, what can be achieved in the next decade. *International Journal of Gynecology & Obstetrics, 70,* 89 – 97.

**Doty, R. L., Ford, M., Preti, G., & Huggins, G. R. (1975). Changes in the intensity and pleasantness of human vaginal odors during the menstrual cycle. *Science, 190,* 1316 – 1318.

Doty, R. L., Snyder, P. J., Huggins, G. R., & Lowry, L. D. (1981). Endocrine, cardiovascular, and psychological correlates of olfactory sensitivity changes during the human menstrual cycle. *Journal of Comparative and Physiological Psychology*, *95*, 45 – 60.

**Durante (2009). Intrasexual competition among women: The influence of same-sex rivals on women's purchasing and risk-taking behavior across the ovulatory cycle. Doctoral Dissertation. Department of Psychology, University of Texas, Austin.

**Durante, K. M., Li, N. P., & Haselton, M. G. (2008). Changes in women's choice of dress across the ovulatory cycle: Naturalistic and laboratory task-based evidence. *Personality and Social Psychology Bulletin, 34,* 1451 – 1460.

**Durante, K. M., Griskevicius, V., Hill, S. E., Perilloux, C., & Li, N. P. (2010). Ovulation, female competition, and product choice: Hormonal influences on consumer behavior. *Journal of Consumer Research, 37.*

**Durante, K. M., & Rae Arsena, A. (2014). Cycle shifts in women's self-perceived attractiveness. Unpublished raw data.

Duval, S. & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455 – 463.

Eastwick, P. W., & Finkel, E. J. (2008). Sex differences in mate preferences revisited: Do people know what they initially desire in a romantic partner? *Journal of Personality and Social Psychology*, *94*, 245.

Eastwick, P.W., Luchies, L.B., Finkel, E.J., & Hunt, L.J. (2013). The predictive validity of ideal partner preferences: A review and meta-analysis. *Psychological Bulletin.* Advance online publication.

Elias, M. (1981). Serum cortisol, testosterone, and testosterone-binding globulin responses to Competitive fighting in human males. *Aggressive Behavior, 7,* 215 – 224.

Ehrlich, P. R. (2000). *Human Natures: Genes, Cultures, and the Human Prospect.* New York: Penguin Books.

Fales, M.R., Aktipis, C. A. & Haselton, M. G. (June, 2012). Can a reduction in ovulation cues be the result of inter-female aggression? Testing alternative hypotheses for the evolution of concealed ovulation in a simulated ancestral environment. Conference of the Human Behavior and Evolution Society. Talk presented at the University of New Mexico, Albuquerque, NM.

Fehring, R. J., & Gaska, N. (1998). Evaluation of the Lady Free Biotester® in determining the fertile period. *Contraception*, *57*, 325-328.

Fehring, R. J., Schneider, M., & Raviele, K. (2006). Variability in the phases of the menstrual Cycle. *Journal of Obstetric, Gynecological, & Neonatal Nursing, 35,* 376 – 384.

*Feinberg, D. R. (2012). Ovulatory cycle shifts in women's preferences for vocal masculinity. Unpublished raw data.

Feinberg, D. R., Jones, B. C., Law Smith, M. J., Moore, F. R., DeBruine, L. M., Cornwell, R. E., Hillier, S. G., & Perrett, D. I (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behavior, 49,* 215 – 222.

Feinberg, D. R., DeBruine, L. M., Jones, B., C., & Perrett, D. I. (2008b). The role of femininity and averageness of voice pitch in aesthetic judgments of women's voices. *Perception, 37,* 615 – 623.

Feingold, A. (1990). Gender differences in effects of physical attractiveness on romantic attractiveness on romantic attraction: A comparison across five research paradigms. *Journal of Personality and Social Psychology, 59,* 981 – 993.

Ferkin, M. H., Lee, D. N., & Leonard, S. T. (2004). The reproductive state of female voles affects their scent marking behavior and the responses of male conspecifics to such marks. *Ethology*, *110*, 257 – 272.

Fessler, D. M. T. (2002). Reproductive immunosuppression and diet: An evolutionary perspective on pregnancy sickness and meat consumption. *Current Anthropology, 43,* 19 – 39.

*Fink, B. (2012). Ovulatory cycle shifts in women's preferences for facial masculinity. Unpublished raw data.

**Fink, B., Hugill, N. & Lange, B. P. (2012). Women's body movements are a potential cue to ovulation. *Personality and Individual Differences, 53,* 759 – 763.

**Fischer, J., Semple, S., Fickenscher, G., Jürgens, R., Kruse, E., Heistermann, M., Ofer, A. (2011). *PLoS ONE, 6,* e24490.

Fiske, P., Rintamaki, P. T., & Karvonen, E. (1998). Mating success in lekking males: A meta-analysis. *Behavioral Ecology, 9,* 328 – 338.

**Flowe, H. D., Swords, E., & Rockey, J. C. (2012). Women's behavioural engagement with a masculine male heightens during the fertile window: Evidence for the cycle shift hypothesis. *Evolution and Human Behavior, 33*, 285-290.

Folstad, I. & Karter, A.J. (1992). Parasites, bright males, and the immunocompetence handicap. *American Naturalist, 139,* 603 – 622.

*Frost, P. (1994). Preference for darker faces in photographs at different phases of the menstrual cycle: Preliminary assessment of evidence for a hormonal relationship. *Perceptual and Motor Skills, 79*, 509 – 514.

Furnham, A., Moutafi, J., & Baguma, P. (2002). A cross-cultural study on the role of weight and waist-to-hip ratio on female attractiveness. *Personality and Individual Differences, 32,* 729 – 745.

Gangestad, S. W. (1993). Human facial beauty: Averageness, symmetry, and parasite resistance. *Human Nature, 4,* 237 – 269.

Gangestad, S. W. & Buss, D. M. (1993). Pathogen prevalence and human mate preferences. *Ethology and Sociobiology, 14,* 89 – 96.

Gangestad, S.W. & Cousins, A.J. (2001). Adaptive design, female mate preferences, and shifts across the ovulatory cycle. *Annual Review of Sex Research, 12,* 145 – 185.

Gangestad, S.W. & Garver-Apgar, C.E. (2013). The nature of female sexuality: Insights into the dynamics of romantic relationships. In J.A. Simpson & L. Campbell (Eds.), *The Oxford Handbook of Relationships.* New York: Oxford University Press.

*Gangestad, S.W., Garver-Apgar, C.E., Simpson, J.A., & Cousins, A.J. (2007). Changes in women's mate preferences across the ovulatory cycle. *Journal of Personality & Social Psychology, 92,* 151 – 163.

Gangestad, S.W. & Simpson, J.A. (2000). The evolution of human mating: Trade-offs and strategic pluralism. *Behavioral and Brain Sciences, 23,* 573 – 644.

*Gangestad, S. W., Simpson, J. A., Cousins, A. J., Garver-Apgar, C. E., & Christensen, P. N. (2004). Women's preferences for male behavioral display change across the menstrual cycle. *Psychological Science, 15*, 203-207.

*Gangestad, S. W. & Thornhill, R. (1998). Menstrual cycle variation in women's preferences for the scent of symmetrical men. *Proceedings of the Royal Society of London, B, 265*, 927-933.

Gangestad, S. W. & Thornhill, R. (2008). Human oestrus. *Proceedings of the Royal Society B, 275,* 991 – 1000.

**Gangestad, S.W., Thornhill, R., & Garver, C.E. (2002). Changes in women's sexual interests and their partner's mate-retention tactics across the ovulatory cycle: Evidence for shifting conflicts of interest. *Proceedings of the Royal Society of London, B, 269,* 975 – 982.

Gangestad, S.W., Thornhill, R., & Garver-Apgar, C.E. (2005a). Women's sexual interests across the ovulatory cycle depend on primary partner developmental instability. *Proceedings of the Royal Society of London, B, 272,* 2023 – 2027.

Gangestad, S.W., Thornhill, R., & Garver-Apgar, C.E. (2005b). Adaptations to ovulation: implications for sexual and social behavior. *Current Directions in Psychological Science, 14,* 312 – 316.

Gangestad, S. W., Thornhill, R., & Garver-Apgar, C. E. (2010a). Fertility in the cycle predicts women's interest in sexual opportunism. *Evolution and Human Behavior*, *31*, 400-411.

Gangestad, S.W., Thornhill, R., & Garver-Apgar, C.E. (2010b). Men's facial masculinity predicts changes in their female partner's sexual interest across the ovulatory cycle, whereas men's intelligence does not. *Evolution and Human Behavior, 31,* 412 – 424.

*Gangestad, S. W., Thornhill, R., & Garver-Apgar (2011). Ovulatory cycle shifts in women's preferences for facial masculinity. Unpublished raw data.

*Garver-Apgar, C. E. & Gangestad, S. W. (2012). Ovulatory cycle shifts in women's mate preferences. Unpublished raw data.

Garver-Apgar, C. E., Gangestad, S. W., Thornhill, R., Miller, R. D., & Olp, J. (2006). Major histocompatibility complex alleles, sexual responsivity, and unfaithfulness in romantic couples. *Psychological Science, 17,* 830 – 835.

Geary, D. C. (2000). Evolution and proximate expression of human paternal investment. *Psychological Bulletin, 126,* 55 – 77.

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.* Advanced online publication.

**Gildersleeve, K., Haselton, M. G., Larson, C. M., & Pillsworth, E. G. (2012). Body odor attractiveness as a cue of impending ovulation in women: Evidence from a study using hormone-confirmed ovulation. *Hormones and Behavior, 61,* 157 – 166.

Gnoth, G., Frank-Herrmann, P. F., Schmoll, A., Godehardt, E., & Freundl, G. (2002). Cycle characteristics after discontinuation of oral contraceptives. *Gynecological Endocrinology, 16*, 307-317.

**Gonzales, J. E. (2010). An investigation of proceptivity, attractivity, and receptivity in women: Searching for estrus in humans. Master's Thesis. Department of Psychology, California State University, Stanislaus, CA.

Griskevicius, V., Cialdini, R. B., Kenrick, D. T. (2006). Peacocks, Picasso, and parental investment: The effects of romantic motives on creativity. *Journal of Personality and Social Psychology*, *91*, 63 – 76.

**Guéguen, N. (2009a). The receptivity of women to courtship solicitation across the menstrual cycle: A field experiment. *Biological Psychology, 80,* 321 – 324.

**Guéguen, N. (2009b). Menstrual cycle phases and female receptivity to a courtship solicitation: an evaluation in a nightclub. *Evolution and Human Behavior, 80,* 321 – 324.

**Guéguen, N. (2012). Gait and menstrual cycle: Ovulating women use sex gaits and walk slowly ahead of men. *Gait & Posture, 35,* 621 – 624.

Guermandi, E. (2001). Reliability of Ovulation Tests in Infertile Women. *Obstetrics & Gynecology, 97,* 92 – 96.

Hale, G. E., Zhao, X., Hughes, C. L., Burger, H. G., Robertson, D. M., & Fraser, I. S. (2007). Endocrine features of menstrual cycles in middle and late reproductive age and the menopausal transition classified according to the staging of reproductive aging (STRAW) staging system. *The Journal of Clinical Endocrinology & Metabolism, 92*, 3060 – 3067.

Harlow, S.D. (2000). Menstruation and menstrual disorders: The epidemiology of menstruation and menstrual dysfunction. In Goldman, M., Hatch, M. (Eds.), *Women and Health*. Academic Press, San Diego, CA, pp. 99–113.

*Harris, C.R. (2011). Menstrual cycle and facial preferences reconsidered. *Sex Roles*, *64*, 669 – 681.

**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*, 509 – 518.

Haselton, M. G. & Gildersleeve, K. (2011). Can men detect ovulation? *Current Directions in Psychological Science*, *20*, 87 – 92.

Haselton, M.G. & Miller, G.F. (2006). Women's fertility across the cycle increases the short-term attractiveness of creative intelligence. *Human Nature, 17,* 50 – 73.

**Haselton, M. G., Mortezaie, M., Pillsworth, E. G., Bleske-Rechek, A., & Frederick, D. A. (2007). Ovulatory shifts in human female ornamentation: Near ovulation, women dress to impress. *Hormones and Behavior*, *51*, 40 – 45.

**Havlíček, J., Dvořáková, R., Bartoš, L., & Flegr, J. (2005). Non-advertised does not mean concealed: Body odour changes across the human menstrual cycle. *Ethology, 112,* 81—90.

Havlíček, J. & Roberts, S.C. (2009). MHC-correlated mate choice in humans: A review. *Psychoneuroendocrinology*, *34*, 497 – 512.

Henrich, J., Heine, S. J., Norenzayan, A. (2010). The weirdest people in the world. *Behavioral and Brain Sciences, 33,* 61 – 83.

Hewlett, B.S. (1988). Sexual selection and paternal investment among Aka pygmies. In L.L. Betzig, M. Borgerhoff Mulder, & P.W. Turke (Eds.), *Human reproductive behavior: A Darwinian perspective* (pp. 263 – 276.) Cambridge, United Kingdom: Cambridge University Press.

Higham, J. P., Hughes, K. D., Brent, L. J. N., Duboc, C., Engelhardt, A., Heistermann, M., Maestripieri, D., Santos, L. R., & Stevens, M. (2011). Familiarity affects the assessment of female facial signals of fertility by free-ranging male rhesus macaques. *Proceedings of the Royal Society, B*.

Hill, E.M. (1988). The menstrual cycle and components of human female sexual behavior. *Journal of Social and Biological Systems*, *11*, 443 – 455.

**Hill, S.E. & Durante, K.M (2009). Do women feel worse to look their best? Testing the relationship between self-esteem and fertility status across the menstrual cycle. *Personality and Social Psychology Bulletin*, *35*, 1592 – 1601.

Hill, K. & Hurtado, A.M. (1989). Hunter-gatherers of the world. *American Scientist*, *77*, 437 – 443.

Hill, K. & Hurtado, A.M. (1996). *Aché life history: The ecology and demography of a foraging* *people.* Hawthorne, NY: Aldine de Gruyter.

*Hodges-Simeon, C. R., Gaulin, S. J. C ., & Puts, D. A. (2010). Different vocal parameters predict perceptions of dominance and attractiveness. *Human Nature, 21,* 406 – 427.

Hromatko, I., Tadinac, M., & Prizmic, H. (2006). Women's hormonal status and mate value influence relationship satisfaction and perceived male attractiveness. *Psychological Topics, 15*, 315-330.

Hromatko, I., Tadinac, M., & Vranic, A. (2008). Femininity and masculinity across the menstrual cycle: A relation to mate value. *Collegium Antropologicum, 32,* 81 – 86.

Hrdy, S.B. (1995b). The primate origins of female sexuality, and their implications for the role of nonconceptive sex in the reproductive strategies of women. *Human Evolution, 10,* 131 – 44.

Huchard, E. & Cowlishaw, G. (2011). Female-female aggression around mating: An extra cost of sociality in a multimale primate society. *Behavioral Ecology, 22,* 1003 – 1011.

Hughes, A. L., & Nei, M. (1988). Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature, 335*, 167 – 170.

Hughes, A. L., & Nei, M. (1989). Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection.*Proceedings of the National Academy of Sciences*, *86*, 958-962.

Hurtado, A.M. & Hill, K.R. (1992). Paternal effect on offspring survivorship among Aché and Hiwi hunter-gatherers: Implications for modeling pair-bond stability. In B.S. Hewlett (Ed.), *Father-child relations: Cultural and biosocial contexts* (pp. 31-55). Hawthorne, NY: Aldine de Gruyter.

*Izbicki, E. V. & Johnson, K. J. (2010, June). Dark, tall, and handsome: Evidence for a female increase in openness to other-race partners during periods of high conception risk. Poster presented at the annual meeting of the Human Behavior and Evolution Society (HBES), Eugene, OR.

*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,* 251 – 267.

Jones, B.C., DeBruine, L.M., Perrett, D.I., Little, A.C., Feinberg D.R., Law Smith, M.J. (2008). Effects of menstrual cycle phase on face preferences. *Archives of Sexual Behavior*, *37*, 78 – 84.

*Jones, B.C., Little, A.C., Boothroyd, L., DeBruine, L.M., Feinberg, D.R., Law Smith, M.J., Cornwell, R.E., Moore, F.R., & Perrett, D.I. (2005a). Commitment to relationships and preferences for femininity and apparent health in faces are strongest on days of the ovulatory cycle when progesterone level is high. *Hormones and Behavior, 48,* 283 – 290.

Jones, B. C., Little, A. C., Penton-Voak, I. S., Tiddeman, B. P., Burt, D. M, & Perrett, D. I. (2001). Facial symmetry and judgments of apparent health: Support for a “good genes” explanation of the attractiveness-symmetry relationship. *Evolution and Human Behavior*, *22*, 417 – 429.

Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2004). When facial attractiveness is only skin deep. *Perception*, *33*, 569 – 576.

Jones, B.C., DeBruine, L.M., Perrett, D.I., Little, A.C., Feinberg D.R., Law Smith, M.J. (2008). Effects of menstrual cycle phase on face preferences. *Archives of Sexual Behavior*, *37*, 78 – 84.

Jones, B. C., Perrett, D. I., Little, A. C., Boothroyd, L., Cornwell, R. E., Feinberg, D. R., Moore, F. R. (2005b). Menstrual cycle, pregnancy and oral contraceptive use alter attraction to apparent health in faces. *Proceedings of the Royal Society B, 272,* 347-354.

Jordan, W.C. & Bruford, M.W. (1998). New perspectives on mate choice and the MHC. *Heredity*, *81*, 127 – 133.

Keddy-Hector, A. C. (1992). Mate choice in non-human primates. *American Zoologist, 32*, 62-70.

Kirchengast, S. & Gartner, M. (2002). Changes in fat distribution (WHR) and body weight across the menstrual cycle. *Collegium Antropologicum*, *26*, 47 – 57.

Klingenberg, C. P. (2003). Developmental instability as a research tool: Using patterns of fluctuating asymmetry to infer the developmental origins of morphological integration. In M. Polak (Ed.), *Developmental instability: Causes and consequences.* Oxford University Press, New York (2003), pp. 427 – 442.

Knott, C.D., Thompson, E.M., & Stumpf, R.M. (2007). Sexual coercion and mating strategies of wild Bornean orangutans. *American Journal of Physical Anthropology, Supplement, 44,* 145.

Koehler, N., Rhodes, G., & Simmons, L.W. (2002). Are human female preferences for symmetrical male faces enhanced when conception is likely? *Animal Behaviour, 64,* 233 – 238.

*Koehler, N., Rhodes, G., & Simmons, L.W. (2006). Do cyclic changes in women's face preferences target cues to long-term healthr *Social Cognition, 24,* 641 – 656.

Koeslag, J.H. (1990). Koinophilia groups sexual creatures into species, promotes stasis, and stabilizes social behavior. *Journal of Theoretical Biology*, *144*, 15 – 35.

Knussmann, R. & Sperwien, A. (1998). Relations between anthropometric characteristics and androgen hormone levels in healthy young men. *Annals of Human Biology, 15,* 131 – 142.

Kuukasjärvi, S., Eriksson, C. J. P., Koskela, E., Mappes, T., Nissinen, K., & Rantala, M. J. (2004). Attractiveness of women's body odors over the menstrual cycle: The role of oral contraceptives and receiver sex. *Behavioral Ecology, 15,* 579—584.

Laeng, B. & Falkenberg, L. (2007). Women's pupillary responses to sexually significant others during the hormone cycle. *Hormones and Behavior*, *52*, 520 – 530.

Lancaster, J. B. & Alvarado, L. C. (June, 2010). The hormonal platform for conception in natural fertility populations: The co-occurrence of lactation, ovulation, and gestation. Conference of the Human Behavior and Evolution Society. Talk presented at the University of Oregon, Eugene, OR.

Lancaster, J. B. & Alvarado, L. C. (2010). The hormonal platform for conception in natural fertility populations: The co-occurrence of lactation, ovulation, and gestation. Paper presented at the Human Behavior and Evolution Society June 2010 Annual Conference, Eugene, Oregon.

Landolt, M. A., Lalumière, M. L., Quinsey, V. L. (1995). Sex differences in intra-sex variations in human mating tactics: An evolutionary approach. *Ethology and Sociobiology*, *16*, 3 – 23.

Landolt, M.A., Lalumière, M.L., and Quinsey, V.L. Sex differences and intra-sex variations in human mating tactics: an evolutionary approach. *Ethology and Sociobiology* 16:3-23, 1995.

Larson, C. M., Haselton, M. G., Gildersleeve, K. A., & Pillsworth, E. G. (2013). Changes in women's feelings about their romantic relationships across the ovulatory cycle. *Hormones and Behavior, 63,* 128-135.

Larson, C. M., Pillsworth, E. G., & Haselton, M. G. (2012). Ovulatory shifts in women’s attractions to primary partners and other en: Further Evidence of the importance of primary partner sexual attractiveness. *PLoS One*, *7*, p. e44456.

Lee, R. D. & Kramer, K. L. (2002). Children's economic roles in the Maya family life cycle: Cain, Caldwell, and Chayanov revisited. *Population and Development Review*, *28*, 475 – 499.

Li, N. P., Bailey, J. M., Kenrick, D. T., & Linsenmeier, J. A. W. (2002). The necessities and luxuries of mate preferences: Testing the tradeoffs. *Journal of Personality and Social Psychology*, *82*, 947 – 955.

Li, N. P., & Kenrick, D. T. (2006). Sex similarities and differences in preferences for short-term mates: What, whether, and why. *Journal of Personality and Social Psychology*, *90*, 468-489.

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*, 862 – 868

*Little, A.C., Jones, B.C., Burriss, R.P. (2007). Preferences for masculinity in male bodies change across the ovulatory cycle. *Hormones & Behavior, 51,* 633 – 639.

*Little, A. C., Jones, B. C., DeBruine, L. M. (2008). Preferences for variation in masculinity in real faces change across the menstrual cycle: Women prefer more masculine faces when they are more fertile. *Personality and Individual Differences, 45*, 478-482.

*Little, A. C., Jones, B. C., Burt, D. M., & Perrett, D. I. (2007). Preferences for symmetry in faces change across the menstrual cycle. *Biological Psychology*, *76*, 209 – 216.

Lovejoy, O. (1981). The Origins of Man. *Science, 211,* 241 – 250.

Loy, J. (1987). The sexual behavior of African monkeys and the question of estrus. In E. Zucker (Ed.), *Comparative behavior of African monkeys* (pp. 175 – 195). New York, NY: Alan R. Liss.

*Luevano, V. X. & Zebrowitz, L. A. (2006, June). *When fighters are lovers too: Predictors of attractiveness as a function of partner-type and fertility level.* Poster presented at the annual meeting of the Human Behavior and Evolution Society (HBES), Philadelphia, PA.

*Lukaszewski, A. W. & Roney, J. R. (2009). Estimated hormones predict women's mate preferences for dominant personality traits. *Personality and Individual Differences, 47*, 191-196.

Lyons, J. C. (2005). *The 2D4D ration and cycle shifts as predictors of self-perceived quality and permissiveness*. (Unpublished doctoral dissertation). University of Texas at Arlington, TX.

Mace, R. (1999). Evolutionary ecology of human life history. *Animal Behavior, 59,* 1 – 10.

Maestripieri, D. & Roney, J. R. (2005). Primate copulation calls and postcopulatory female choice. *Behavior Ecology*, *16*, 106 – 113.

Maner, J. K. & McNulty, J. K. (2013). Attunement to the fertility status of same-sex rivals: Women's testosterone responses to olfactory ovulation cues. *Evolution and Human Behavior.*

Manning, J. T., Scutt, D., Whitehouse, G. H., Leinster, S. J., & Walton, J. M. (1996). Asymmetry and the menstrual cycle in women. *Ethology and Sociobiology*, *17*, 129 – 143.

Manson, W. C., (1986). Sexual cyclicity and concealed ovulation. *Journal of Human Evolution*, *15*, 21 – 30.

Martello, M. F. D. & Maloney, L. T. (2006). Where are kin recognition signals in the human face? *Journal of Vision*, *6*, 1356 – 1366.

Matsumoto-Oda, A. (1999). Female choice in the opportunistic mating of wild chimpanzees (*Pan troglodytes schweinfurthii*) at Mahale. *Behavioral Ecology and Sociobiology, 46,* 258 – 266.

McClellan, M.E., Blackwell, A.D., Garrett, E.L., & Sugiyama, L.S. (2007, May). *Women's preferences for male body types in long-term and short-term mating contexts*. Paper presented at the annual meeting of the Human Behavior and Evolution Society (HBES), Williamsburg, VA.

McDaniel, M.A. (2005). Big-brained people are smarter: A meta-analysis of the relationship between in vivo brain volume and intelligence. *Intelligence,* 33, 337–346.

McDonald, M. M., Asher, B. D., Kerr, N. L., & Navarrete, C. D. (2011). Fertility and intergroup bias in racial and minimal group contexts: Evidence for shared architecture. *Psychological Science, 22,* 860-865.

* McDonald, M. M. & Navarrete, C. D. (2012). Ovulatory cycle shifts in women's preferences for same-race versus other-race male avatars. Unpublished raw data.

**Mikac, U. (2014). Cycle shifts in women's sexiness of dress in a computerized task. Unpublished raw data.

Miller, G.F. (2000). Sexual selection for indicators of intelligence. In G. Bock, J. A. Goode, & K. Webb (Eds.) *The nature of intelligence,* (pp. 260 – 270). Novartis Foundation Symposium 233. NY: Wiley.

*Miller, G. (2003, June). *A great sense of humor is a good genes indicator: Ovulatory cycle* *effects on the sexual attractiveness of male humor ability*. Poster presented at the annual meeting of the Human Behavior and Evolution Society (HBES), Lincoln, NE.

**Miller, J. K. & Maner, S. L. (2010a). Scent of a woman: Men's testosterone responses to olfactory ovulation cues. *Psychological Science, 21,* 276—283.

**Miller, J. K. & Maner, S. L. (2010b). Ovulation as a male mating prime: Subtle signs of women's fertility influence men's mating cognition and behavior. *Journal of Personality and Social Psychology, 100,* 295—308.

Miller, G.F. & Penke, L. (2006). The evolution of human intelligence and the coefficient of additive human brain size. *Intelligence, 35,* 97 – 114.

Miller, G. F., & Penke, L. (2007). The evolution of human intelligence and the coefficient of additive genetic variance in human brain size. *Intelligence*, *35*, 97 – 114.

**Miller, G., Tybur, J. M., & Jordan, B. D. (2007). Ovulatory cycle effects on tip earnings by lap dancers: Economic evidence for human estrus? *Evolution and Human Behavior*, *28*, 375 – 381.

Møller, A.P. & Swaddle, J.P. (1997). *Asymmetry, Developmental Stability, and Evolution*. Oxford, United Kingdom: Oxford University Press.

Møller, A.P. & Thornhill, R. (1997). A meta-analysis of the heritability of developmental stability. *Journal of Evolutionary Biology*, *10*, 1 – 16.

Møller, A.P. & Thornhill, R. (1998). Bilateral symmetry and sexual selection: A meta-analysis. *The American Naturalist*, *151*, 174 – 192.

*Moore, F. R. (2011). Ovulatory cycle shifts in women's face preferences. Unpublished raw data.

*Moore, F. R., Cornwell, R. E., Law Smith, M. J., Al Dujaili, E. A. S., Sharp, M., & Perrett, D. I.

(2011). Evidence for the stress-linked immunocompetence handicap hypothesis in human male faces. *Proceedings of the Royal Society B, 278*, 774-780.

*Morrison, E. R., Clark, A. P., Gralewski, L., Campbell, N., & Penton-Voak, I. S. (2010). Women's probability of conception is associated with their preference for flirtatious but not masculine facial movement. *Archives of Sexual Behavior, 39*, 1297-1304.

Nassaralla, C. L., Stanford, J. B., Daly, D. K., Schneider, M., Schliep, K. C., & Fehring, R. J. (2011). Characteristics of the menstrual cycle after discontinuation of oral contraceptives. *Journal of Women's Health, 20*, 169-177.

Navarrete, C. D., Fessler, D. M. T., Santos Fleischmann, D., & Geyer, J. (2009). Race bias tracks Conception risk across the menstrual cycle. *Psychological Science, 20,* 661 – 665.

Neel, J. V. & Weiss, K. M. (1975). The genetic structure of a tribal population, the Yanomama Indians. *American Journal of Physical Anthropology*, *42*, 25–51.

Nunn, C. L. (1999). The evolution of exaggerated sexual swellings in primates and the graded signal hypothesis. *Animal Behavior*, *58,* 229 – 246.

O'Connell, A. A. & McCoach, D. B. (Eds.) (2008). *Multilevel modeling of educational data.* Greenwich, CT: Information Age Publishing.

Oinonen, K.A., Klemencic, N., & Mazmanian, D. (2008). The periovulatory sociosexuality tactic shift (PSTS): Activational hormonal mechanisms in two female sexual strategies. In G.A. Conti (Ed.), *Progress in Biological Psychology Research* (pp. 139 – 158). Nova Science Publishers, Inc.

*Oinonen, K. A. & Mazmanian, D. (2007). Facial symmetry detection ability changes across the ovulatory cycle. *Biological Psychology, 75,* 136 – 145.

Oliver, M.K. & Piertney, S.B. (2006). Isolation and characterization of a MHC class II DRB locus in the European water vole (*Arvicola terrestris*). *Immunogenetics*, *58*, 390 – 395.

Pawlowski, B. (1999). Loss of oestrus and concealed ovulation in human evolution: The case against the sexual-selection hypothesis. *Current Anthropology*, *40*, 257 – 276.

*Pawlowski, B. & Jasienska, G. (2005). Women's preferences for sexual dimorphism in height depend on menstrual cycle phase and expected duration of relationship. *Biological Psychology, 70*, 38-43.

Penn, D. J. (2002). The scent of genetic compatibility: Sexual selection and the major Histocompatibility complex. *Ethology, 108,* 1 – 21.

Penn, D. J., Damjanovich, K. & Potts, W. K. (2002). MHC-heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the Royal Society, B*, *99,* 11260 – 11264.

Penton-Voak, I. S. & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior*, *25*, 229 – 241.

*Penton-Voak, I. S. & Perrett, D. I. (2000). Female preference for male faces change cyclically: Further evidence. *Evolution and Human Behavior, 21*, 39-48.

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

*Perrett, D. I., Jones, B. C., Little, A. C., Boothroyd, L., Cornwell, R. E., Feinberg, D. R., Tiddeman, S. W., Pitman, R. M., Hillier, S. G., Burt, D. M., & Stirrat, M. R. (2013). Pregnancy, the menstrual cycle, and hormonal contraceptive use alter attraction to apparent health in faces. Unpublished manuscript.

Peters, M., Rhodes, G., & Simmons, L. W. (2008). Does attractiveness in men provide clues to semen quality? *Journal of Evolutionary Biology, 21*, 572-579.

Peters, M., Simmons, L. W., & Rhodes, G. (2009). Preferences across the ovulatory cycle for masculinity and symmetry in photographs of male faces and bodies. *PLoS ONE, 4(1)*: e4138.

Pieta, K. (2008). Female mate preferences among Pan troglodytes schweinfurthii of Kanyawara, Kibale National Park, Uganda. *International Journal of Primatology*, *29*, 845-864.

Pillsworth, E. G., & Haselton, M. G. (2006a). Male sexual attractiveness predicts differential ovulatory shifts in female extra-pair attraction and male mate retention. *Evolution and Human Behavior, 27,* 247 – 258.

Pillsworth, E. G., & Haselton, M. G. (2006b). Women's sexual strategies: The evolution of long-term bonds and extrapair sex. *Annual Review of Sex Research, 17,* 59 – 100.

Pillsworth, E. G., Haselton, M.G., & Buss, D. M. (2004). Ovulatory shifts in female sexual desire. *The Journal of Sex Research*, *41*, 55 – 65.

Pipitone, R. N. & Gallup, G. G. (2008). Women's voice attractiveness varies across the menstrual cycle. *Evolution and Human Behavior*, *29*, 268 – 274.

Pitkin, R. M. (1977). Nutritional influences during pregnancy. *The Medical Clinics of North America, 61,* 3 – 15.

Poran, N. S. (1994). Cyclic attractivity of human female odours. *Advances in the Biosciences, 93,* 555—560.

Pound, N., Penton-Voak, I.S., & Surridge, A.K. (2009). Testosterone responses to competition in men are related to facial masculinity. *Proceedings of the Royal Society, B*, *276*, 153 – 159.

Prokosch, M.D., Coss, R.G., Scheib, J.E., & Blozis, S.A. (2009). Intelligence and mate choice: Intelligent men are always appealing. *Evolution and Human Behavior, 30,* 11 – 20.

Prokosch, M. D., Yeo, R. A., & Miller, G. F., (2005). Intelligence tests with higher *g*-loadings show higher correlations with body symmetry: Evidence for a general fitness factor mediated by developmental stability. *Intelligence*, *33*, 203 – 213.

*Provost, M. P., Quinsey, V. L. & Troje, N. F. (2008). Short-term mating strategies and attraction to masculinity in point-light walkers. *Evolution and Human Behavior, 29*, 65-69.

Pusey, A. & Wolf, M. (1996). Inbreeding avoidance in animals. *Trends in Ecology and Evolution*, *11*, 201 – 206.

*Puts, D. A. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evolution and Human Behavior, 26*, 388-397.

Puts, D. A. (2010). Beauty and the beast: mechanisms of sexual selection in humans. *Evolution and Human Behavior*, *31*, 157-175.

Puts, D. A., Bailey, D. H., Cárdenas, R. A., Burriss, R. P., Welling, L. L. M., Wheatley, J. R., Dawood, K. (2013). Women's attractiveness changes with estradiol and progesterone across the ovulatory cycle. *Hormones and Behavior, 63,* 13 – 19.

Rabow, M. W. (1987). The agony of defeat: The pulsatile release and response of testosterone To aggressive athletics. Honors Thesis, Department of Anthropology, Harvard University, Cambridge, MA.

*Rantala, M. J., Eriksson, C. J. P., Vainikka, A., & Kortet, R. (2006). Male steroid hormones and female preference for male body odor. *Evolution and Human Behavior, 27*, 259-269.

*Rantala, M. J., Polkki, M., & Rantala, L. M. (2010). Preferences for human male body hair changes across the menstrual cycle and menopause. *Behavioral Ecology, 21*, 419-423.

Raudenbush, S. W. & Bryk, A. S. (1985). Empirical bayes meta-analysis. *Journal of Educational and Behavioral Statistics, 10*, 75-98.

Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical Linear Models* (Second Edition). Thousand Oaks: Sage Publications.

Raudenbush, S. W., Bryk, T., & Congdon, R. (2009). HLM 6.08 [computer software]. Chicago: Scientific Software International.

Regan, P. C. (1998). What if you can't get what you want? Willingness to compromise ideal mate selection standards as a function of sex, mate value, and relationship context. *Personality and Social Psychology Bulletin, 24,* 1294 – 1303.

Reichert, K. E., Heistermann, M., Hodges, J.K., Boesch, C., & Hohmann, G. (2002). What females tell males about their reproductive status: Are morphological and behavioral cues reliable signals of ovulation in bonobos (*Pan paniscus*)? *Ethology, 108,* 583 – 600.

Regan, P.C. (1996). Rhythms of desire: The association between ovulatory cycle phases and female sexual desire. *The Canadian Journal of Human Sexuality, 5,* 145 – 156.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57,* 199 – 226.

Rhodes, G., Sumich, A., & Byatt, G. (1999). Are average facial configurations attractive only\ because of their symmetry? *Psychological Science*, *10*, 52-58.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57,* 199 – 226.

Riggs, D. S. & O'Leary, K. D. (1996). Aggression Between Heterosexual Dating Partners An Examination of a Causal Model of Courtship Aggression. *Journal of Interpersonal Violence*, *11*, 519 – 540.

Rikowski, A. & Grammer, K. (1999). Human body odour, symmetry, and attractiveness. *Proceedings of the Royal Society of London, B, 266*, 869-874.

Rissman, E. F. & Crews, D. (1988). Hormonal correlates of sexual behavior in the female musk shrew: The role of estradiol. *Physiology & Behavior*, *44*, 1 – 7.

Roberts, M. L., Buchanan, K. L., & Evans, M. R. (2004). Testing the immunocompetence handicap hypothesis: A review of the evidence. *Animal Behavior*, *68*, 227 – 239.

**Röder, S., Brewer, G., & Fink, B. (2009). Menstrual cycle shifts in women's self-perception and motivation: A daily report method. *Personality and Individual Differences, 47,* 616 – 619.

Roberts, S. C., Havlíček, J., Flegr, J., Hruskova, M., Little, A. C., Jones, B.C., Perrett, D. I., & Petrie, M. (2004). Female facial attractiveness increases during the fertile phase of the menstrual cycle. *Proceedings of the Royal Society*, *B*, 271, 270 – 272.

Roney, J. R., Hanson, K. N., Durante, K. M., & Maestripieri, D. (2006). Reading men's faces: Women's mate attractiveness judgments track men's testosterone and interest in infants. *Proceedings of the Royal Society, B, 273*, 2169 – 2175.

*Roney, J. R. & Simmons, Z. L. (2008). Women's estradiol predicts preference for facial cues of mens' testosterone. *Hormones and Behavior, 53*, 14-19.

*Roney, J. R., Simmons, Z. L., & Gray, P. B. (2011). Changes in estradiol predict within-women shifts in attraction to facial cues of men's testosterone. *Psychoneuroendocrinology, 36*, 742-749.

Rowe, L., & Houle, D. (1996). The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *263*, 1415-1421.

Rupp, H.A., James, T.W., Ketterson, E.D., Sengelaub, D.R., Janssen, E., & Heiman, J.R. (2009). Neural activation in women in response to masculinized male faces: Mediation by hormones and psychosexual factors. *Evolution and Human Behavior, 30,* 1 – 10.

*Rupp, H.A., Librach, G. R., Feipel, N. C., Ketterson, E. D., Sengelaub, D. R., & Heiman, J. R. (2009). Partner status influences women's interest in the opposite sex. *Human Nature, 20,* 93 – 104.

**Saad, G., & Stenstrom, E. (2012). Calories, beauty, and ovulation: The effects of the menstrual cycle on food and appearance-related consumption. *Journal of Consumer Psychology, 22,* 102 – 113.

Said, C.P. & Todorov, A. (2011). A statistical model of facial attractiveness. *Psychological Science, 22,* 1183 – 90.

Samson, N., Fink, B., & Matts, P. (2011). Does a woman's skin color indicate her fertility levels? *Swiss Journal of Psychology, 70,* 199 – 202.

Scelza, B. A. (2011). Female choice and extra-pair paternity in a traditional human population. *Biology Letters*.

**Schwarz, S. & Hassebrauck, M. (2008). Self-perceived and observed variations in women's attractiveness throughout the menstrual cycle—a diary study. *Evolution and Human Behavior*, *29*, 282 – 288.

Scutt, D. & Manning, J. T. (1996). Ovary and ovulation: Symmetry and ovulation in women. *Human Reproduction*, *11*, 2477 – 2480.

Sear R. & Mace, R. (2008). Who keeps children alive? A review of the effects of kin on child survival. *Evolution and Human Behavior*, *29*, 1 – 18.

Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., & Gurven, M. (2008). Human

Adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society of London, B, 276,* 575 – 584.

Sellen, D.W. (2007). Evolution of infant and young child feeding: Implications for contemporary public health. *Annual Review of Nutrition*, *27*, 123 – 148.

Sheldon, M.S., Cooper, M.L., Geary, D.C., Hoard, M., & DeSoto, M.C. (2006). Fertility cycle patterns in motives for sexual behavior. *Personality and Social Psychology Bulletin*, *32*, 1659 – 1673.

Sillen-Tullberg, B. & Møller, A. P. (1993). The relationship between concealed ovulation and mating systems in anthropoid primates: a phylogenetic analysis. *The American Naturalist*, *141*, 1 – 25.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed Flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22,* 1359 – 1366.

Singh, D. (1993a). Adaptive significance of female physical attractiveness: Role of waist-to-hip ratio. *Journal of Personality and Social Psychology, 65,* 293 – 307.

*Singh, D. & Bailey, D. (2006, June). Do women's judgments of male attractiveness based on testosterone-mediated body parts depend on menstrual cycle phase? Poster presented at the annual meeting of the Human Behavior and Evolution Society (HBES), Philadelphia, PA.

**Singh, D. & Bronstad, P. M. (2001). Female body odour is a potential cue to ovulation. *Proceedings of the Royal Society of London B, 268,* 797—801.

Singh, D. & Young, R. K. (1995). Body weight, waist-to-hip ratio, breasts, and hips: Role in judgments of female attractiveness and desirability for relationships. *Ethology and*

Sociobiology, 16, 483 – 507.

Small, M.F. (1989). Female choice in nonhuman primates. *American Journal of Anthropology, 32,* 103 – 127.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation modeling. In S. Leinhart (Ed.), *Sociological methodology 1982* (pp. 290 – 312). San Francisco: Jossey-Bass.

Soler, C., Núñez, M. Gutiérrez, R., Núñez, J., Medina, P. Sancho, M., Núñez, A. (2003). Facial attractiveness in men provides clues to semen quality. *Evolution and Human Behavior, 24*, 199 – 207.

Steklis, H. D. & Whiteman, C. H. (1989). Loss of estrus in human evolution: Too many answers, too few questions. *Ethology and Sociobiology*, *10*, 417 – 434.

Strassman, B. (1997). The biology of menstruation in Homo sapiens: Total lifetime menses, Fecundity, and nonsynchrony in a natural-fertility population. *Current Anthropology, 38,* 123 – 129.

Stumpf, R. M., & Boesch, C. (2005). Does promiscuous mating preclude female choice? Female sexual strategies in chimpanzees (Pan troglodytes verus) of the Taï National Park, Côte d'Ivoire. *Behavioral ecology and sociobiology*, *57*, 511-524.

Szalay, F. S. & Costello, R. K. (1990). Evolution of permanent estrus displays in hominids. *Journal of Human Evolution*, *20*, 439 – 464.

*Teatero, M. (2009). *Mating strategies across the menstrual cycle: Preferences, jealousy, and masculinity* (Unpublished thesis). Lakehead University, Ontario.

Testart, J. & Frydman, R. (1982). Minimum time lapse between luteinizing hormone surge or human chorionic gonadotrophin administration and follicular rupture. *Fertility and Sterility, 37,* 50—53.

Thompson, A. P. (1983). Extramarital sex: A review of the research literature. *The Journal of Sex Research, 19,* 1 – 22.

*Thornhill, R., Chapman, J.F., & Gangestad, S.W. (2013). Women's preferences for men's scents associated with testosterone and cortisol levels: Patterns across the ovulatory cycle. *Evolution and Human Behavior, 34,* 216 – 221.

Thornhill, R. & Gangestad, S. W. (1994). Human fluctuating asymmetry and sexual behavior. *Psychological Science*, *5*, 297 – 302.

Thornhill, R. & Gangestad, S.W. (1999a). Facial attractiveness. *Trends in Cognitive Sciences, 3,* 452 – 460.

*,**Thornhill, R. & Gangestad, S. W. (1999b). The scent of symmetry: A human sex pheromone that signals fitness? *Evolution and Human Behavior, 20,* 175—201.

Thornhill, R. & Gangestad, S.W. (2008). *The Evolutionary Biology of Human Female Sexuality.* New York: Oxford University Press.

*Thornhill, R., Gangestad, S. W., Miller, R., Scheyd, G., McCollough, J. K., & Franklin, M. (2003). Major histocompatibility complex genes, symmetry, and body scent attractiveness in men and women. *Behavioral Ecology, 14,* 668—678.

Thornhill, R. & Moller, A. P. (1997). Developmental stability, disease, and medicine. *Biological Reviews, 72*, 497-548.

Todd, P. M., Penke, L., Fasolo, B., & Lenton, A. P. (2007). Different cognitive processes underlie human mate choices and mate preferences. *Proceedings of the National Academy of Sciences*, *104*, 15011-15016.

Trouton, G. T., Guitar, A. E., Carmen, R. A., Geher, G., & Grandis, T. L. (2012). *Journal of Social, Evolutionary, and Cultural Psychology, 6,* 469 – 479.

Van Dongen, S. (2006). Fluctuating asymmetry and developmental instability in evolutionary biology: Past, present, and future. *Journal of Evolutionary Biology, 19*, 1727-1743.

Van Dongen, S. & Gangestad, S. W. (2011). Human fluctuating asymmetry in relation to health and quality: A meta-analysis. *Evolution and Human Behavior, 32*, 380-398.

*Vaughn, J. E., Bradley, K. I., Byrd-Craven, J., & Kennison, S. M. (2010). The effect of mortality salience on women's judgments of male faces. *Evolutionary Psychology, 8*, 477-491.

von Rueden, C., Gurven, M., & Kaplan, H. (2011). Why do men seek high social status? Fitness payoffs to dominance and prestige. *Proceedings of the Royal Society B, 278,* 2223 – 2232.

Waynforth, D. (2001). Mate choice trade-offs and women's preferences for physically attractive men. *Human Nature*, *12*, 207 – 219.

Wedekind, C., Seebeck, T., Bettens, F., & Paepke, A. J. (2005). MHC-dependent mate preferences in humans. *Proceedings of the Royal Society, B, 260*, 245 – 249.

Wegienka, G. & Day Baird, D., (2005). A comparison of recalled date of last menstrual period with prospectively recorded dates. *Journal of Women's Health*, *14*, 248 – 252.

Weinberg, C.R. & Wilcox, A.J. (1995). A model for estimating the potency and survival of human gametes *in Vivo*. *Biometrics, 51,* 405 – 412.

*Welling, L. L. M., Jones, B. C., DeBruine, L. M., Conway, C. A., Law Smith, M. J.,Little, A. C., & Al-Dujaili, E. A. S. (2007). Raised salivary testosterone in women is associated with increased attraction to masculine faces. *Hormones and Behavior, 52,* 156-161.

Wilcox, A.J., Dunson, D., & Day Baird, D. (2000). The timing of the "fertile window" in the menstrual cycle: Day specific estimates from a prospective study. *British Medical Journal*, *321*, 1259 – 1262.

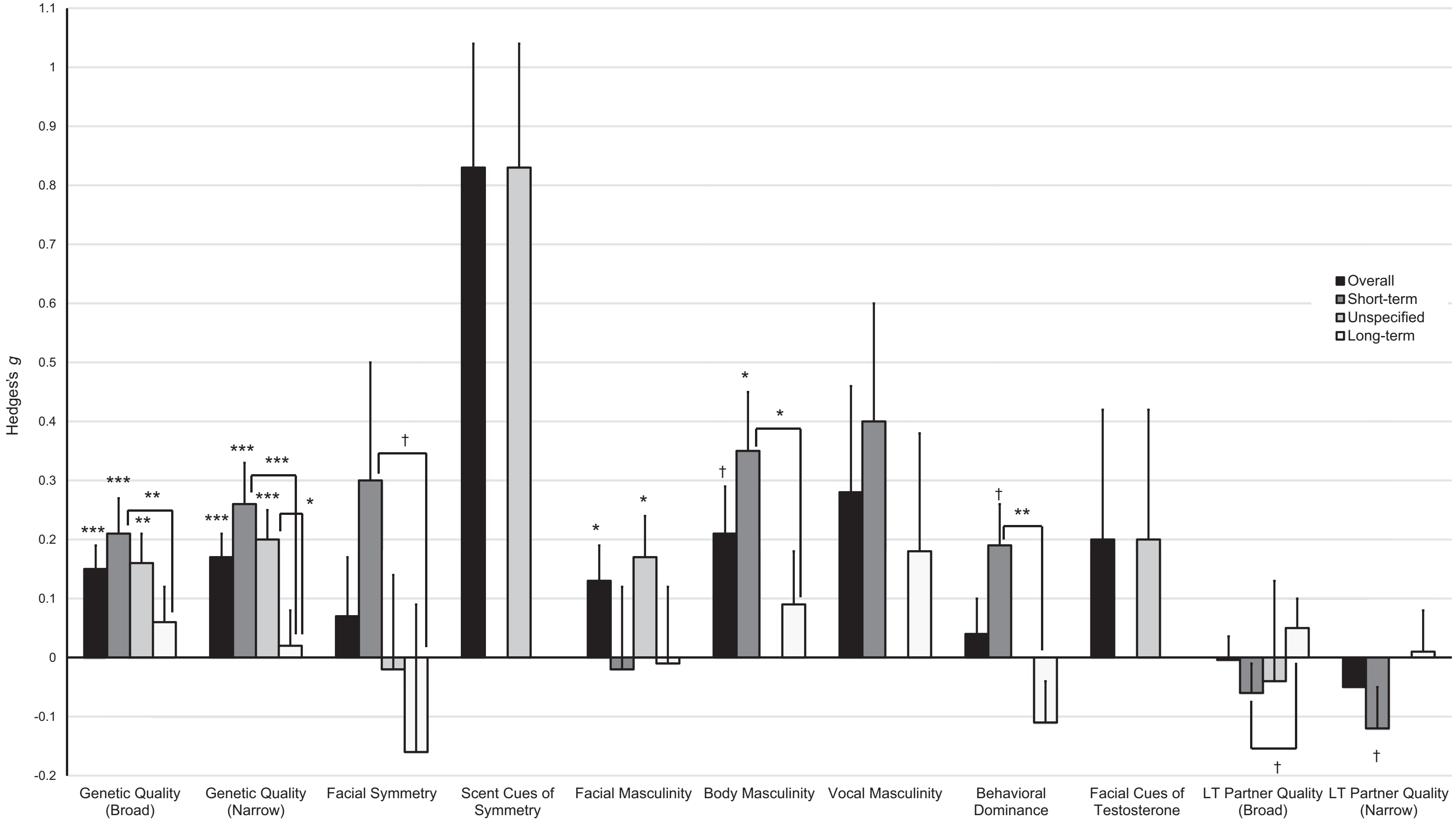
Wilcox, A.J., Dunson, D.B., Weinberg, C.R., Trussell, J., & Day Baird, D. (2001). Likelihood of conception with a single act of intercourse: Providing benchmark rates for assessment of post-coital contraceptives. *Contraception, 63,* 211 – 215.

Wilcox, A. J., Day Baird, D., Dunson, D. B., McConnaughey, D. R., Kesner, J. S., & Weinberg, C. R. (2004). On the frequency of intercourse around ovulation: Evidence for biological influences. *Human Reproduction*, *19*, 1539 – 1543.

Wilcox, A. J., Weinberg, C. R., & Day Baird, D. (1995). Timing of sexual intercourse in relation to ovulation: Effects on the probability of conception, survival of pregnancy, and sex of the baby. *The New England Journal of Medicine, 333*, 1517-1521.

Wilson, E.K., Rogler, J.C., & Erb, R.E. (1979). Effect of sexual experience, location, malnutrition, and repeated sampling on concentrations of testosterone in blood plasma of *Gallus domesticus* roosters. *Poultry Science, 58,* 178 – 186.

Wood, D. & Brumbaugh, C.C. (2009). Using revealed mate preferences to evaluate market force and differential preference explanations for mate selection. *Journal of Personality and Social Psychology, 96,* 1226 – 1244.



† $p \leq .10$. * $p \leq .05$. ** $p \leq .01$. *** $p \leq .001$.

Figure 1. Summary of results from all analyses examining cycle shifts in women's preferences for hypothesized cues of genetic and long-term partner quality in ancestral males. For each sample, the weighted mean Hedges's *g* is presented overall (across relationship contexts) and separately for short-term, unspecified, and long-term relationship contexts. Errors bars represent standard error. LT = long-term.

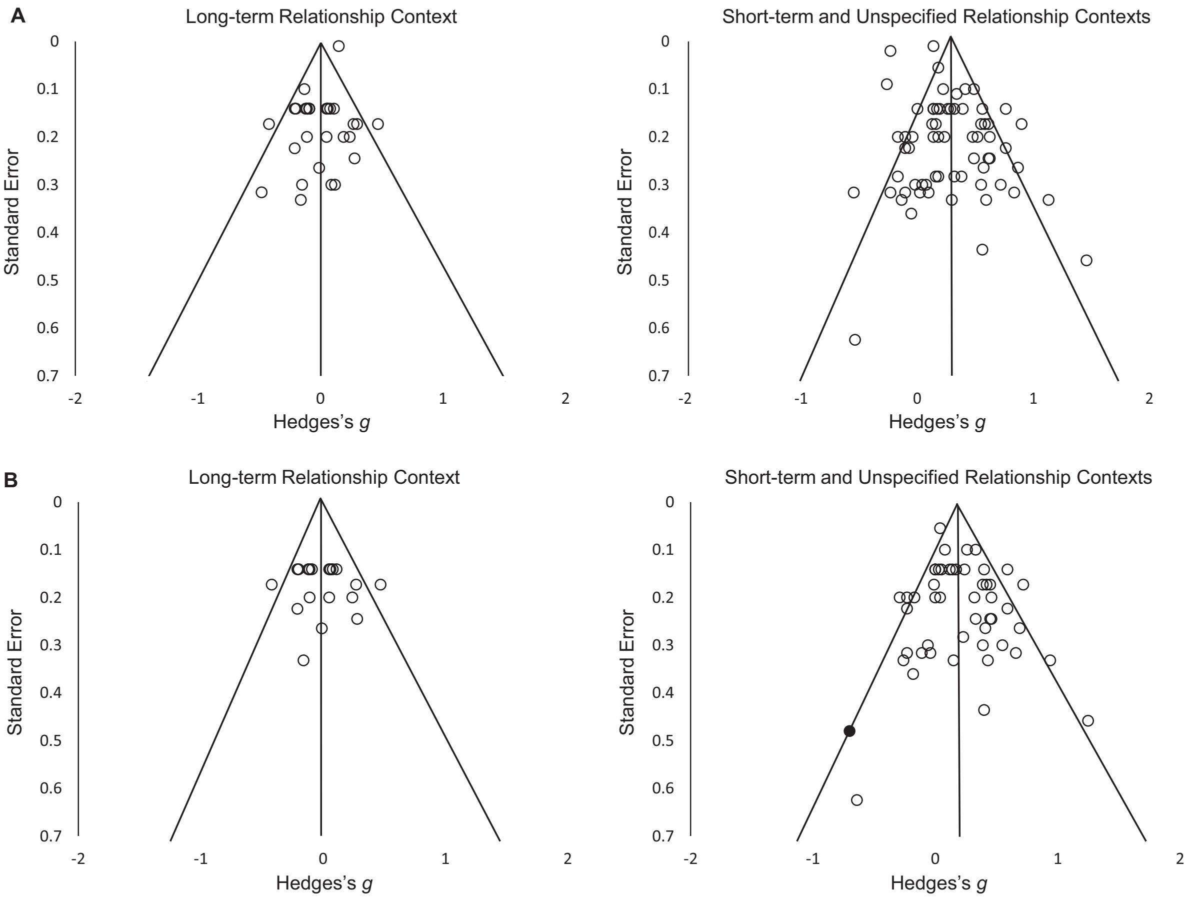


Figure 2. Funnel plots to examine evidence for an underrepresentation of small effects among the sample of effects for which the ovulatory shift hypothesis predicts relationship context-dependent cycle shifts—namely, effects assessing cycle shifts in preferences for characteristics hypothesized to have reflected genetic quality in ancestral males. Effects assessing cycle shifts in a long-term relationship context (no cycle shift predicted) are plotted separately from effects assessing cycle shifts in a short-term or unspecified relationship context (positive cycle shift predicted). Empty circles represent observed effects. Filled circle represents imputed putative missing effect. (A) Sample of effects that included a broad set of mate preference measures. (B) Sample of effects that included a narrow set of mate preference measures.

Figure 3.

Study	Effect (Relationship Context)	Effect Size (*g*)
Miller (2003)*	Facial Masculinity (ST)	-0.65
Rupp, Librach, et al. (2009)	Facial Masculinity (U)	-0.64
Morrison et al. (2010)*	Male Facial Movements (ST)	-0.38
Miller (2003)*	Body Muscularity (ST)	-0.35
Rantala et al. (2010)*	Torso Hair (U)	-0.35
Caryl et al. (2009)*	Strong (U)	-0.29
Izbicki & Johnson (2010)	Facial Masculinity (ST)	-0.29
Hromatko et al. (2006)	Facial Symmetry (U)	-0.26
Oinonen & Mazmanian (2007)	Facial Symmetry (U)	-0.23
Roney, Simmons, & Gray (2011)	Facial Masculinity (U)	-0.23
Rantala et al. (2006)	Scent Cues of Testosterone (U)	-0.18
Moore et al. (2011), Study 2	Facial Cues of Testosterone (U)	-0.17
Miller (2003)*	Tall (ST)	-0.15
Izbicki & Johnson (2010)*	Strong (ST)	-0.13
Moore (2011)	Facial Symmetry (U)	-0.11
Izbicki & Johnson (2010)*	Dominant (ST)	-0.09
Miller (2003)*	Big Ego (ST)	-0.09
Fink (2012)	Facial Masculinity (U)	-0.06
Moore (2011)	Facial Masculinity (U)	-0.04
Vaughn et al. (2010)	Facial Masculinity (U)	-0.01
Izbicki & Johnson (2010)*	Facial Darkness (ST)	0.00
Beaulieu (2007), Study 2*	Dominance (U)	0.00
Bressan & Stranieri (2008)	Facial Masculinity (U)	0.00
Cárdenas & Harris (2007)	Facial symmetry (ST)	0.00
Singh & Bailey (2006)	Male SHR (ST)	0.00
Caryl et al. (2009)*	Aggressive (U)	0.02
Harris (2011)	Facial Masculinity (U)	0.03
Caryl et al. (2009)*	Arrogant (U)	0.04
Koehler et al. (2006)	Facial Symmetry (U)	0.04
Perrett et al. (2013), Study 1	Facial Masculinity (U)	0.04
Gangestad et al. (2007)	Soc. Respected/Influential (ST)	0.05

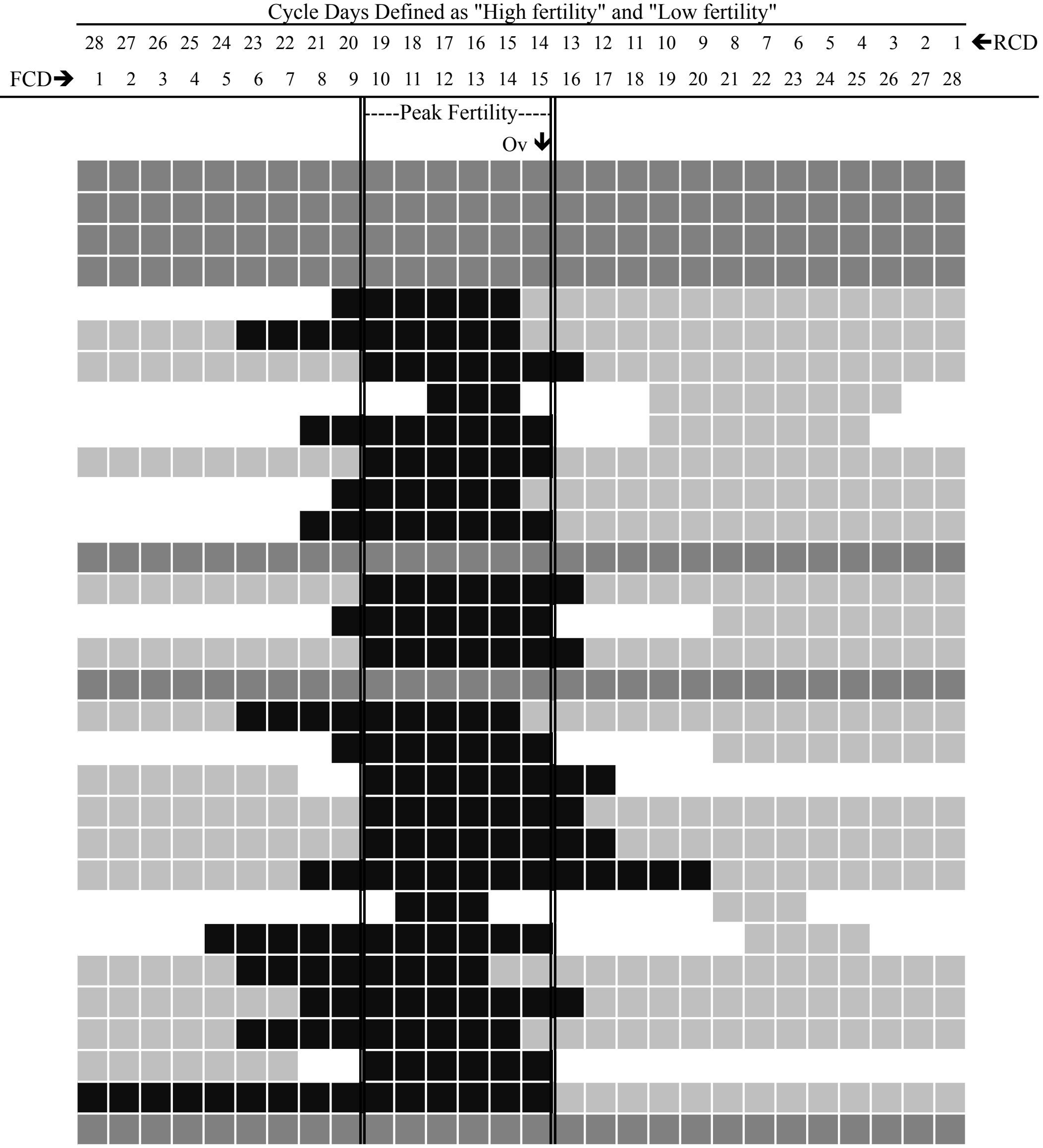


Figure 3. (*continued*)

Study	Effect (Relationship Context)	Effect Size (*g*)
Gangestad et al. (2011)	Facial Masculinity (U)	0.08
Luevano& Zebrowitz (2006)*	Dominant (ST)	0.09
Gangestad et al. (2004)	Competitiveness (ST)	0.12
Gangestad et al. (2007)	Arrogant/Self-centered (ST)	0.14
Garver-Apgar & Gangestad (2012)	Soc. Presence/Competitiveness (ST)	0.15
Koehler et al. (2006)*	Facial Averageness (U)	0.15
Caryl et al. (2009)*	Conceited (U)	0.17
Gangestad et al. (2007)	Muscular (ST)	0.17
Frost (1994)*	Facial Darkness (U)	0.19
Penton-Voak et al. (1999), Study 2	Facial Masculinity (ST)	0.23
Gangestad et al. (2007)	Confrontative (ST)	0.24
Welling et al. (2007)	Facial Masculinity (U)	0.26
Luevano & Zebrowitz (2006)	Facial Masculinity (ST)	0.30
Perrett et al. (2013), Study 2	Facial Masculinity (ST)	0.32
Havlíček et al. (2005)	Scent Cues of Dominance (U)	0.33
Jones, Little, et al. (2005), Study 2	Facial Masculinity (U)	0.33
Lukaszewski & Roney (2009)*	Dominant (ST)	0.36
Penton-Voak & Perrett (2000)	Facial Masculinity (U)	0.39
Singh & Bailey (2006)	Male-typical WHR (ST)	0.39
Gangestad et al. (2004)	Social Presence (ST)	0.40
Johnston et al. (2001)	Facial Masculinity (U)	0.40
Little, Jones, et al. (2007), Study 1	Facial Symmetry (U)	0.41
Puts (2005)	Vocal Masculinity (ST)	0.42
Roney, Simmons, & Gray (2011)	Facial Cues of Testosterone (U)	0.43
Feinberg (2012)	Vocal Masculinity (ST)	0.45
Penton-Voak et al. (1999), Study 1	Facial Masculinity (U)	0.45
Provost et al. (2008)*	Male-typical Walk (U)	0.45
Pawlowski & Jasienska (2005)	Taller Man relative to Self (ST)	0.46
Roney & Simmons (2008)	Facial Cues of Testosterone (U)	0.46

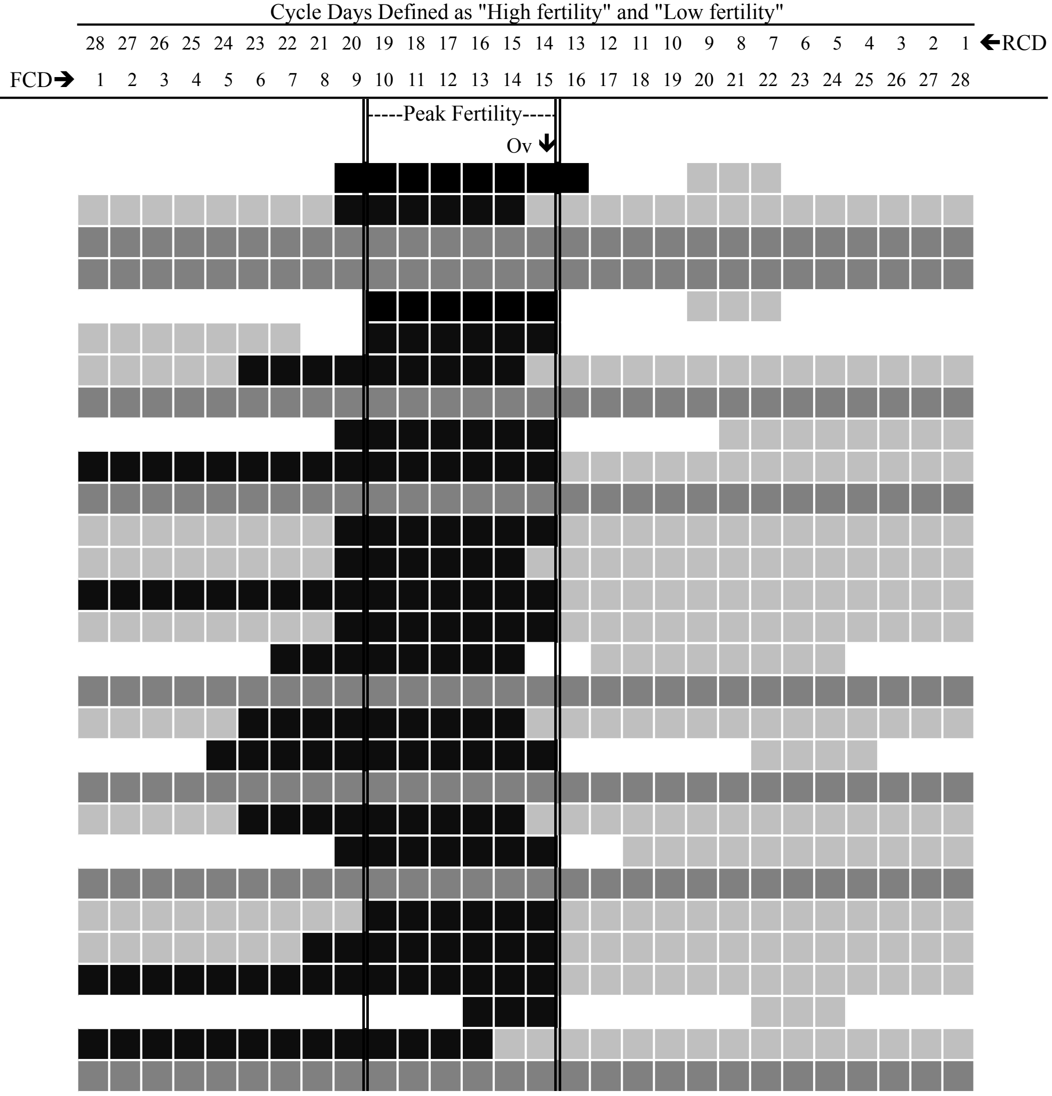
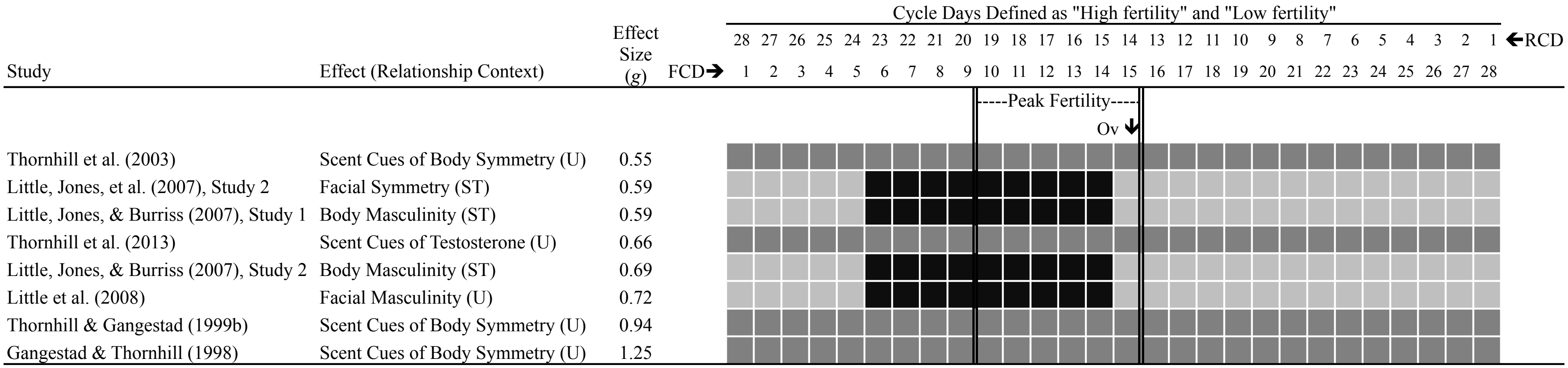


Figure 3. (*continued*)



Cycle Days Defined as "High fertility" and "Low fertility"

				28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	←RCD
Study	Effect (Relationship Context)	Effect Size (*g*)	FCD→	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	
Thornhill et al. (2003)	Scent Cues of Body Symmetry (U)	0.55		▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	
Little, Jones, et al. (2007), Study 2	Facial Symmetry (ST)	0.59		░	░	░	░	░	■	■	■	■	■	■	■	■	■	░	░	░	░	░	░	░	░	░	░	░	░	░	░	
Little, Jones, & Burriss (2007), Study 1	Body Masculinity (ST)	0.59		░	░	░	░	░	■	■	■	■	■	■	■	■	■	░	░	░	░	░	░	░	░	░	░	░	░	░	░	
Thornhill et al. (2013)	Scent Cues of Testosterone (U)	0.66		▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	
Little, Jones, & Burriss (2007), Study 2	Body Masculinity (ST)	0.69		░	░	░	░	░	■	■	■	■	■	■	■	■	■	░	░	░	░	░	░	░	░	░	░	░	░	░	░	
Little et al. (2008)	Facial Masculinity (U)	0.72		░	░	░	░	░	■	■	■	■	■	■	■	■	■	░	░	░	░	░	░	░	░	░	░	░	░	░	░	
Thornhill & Gangestad (1999b)	Scent Cues of Body Symmetry (U)	0.94		▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	
Gangestad & Thornhill (1998)	Scent Cues of Body Symmetry (U)	1.25		▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	

Figure 3. High- and low-fertility cycle phase definitions for effects assessing cycle shifts in preferences for hypothesized cues of genetic quality in a short-term (ST) or unspecified (U) relationship context (where a cycle shift was predicted. Effects marked with asterisks were included only in the broad sample. Effects not marked with asterisks were included in both the broad and narrow samples. Black boxes and light gray boxes indiate cycle days defined as high-fertility and low-fertility, respectively. White (unfilled) boxes indicate days that fell outside of high- and low-fertility windows and were therefore excluded from analysis. Dark gray boxes indicate that fertility was treated as a continuous variable, and therefore all cycle days were included in analyses. High- and low-fertility windows are displayed in terms of forward cycle days (FCD; days since last menstrual onset) and reverse cycle day (RCD; days until next menstrual onset) for studies that used the forward counting or reverse counting method, respectively. High-fertility windows are displayed in terms of days from ovulation, and low-fertility windows are displayed in terms of FCD, for studies that used luteinizing hormone tests to verify impending ovulation. To enable comparing high- and low-fertility windows across these three methods, we have assumed a 28-day cycle length, with ovulation (Ov) occurring on FCD 14/RCD 15. We have demarcated a suggested "peak fertility" window with double lines. This window includes the 6 days with the highest average conception probabilities for regularly cycling women as reported by Wilcox et al. (2001). SHR = shoulder-to-hip ratio; WHR = waist-to-hip ratio.

Figure 4.

Funnel Plot of Effects Assessing Cycle Shifts in Women's Other-rated Attractiveness

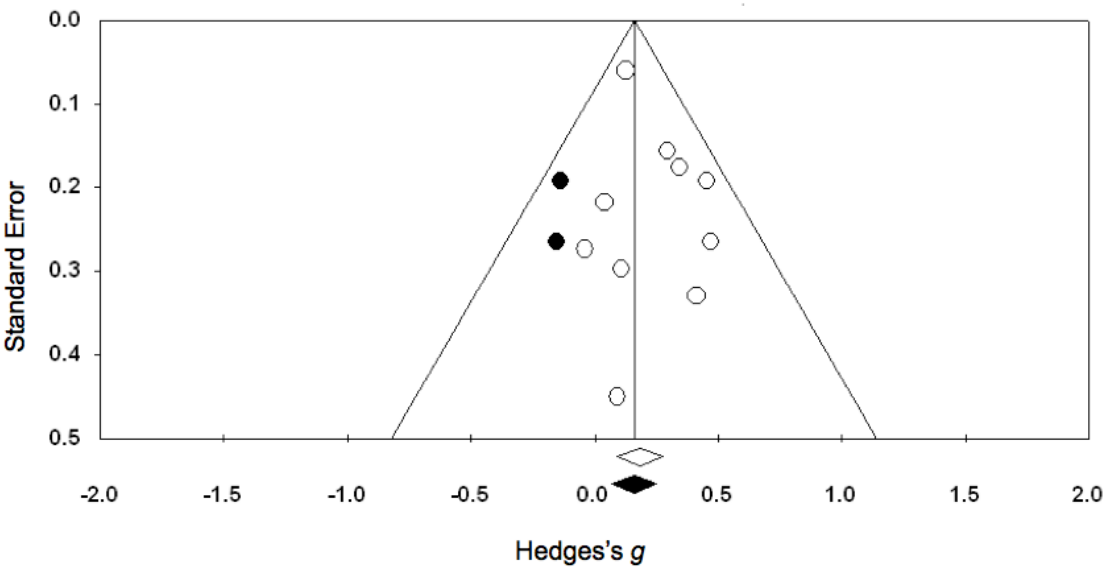


Figure 5.

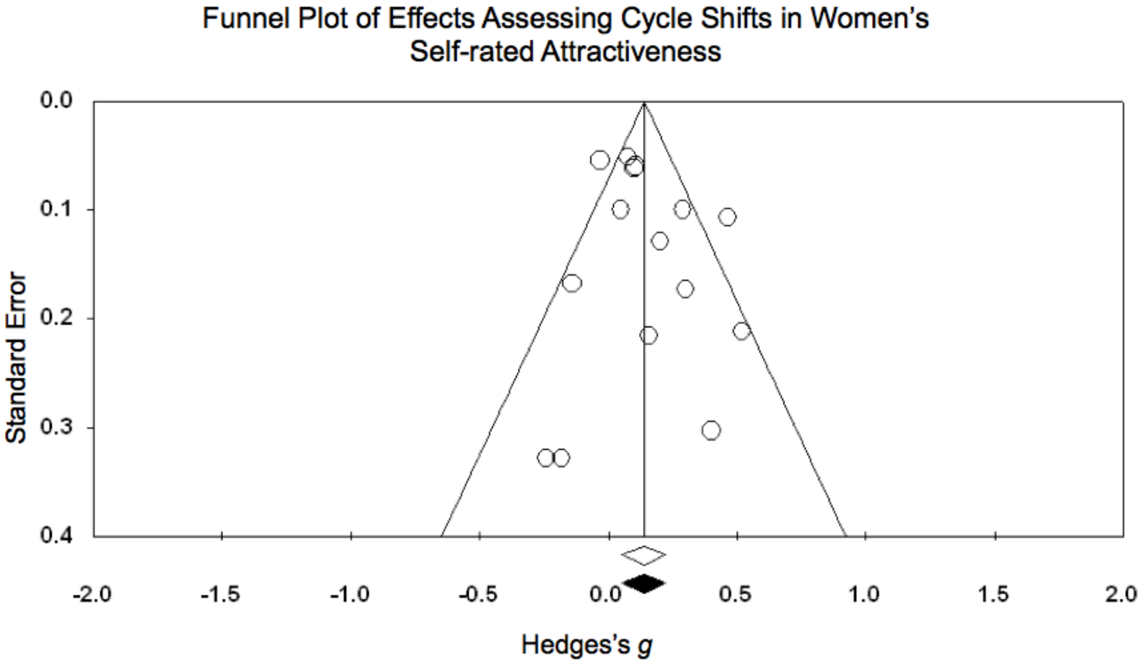

Funnel Plot of Effects Assessing Cycle Shifts in Women's Self-rated Attractiveness
Standard Error
0.0
0.1
0.2
0.3
0.4
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Hedges's *g*

Figure 6.

Funnel Plot of Effects Assessing Cycle Shifts in Women's Attractive Behavior and Attractiveness-enhancing Behavior and Motivation

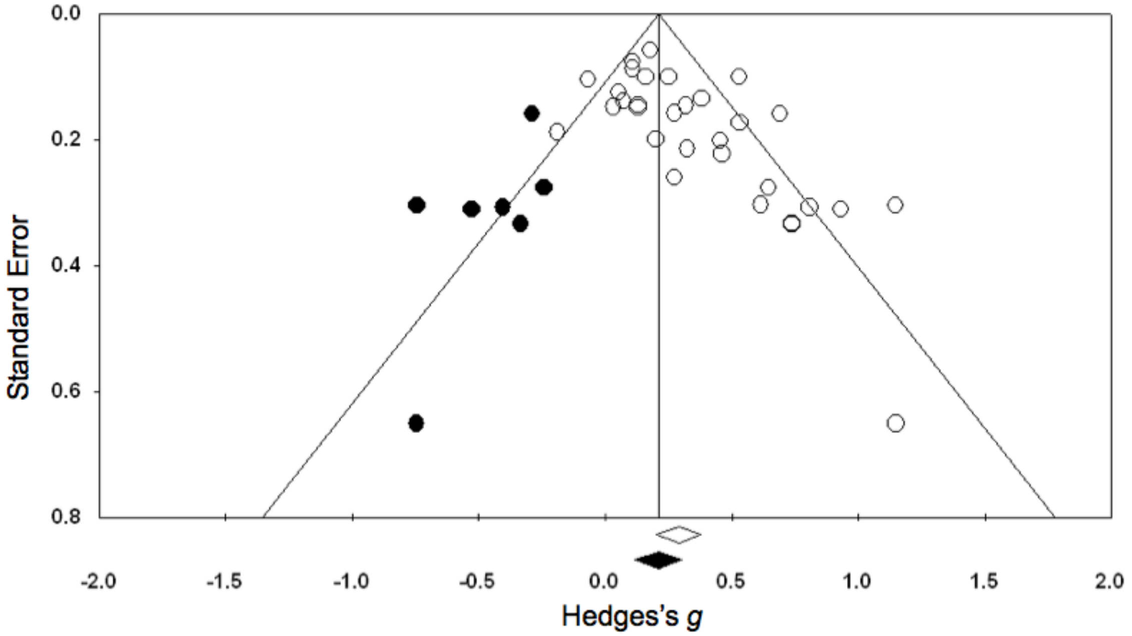


Table 1

Studies Assessing Ovulation-Related Cycle Shifts in Mate Preferences: Basic Characteristics, Effect Size, and Inclusion in Analyses

								Inclusion in analyses										
			Sample size					Long-term partner quality		Genetic quality								
Study and effects	Relationship context	*N*	High-fertility *n*	Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Beaulieu (2007), Study 2																		
Relationship skills (composite of kind, understanding, loyal, generous)	U	92	33	59	0.00	0.05		✓										
Dominance (composite of dominant, powerful, aggressive)	U	92	33	59	−0.20	0.05				✓								
Education (composite of educated, cultured, intelligent)	U	92	33	59	M	M												
Good financial prospects (composite of wealthy, good financial prospects)	U	92	33	59	M	M												
Beaulieu (2007), Study 4																		
Relationship skills (composite of kind, understanding, loyal, generous)	ST, LT	33	Within participants		M	M												
Dominance (composite of dominant, powerful, aggressive)	ST, LT	33	Within participants		M	M												
Education (composite of educated, cultured, intelligent)																		
Good financial prospects (composite of wealthy, good financial prospects)	ST, LT	33	Within participants		M	M												

(*table continues*)

Table 1 (*continued*)

								Inclusion in analyses										
		Sample size						Long-term partner quality		Genetic quality								
Study and effects	Relationship context	*N*	High-fertility *n*	Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Bressan & Stranieri (2008)																		
Facial masculinity	U	198	97	101	0.00	0.02				✓	✓			✓				
Bullock (2000), Chapter 4																		
Chin length	U	60	M	M	M	M												
Cárdenas & Harris (2007)																		
Facial symmetry	ST	53	Within participants		0.00	0.02				✓	✓	✓						
Caryl et al. (2009)																		
Pupil size	U	50	22	28	M	M												
Arrogant	U	50	22	28	0.04	0.08				✓								
Ingenious	U	50	22	28	−0.22	0.08	4											
Aggressive	U	50	22	28	0.02	0.08				✓								
Strong	U	50	22	28	−0.29	0.08				✓								
Conceited	U	50	22	28	0.17	0.08				✓								
Enterprising	U	50	22	28	−0.14	0.08	4											
Inventive	U	50	22	28	0.04	0.08	4											
Warm	U	50	22	28	−0.11	0.08		✓										
Sensitive	U	50	22	28	M	M												
Sentimental	U	50	22	28	M	M												
Sympathetic	U	50	22	28	M	M												
Jolly	U	50	22	28	M	M												
Helpful	U	50	22	28	M	M												
Appreciative	U	50	22	28	M	M												
Considerate	U	50	22	28	M	M												
Cooperative	U	50	22	28	M	M												
Friendly	U	50	22	28	M	M												
Talkative	U	50	22	28	M	M												
Forgiving	U	50	22	28	M	M												
Emotional	U	50	22	28	M	M												
Foresighted	U	50	22	28	M	M												
Shrewd	U	50	22	28	M	M												
Industrious	U	50	22	28	M	M												
Assertive	U	50	22	28	M	M												
Forceful	U	50	22	28	M	M												
Timid	U	50	22	28	M	M												
Dependent	U	50	22	28	M	M												
Fickle	U	50	22	28	M	M												
Frivolous	U	50	22	28	M	M												
Opportunistic	U	50	22	28	M	M												
Hardheaded	U	50	22	28	M	M												
Confident	U	50	22	28	M	M												
DeBruine et al. (2005)																		
Facial self-resemblance	U	43	21	22	−0.09	0.09	4											

(*table continues*)

Table 1 (*continued*)

								Inclusion in analyses										
			Sample size					Long-term partner quality		Genetic quality								
Study and effects	Relationship context	*N*	High-fertility *n*	Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Feinberg et al. (2006)																		
Vocal masculinity	U	26	Within participants		M	M												
Feinberg (2012)																		
Vocal masculinity	ST	22	Within participants		0.45	0.06				✓	✓					✓		
Vocal masculinity	LT	22	Within participants		−0.21	0.05				✓	✓					✓		
Fink (2012)																		
Facial masculinity	U	20	Within participants		−0.06	0.09				✓	✓			✓				
Flowe et al. (2012)																		
Behavioral masculinity	U	106	45	61	M	M												
Frost (1994)																		
Darker skin tone	U	36	15	21	0.19	0.11				✓								
Gangestad et al. (2004)																		
Social presence	ST	237	Fertility continuous		0.40	0.02				✓	✓						✓	
Social presence	LT	237	Fertility continuous		0.08	0.02				✓	✓						✓	
Direct intrasexual competitiveness	ST	237	Fertility continuous		0.12	0.02				✓	✓						✓	
Direct intrasexual competitiveness	LT	237	Fertility continuous		−0.11	0.02				✓	✓						✓	
Gangestad et al. (2007)																		
Muscular	ST	237	Fertility continuous		0.17	0.02				✓	✓				✓			
Muscular	LT	237	Fertility continuous		−0.09	0.02				✓	✓				✓			
Confrontative (with other men)	ST	237	Fertility continuous		0.24	0.02				✓	✓						✓	
Confrontative (with other men)	LT	237	Fertility continuous		−0.12	0.02				✓	✓						✓	
Socially respected and influential	ST	237	Fertility continuous		0.05	0.02				✓	✓						✓	
Socially respected and influential	LT	237	Fertility continuous		−0.21	0.02				✓	✓						✓	
Arrogant and self-centered	ST	237	Fertility continuous		0.14	0.02				✓	✓						✓	
Arrogant and self-centered	LT	237	Fertility continuous		−0.20	0.02				✓	✓						✓	
Intelligent	ST	243	Fertility continuous		−0.22	0.02	4											
Intelligent	LT	243	Fertility continuous		−0.12	0.02	4											
Faithful	ST	243	Fertility continuous		−0.22	0.02		✓	✓									

(*table continues*)

Table 1 (*continued*)

								Inclusion in analyses										
		Sample size						Long-term partner quality		Genetic quality								
Study and effects	Relationship context	*N*	High-fertility *n*	Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Faithful	LT	243	Fertility continuous		0.04	0.02		✓	✓									
Warm (kind and understanding)	ST	243	Fertility continuous		−0.16	0.02		✓	✓									
Warm (kind and understanding)	LT	243	Fertility continuous		0.05	0.02		✓	✓									
Likely to be financially successful	ST	243	Fertility continuous		−0.15	0.02		✓	✓									
Likely to be financially successful	LT	243	Fertility continuous		−0.10	0.02		✓	✓									
Likely to be a good parent	ST	243	Fertility continuous		0.04	0.02		✓	✓									
Likely to be a good parent	LT	243	Fertility continuous		0.06	0.02		✓	✓									
Gangestad et al. (2011)																		
Facial masculinity	U	59	Within participants		0.08	0.01				✓	✓			✓				
Gangestad & Thornhill (1998)																		
Scent cues of body symmetry	U	28	Fertility continuous		1.25	0.21				✓	✓		✓					
Garver-Apgar & Gangestad (2012)																		
Average of social presence and direct intrasexual competitiveness	ST	18	Within participants		0.15	0.11				✓	✓						✓	
Average of social presence and direct intrasexual competitiveness	LT	18	Within participants		−0.16	0.11				✓	✓						✓	
Harris (2011)																		
Facial masculinity	U	258	80	178	0.03	0.02				✓	✓			✓				
Haselton & Miller (2006)																		
Wealth versus creativity	ST, LT	41	Fertility continuous		M	M	5											
Havlíček et al. (2005)																		
Scent cues of dominance (Narcissism scale from CPI)	U	65	30	35	0.33	0.06				✓	✓							
Hromatko et al. (2006)																		
Facial symmetry	U	64	11	53	−0.26	0.11				✓	✓	✓						

(*table continues*)

Table 1 (*continued*)

								Inclusion in analyses										
			Sample size					Long-term partner quality		Genetic quality								
Study and effects	Relationship context	N	High-fertility n	Low-fertility n	Effect Size (g)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Izbicki & Johnson (2010)																		
Facial masculinity	ST	42	Within participants		−0.29	0.04				✓	✓			✓				
Facial masculinity	LT	42	Within participants		−0.42	0.03				✓	✓			✓				
Strong	ST	42	Within participants		−0.13	0.02				✓								
Strong	LT	42	Within participants		−0.13	0.01				✓								
Warm	ST	42	Within participants		0.05	0.02		✓										
Warm	LT	42	Within participants		0.10	0.02		✓										
Mature	ST	42	Within participants		−0.21	0.03	4											
Mature	LT	42	Within participants		−0.13	0.02	4											
Socially competent	ST	42	Within participants		−0.11	0.03	4											
Socially competent	LT	42	Within participants		0.02	0.03	4											
Nurturant	ST	42	Within participants		0.02	0.03		✓										
Nurturant	LT	42	Within participants		0.26	0.03		✓										
Threatening	ST	42	Within participants		0.00	0.02	4											
Threatening	LT	42	Within participants		0.07	0.02	4											
Dominant	ST	42	Within participants		0.02	0.03				✓								
Dominant	LT	42	Within participants		−0.09	0.02				✓								
Dark	ST	42	Within participants		0.00	0.01				✓								
Dark	LT	42	Within participants		0.15	0.01				✓								
Johnston et al. (2001)																		
Facial masculinity	U	29	Within participants		0.40	0.19				✓	✓			✓				
Jones, Little, et al. (2005), Study 2																		
Facial masculinity	U	328	169	159	0.33	0.01				✓	✓			✓				
Koehler et al. (2002)																		
Facial symmetry	ST, LT	29	Within participants		M	M												

(*table continues*)

Table 1 (*continued*)

								Inclusion in analyses										
		Sample size						Long-term partner quality		Genetic quality								
Study and effects	Relationship context	*N*	High-fertility *n*	Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Koehler et al. (2006)																		
Facial averageness	U	50	Within participants		0.15	0.04				✓								
Facial symmetry	U	50	Within participants		0.04	0.04				✓	✓	✓						
Li et al. (2006)																		
Multiple traits	ST, LT	54	Within participants		M	M	5											
Little, Jones, et al. (2007), Study 1																		
Facial symmetry	U	31	Within participants		0.41	0.07				✓	✓	✓						
Little, Jones, et al. (2007), Study 2																		
Facial symmetry	ST	210	63	147	0.59	0.02				✓	✓	✓						
Facial symmetry	LT	210	63	147	0.05	0.02				✓	✓	✓						
Little, Jones, & Burriss (2007), Study 1																		
Body masculinity	ST	97	36	61	0.59	0.05				✓	✓				✓			
Body masculinity	LT	97	36	61	0.05	0.04				✓	✓				✓			
Little, Jones, & Burriss (2007), Study 2																		
Body masculinity	ST	17	Within participants		0.69	0.07				✓	✓				✓			
Body masculinity	LT	17	Within participants		0.28	0.06				✓	✓				✓			
Little et al. (2008)																		
Facial masculinity	U	150	54	96	0.72	0.03				✓	✓			✓				
Luevano & Zebrowitz (2006)																		
Dominant	ST	25	Within participants		0.09	0.04				✓								
Dominant	LT	25	Within participants		0.30	0.03				✓								
Facial masculinity	ST	25	Within participants		−0.23	0.04				✓	✓			✓				
Facial masculinity	LT	25	Within participants		0.06	0.02				✓	✓			✓				
Warm	ST	25	Within participants		−0.25	0.03		✓										
Warm	LT	25	Within participants		0.00	0.03		✓										
Lukaszewski & Roney (2009)																		
Dominant	ST	111	Fertility continuous		0.36	0.04				✓								
Dominant	LT	111	Fertility continuous		0.19	0.04				✓								

(*table continues*)

Table 1 (*continued*)

								Inclusion in analyses										
			Sample size					Long-term partner quality		Genetic quality								
Study and effects	Relationship context	*N*	High-fertility *n*	Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Kind	ST	111	Fertility continuous		−0.02	0.04		✓										
Kind	LT	111	Fertility continuous		−0.05	0.04		✓										
Trustworthy	ST	111	Fertility continuous		−0.02	0.04		✓										
Trustworthy	LT	111	Fertility continuous		−0.05	0.04		✓										
McClellan et al. (2007)																		
Body masculinity	ST, LT, U	24	10	14	M	M												
Age	ST, LT, U	24	10	14	M	M												
McDonald & Navarrete (2012), Sample 1																		
Body muscularity	U	80	42	38	M	M												
Same-race (vs. other-race) face	U	80	42	38	M	M												
McDonald & Navarrete (2012), Sample 2																		
Body muscularity	U	81	43	38	M	M												
Same-race (vs. other-race) face	U	81	43	38	M	M												
Miller (2003)																		
Intelligent	ST	45	Fertility continuous		0.11	0.09	4											
Intelligent	LT	45	Fertility continuous		0.01	0.09	4											
Future kids' intelligence	ST	45	Fertility continuous		0.08	0.09	4											
Future kids' intelligence	LT	45	Fertility continuous		0.38	0.10	4											
Mathematical problem-solving ability	ST	45	Fertility continuous		0.28	0.09	4											
Mathematical problem-solving ability	LT	45	Fertility continuous		0.28	0.09	4											
Good grades	ST	45	Fertility continuous		0.50	0.10	4											
Good grades	LT	45	Fertility continuous		0.31	0.09	4											
Creative/imaginative	ST	45	Fertility continuous		−0.36	0.10	4											
Creative/imaginative	LT	45	Fertility continuous		0.05	0.09	4											
Future kids' sense of humor	ST	45	Fertility continuous		−0.22	0.09	4											

(*table continues*)

Table 1 (*continued*)

			Sample size					Inclusion in analyses										
								Long-term partner quality		Genetic quality								
Study and effects	Relationship context	*N*	High-fertility *n*	Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Future kids' sense of humor	LT	45	Fertility continuous		0.34	0.10	4											
Social sensitivity	ST	45	Fertility continuous		−0.32	0.09	4											
Social sensitivity	LT	45	Fertility continuous		0.03	0.09	4											
Adaptable to situations and challenges	ST	45	Fertility continuous		−0.35	0.10	4											
Adaptable to situations and challenges	LT	45	Fertility continuous		0.15	0.09	4											
Big ego	ST	45	Fertility continuous		−0.09	0.09				✓								
Big ego	LT	45	Fertility continuous		0.12	0.09				✓								
Body muscularity	ST	45	Fertility continuous		−0.35	0.10				✓								
Body muscularity	LT	45	Fertility continuous		0.09	0.09				✓								
Facial masculinity	ST	45	Fertility continuous		−0.65	0.10				✓								
Facial masculinity	LT	45	Fertility continuous		−0.48	0.10				✓								
Tall	ST	45	Fertility continuous		−0.15	0.09				✓								
Tall	LT	45	Fertility continuous		−0.15	0.09				✓								
Happy	ST	45	Fertility continuous		−0.46	0.10	4											
Happy	LT	45	Fertility continuous		−0.09	0.09	4											
Exciting/spontaneous	ST	45	Fertility continuous		−0.16	0.09	4											
Exciting/spontaneous	LT	45	Fertility continuous		0.24	0.09	4											
Talkative/extraverted	ST	45	Fertility continuous		−0.13	0.09	4											
Talkative/extraverted	LT	45	Fertility continuous		0.26	0.09	4											
Likelihood of being unfaithful (reverse-coded)	ST	45	Fertility continuous		0.01	0.09		✓										
Likelihood of being unfaithful (reverse-coded)	LT	45	Fertility continuous		−0.18	0.09		✓										
Future money making	ST	45	Fertility continuous		−0.06	0.09		✓										

(*table continues*)

Table 1 (*continued*)

								Inclusion in analyses										
		Sample size						Long-term partner quality		Genetic quality								
Study and effects	Relationship context	*N*	High-fertility *n*	Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Future money making	LT	45	Fertility continuous		0.12	0.09		✓										
Good at playing with and caring for kids	ST	45	Fertility continuous		−0.37	0.10		✓										
Good at playing with and caring for kids	LT	45	Fertility continuous		−0.13	0.09		✓										
Future career success	ST	45	Fertility continuous		0.05	0.09		✓										
Future career success	LT	45	Fertility continuous		0.41	0.10		✓										
Sympathetic/kind	ST	45	Fertility continuous		−0.21	0.09		✓										
Sympathetic/kind	LT	45	Fertility continuous		0.12	0.09		✓										
Constructive in arguments	ST	45	Fertility continuous		0.14	0.09	4											
Constructive in arguments	LT	45	Fertility continuous		0.44	0.10	4											
Neat/organized	ST	45	Fertility continuous		0.19	0.09	4											
Neat/organized	LT	45	Fertility continuous		0.27	0.09	4											
Moody/irritable	ST	45	Fertility continuous		−0.05	0.09	4											
Moody/irritable	LT	45	Fertility continuous		0.13	0.09	4											
Fun at sex	ST	45	Fertility continuous		−0.49	0.10	4											
Fun at sex	LT	45	Fertility continuous		0.04	0.09	4											
Sexually experienced	ST	45	Fertility continuous		−0.23	0.09	4											
Sexually experienced	LT	45	Fertility continuous		−0.08	0.09	4											
Likelihood of using threats to get sex	ST	45	Fertility continuous		0.26	0.09	4											
Likelihood of using threats to get sex	LT	45	Fertility continuous		0.47	0.10	4											
Moore et al. (2011), Study 2																		
Facial cues of testosterone	U	43	Within participants		−0.17	0.04				✓	✓							✓
Facial cues of cortisol	U	43	Within participants		1.18	0.06	4											
Moore (2011)																		
Intelligent	U	112	72	40	−0.16	0.21	4											

(*table continues*)

Table 1 (*continued*)

								Inclusion in analyses										
		Sample size						Long-term partner quality		Genetic quality								
Study and effects	Relationship context	*N*	High-fertility *n*	Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Facial masculinity	U	446	257	189	−0.04	0.10				✓	✓			✓				
Facial symmetry	U	446	257	189	−0.11	0.10				✓	✓	✓						
Morrison et al. (2010)																		
Male-typical facial movements	ST	47	Fertility continuous		−0.38	0.09				✓								
Flirtatious facial movements	ST	47	Fertility continuous		0.96	0.11	4											
Navarrete et al. (2009)																		
Body muscularity	U	21	9	12	M	M												
Same-race (vs. other-race) face	U	21	9	12	M	M												
Oinonen et al. (2008)																		
Facial symmetry	ST, LT, U	38	19	19	M	M												
Oinonen & Mazmanian (2007)																		
Facial symmetry	U	16	Within participants		−0.23	0.05				✓	✓	✓						
Pawlowski & Jasienska (2005)																		
Taller man relative to self	ST	99	37	62	0.46	0.04				✓	✓				✓			
Taller man relative to self	LT	108	39	69	0.24	0.04				✓	✓				✓			
Penton-Voak & Perrett (2000)																		
Facial masculinity	U	139	55	84	0.39	0.03				✓	✓			✓				
Penton-Voak et al. (1999), Study 1																		
Facial masculinity	U	39	Within participants		0.45	0.03				✓	✓			✓				
Penton-Voak et al. (1999), Study 2																		
Facial masculinity	ST	23	Within participants		0.23	0.08				✓	✓			✓				
Facial masculinity	LT	26	Within participants		−0.01	0.07				✓	✓			✓				
Perrett et al. (2013), Study 1																		
Facial masculinity	U	1290	527	763	0.04	0.00				✓	✓			✓				
Perrett et al. (2013), Study 2																		
Facial masculinity	ST	29	Within participants		0.32	0.04				✓	✓			✓				
Facial masculinity	LT	29	Within participants		0.27	0.03				✓	✓			✓				
Peters et al. (2008)																		
Face and body cues of semen quality	ST	25	Within participants		M	M												

(*table continues*)

Table 1 (*continued*)

								Inclusion in analyses										
		Sample size						Long-term partner quality		Genetic quality								
Study and effects	Relationship context	*N*	High-fertility *n*	Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Face and body averageness	ST	25	Within participants		M	M												
Face and body masculinity	ST	25	Within participants		M	M												
Face and body symmetry	ST	25	Within participants		M	M												
Peters et al. (2009)																		
Facial masculinity	ST	25	Within participants		0.00	0.08	7											
Body masculinity	ST	25	Within participants		−0.11	0.08	7											
Facial symmetry	ST	25	Within participants		0.06	0.08	7											
Body symmetry	ST	25	Within participants		−0.03	0.08	7											
Prokosch et al. (2009)																		
Creativity and intelligence	ST, LT	204	Fertility continuous		M	M	4											
Provost et al. (2008)																		
Male-typical walk	U	20	Within participants		0.45	0.05				✓								
Puts (2005)																		
Vocal masculinity	ST	137	Fertility continuous		0.42	0.03				✓	✓					✓		
Vocal masculinity	LT	137	Fertility continuous		0.47	0.03				✓	✓					✓		
Vocal cues of perceived physical dominance	ST, LT	136	38	98	M	M												
Vocal cues of perceived social dominance	ST, LT	136	38	98	M	M												
Rantala et al. (2006)																		
Scent cues of testosterone	U	36	11	25	−0.18	0.13				✓	✓							
Rantala et al. (2010)																		
Torso hair	U	186	62	124	−0.35	0.02				✓								
Rikowski & Grammer (1999)																		
Scent cues of body symmetry	U	40	14	26	M	M												
Scent cues of facial symmetry	U	40	14	26	M	M												
Roney & Simmons (2008)																		
Facial cues of testosterone	U	74	Fertility continuous		0.46	0.06				✓	✓							✓

(*table continues*)

Table 1 (*continued*)

								Inclusion in analyses										
		Sample size						Long-term partner quality		Genetic quality								
Study and effects	Relationship context	*N*	High-fertility *n*	Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Roney et al. (2011)																		
Facial cues of testosterone	U	18	Within participants		0.43	0.11				✓	✓							✓
Facial masculinity	U	18	Within participants		−0.23	0.10				✓	✓			✓				
Rupp, Librach, et al. (2009)																		
Facial masculinity	U	13	Fertility continuous		−0.64	0.39				✓	✓			✓				
Rupp, James, et al. (2009)																		
Facial masculinity	U	12	Within participants		M	M	6											
Singh & Bailey (2006)																		
Male-typical shoulder-to-hip ratio	ST	64	49	15	0.00	0.04				✓	✓				✓			
Male-typical shoulder-to-hip ratio	LT	130	91	39	0.11	0.02				✓	✓				✓			
Male-typical waist-to-hip ratio	ST	64	49	15	0.39	0.09				✓	✓				✓			
Male-typical waist-to-hip ratio	LT	130	91	39	−0.11	0.04				✓	✓				✓			
Soler et al. (2003), Study 1																		
Facial cues of semen quality	LT	52	8	44	M	M												
Soler et al. (2003), Study 2																		
Facial cues of semen quality	LT	76	30	46	M	M												
Teatero (2009)																		
Kindness	ST	14	Within participants		0.31	0.13		✓										
Kindness	LT	14	Within participants		0.01	0.13		✓										
Faithfulness	ST	14	Within participants		0.35	0.14		✓										
Faithfulness	LT	14	Within participants		0.45	0.14		✓										
Social status	ST	14	Within participants		0.17	0.13	4											
Social status	LT	14	Within participants		−0.03	0.13	4											
Financial resources	ST	14	Within participants		0.11	0.13		✓										

(*table continues*)

Table 1 (*continued*)

								Inclusion in analyses										
		Sample size						Long-term partner quality		Genetic quality								
Study and effects	Relationship context	*N*	High-fertility *n*	Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Broad set of measures	Narrow set of measures	Broad set of measures	Narrow set of measures	Facial symmetry	Scent cues of symmetry	Facial masculinity	Body masculinity	Vocal masculinity	Behavioral dominance	Facial cues of testosterone
Financial resources	LT	14	Within participants		0.26	0.13		✓										
Sense of humor	ST	14	Within participants		0.20	0.13	4											
Sense of humor	LT	14	Within participants		0.39	0.14	4											
Good parent	ST	14	Within participants		0.39	0.14		✓										
Good parent	LT	14	Within participants		−0.12	0.13		✓										
Intelligence	ST	14	Within participants		0.03	0.13	4											
Intelligence	LT	14	Within participants		0.00	0.13	4											
Thornhill & Gangestad (1999b)																		
Scent cues of body symmetry	U	48	Fertility continuous		0.94	0.11				✓	✓		✓					
Thornhill et al. (2013)																		
Scent cues of testosterone	U	48	Fertility continuous		0.66	0.10				✓	✓							
Scent cues of cortisol	U	48	Fertility continuous		0.55	0.09	4											
Thornhill et al. (2003)																		
Scent cues of body symmetry	U	65	Fertility continuous		0.55	0.09				✓	✓		✓					
Vaughn et al. (2010)																		
Facial masculinity	U	139	60	79	−0.01	0.03				✓	✓			✓				
Welling et al. (2007)																		
Facial masculinity	U	70	Within participants		0.26	0.01				✓	✓			✓				

Note. Checkmarks in the "Inclusion in analyses" columns indicate in which analyses an effect was included (a blank indicates that the effect was not included in a given analysis). For nonmissing effects excluded from all analyses, the "Reason for exclusion" column indicates the specific inclusion criterion that the effect did not satisfy (i.e., the reason it was excluded). For example, "4" refers to Inclusion Criterion 4. ST = short term; LT = long term; U = unspecified; M = missing data; CPI = California Psychological Inventory.

Table 2

Studies Assessing Ovulation-Related Cycle Shifts in Mate Preferences: Study and Effect Characteristics

Study and effects	Relationship context	Country	Sample type	Setting	Design	Average conception probability		Cycle position estimation method	Stimuli	Method to determine amount of male trait	Stated versus revealed preferences	Task	Number of trials
						High fertility	Low fertility						
Beaulieu (2007), Study 2													
Relationship skills (composite of kind, understanding, loyal, generous)	U	United States	Col	L	B	0.09	0.01	Forw	SR	N/A	Stated	Ratings	1
Dominance (composite of dominant, powerful, aggressive)	U	United States	Col	L	B	0.09	0.01	Forw	SR	N/A	Stated	Ratings	1
Education (composite of educated, cultured, intelligent)	U	United States	Col	L	B	0.09	0.01	Forw	SR	N/A	Stated	Ratings	1
Good financial prospects (composite of wealthy, good financial prospects)	U	United States	Col	L	B	0.09	0.01	Forw	SR	N/A	Stated	Ratings	1
Beaulieu (2007), Study 4													
Relationship skills (composite of kind, understanding, loyal, generous)	ST, LT	United States	Col	L	B	0.09	0.01	Forw	SR	N/A	Stated	Ratings	1
Dominance (composite of dominant, powerful, aggressive)	ST, LT	United States	Col	L	B	0.09	0.01	Forw	SR	N/A	Stated	Ratings	1
Education (composite of educated, cultured, intelligent)	ST, LT	United States	Col	L	B	0.09	0.01	Forw	SR	N/A	Stated	Ratings	1
Good financial prospects (composite of wealthy, good financial prospects)	ST, LT	United States	Col	L	B	0.09	0.01	Forw	SR	N/A	Stated	Ratings	1
Bressan & Stranieri (2008)													
Facial masculinity	U	Italy	Col	L	B	0.06	0.01	Forw	FP	SSR	Revealed	Ratings	12
Bullock (2000), Chapter 4													
Chin length	U	Canada	Col	L	B	0.06	0.02	Forw	FP	Manip	Revealed	Ratings	5
Cárdenas & Harris (2007)													
Facial symmetry	ST	United States, Chile	Col	L	W	0.09	0.01	Forw	FP	Manip	Revealed	TOFC	54
Caryl et al. (2009)													
Pupil size	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Arrogant	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Ingenious	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Aggressive	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1

(table continues)

Table 2 (*continued*)

Study and effects	Relationship context	Country	Sample type	Setting	Design	Average conception probability		Cycle position estimation method	Stimuli	Method to determine amount of male trait	Stated versus revealed preferences	Task	Number of trials
						High fertility	Low fertility						
Strong	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Conceited	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Enterprising	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Inventive	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Warm	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Sensitive	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Sentimental	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Sympathetic	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Jolly	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Helpful	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Appreciative	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Considerate	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Cooperative	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Friendly	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Talkative	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Forgiving	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Emotional	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Foresighted	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Shrewd	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Industrious	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Assertive	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Forceful	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Timid	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Dependent	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Fickle	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Frivolous	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Opportunistic	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Hardheaded	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
Confident	U	Scotland	Col	L	B	0.06	0.02	Forw	SR	N/A	Stated	Ratings	1
DeBruine et al. (2005)													
Facial self-resemblance	U	Scotland	Col/Com	L	B	0.06	0.02	Forw	FP	Manip	Revealed	TOFC	36
Feinberg et al. (2006)													
Vocal masculinity	U	Scotland	Col/Com	L	W	0.07	0.02	Rev, VM	VR	Manip	Revealed	Ratings	8
Feinberg (2012)													
Vocal masculinity	ST	Canada	Col	L	W	0.07	0.02	Rev, VM	VR	Manip	Revealed	TOFC	4
Vocal masculinity	LT	Canada	Col	L	W	0.07	0.02	Rev, VM	VR	Manip	Revealed	TOFC	4
Fink (2012)													
Facial masculinity	U	Austria, United Kingdom	M	M	W	0.06	0.02	Forw	FP	Manip	Revealed	Slider	1
Flowe et al., 2012													
Behavioral masculinity	U	United Kingdom	Col	L	B	0.08	0.01	Forw	VB, D	Manip	Revealed	Ratings	1
Frost (1994)													
Darker skin tone	U	Canada	Col	L	B	0.08	0.01	Forw	FP	Manip	Revealed	TOFC	3

(*table continues*)

Table 2 (*continued*)

Study and effects	Relationship context	Country	Sample type	Setting	Design	Average conception probability: High fertility	Average conception probability: Low fertility	Cycle position estimation method	Stimuli	Method to determine amount of male trait	Stated versus revealed preferences	Task	Number of trials
Gangestad et al. (2004)													
Social presence	ST	United States	Col	L	B	Continuous		Forw	VB	Meas	Revealed	Ratings	38
Social presence	LT	United States	Col	L	B	Continuous		Forw	VB	Meas	Revealed	Ratings	38
Direct intrasexual competitiveness	ST	United States	Col	L	B	Continuous		Forw	VB	Meas	Revealed	Ratings	38
Direct intrasexual competitiveness	LT	United States	Col	L	B	Continuous		Forw	VB	Meas	Revealed	Ratings	38
Gangestad et al. (2007)													
Muscular	ST	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Muscular	LT	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Confrontative (with other men)	ST	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Confrontative (with other men)	LT	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Socially respected and influential	ST	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Socially respected and influential	LT	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Arrogant and self-centered	ST	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Arrogant and self-centered	LT	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Intelligent	ST	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Intelligent	LT	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Faithful	ST	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Faithful	LT	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Warm (kind and understanding)	ST	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Warm (kind and understanding)	LT	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Likely to be financially successful	ST	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Likely to be financially successful	LT	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Likely to be a good parent	ST	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Likely to be a good parent	LT	United States	Col	L	B	Continuous		Forw	VB	SSR	Revealed	Ratings	38
Gangestad et al. (2011)													
Facial masculinity	U	United States	M	L	W	0.23	0.02	LH	FP	Manip	Revealed	Slider	1
Gangestad & Thornhill (1998)													
Scent cues of body symmetry	U	United States	Col/Com	L	B	Continuous		Forw	SS	Meas	Revealed	Ratings	42

(*table continues*)

Table 2 (*continued*)

Study and effects	Relationship context	Country	Sample type	Setting	Design	Average conception probability		Cycle position estimation method	Stimuli	Method to determine amount of male trait	Stated versus revealed preferences	Task	Number of trials
						High fertility	Low fertility						
Garver-Apgar & Gangestad (2012)													
Average of social presence and direct intrasexual competitiveness	ST	United States	Col	L	W	0.26	0.02	LH	VB	Meas	Revealed	Ratings	38
Average of social presence and direct intrasexual competitiveness	LT	United States	Col	L	W	0.26	0.02	LH	VB	Meas	Revealed	Ratings	38
Harris (2011)													
Facial masculinity	U	Canada, United States	Com	F	B	0.07	0.01	Forw	FP	Manip	Revealed	MOFC	5
Haselton & Miller (2006)													
Wealth versus creativity	ST, LT	United States	Col	L	B	Continuous		Forw	D	Manip	Revealed	TOFC, Ratings	4
Havlíček et al. (2005)													
Scent cues of dominance (Narcissism scale from CPI)	U	Czech Republic	Col	L	B	0.08	0.02	Forw	SS	Meas	Revealed	Ratings	10
Hromatko et al. (2006)													
Facial symmetry	U	Croatia	M	L	B	0.09	0.01	Forw & Rev	FP	Manip	Revealed	Ratings	40
Izbicki & Johnson (2010)													
Facial masculinity	ST	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Facial masculinity	LT	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Strong	ST	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Strong	LT	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Warm	ST	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Warm	LT	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Mature	ST	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Mature	LT	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Socially competent	ST	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Socially competent	LT	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Nurturant	ST	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Nurturant	LT	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Threatening	ST	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Threatening	LT	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Dominant	ST	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Dominant	LT	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Dark	ST	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Dark	LT	United States	M	L	W	0.08	0.02	Rev	FP	PR	Revealed	Ratings	12
Johnston et al. (2001)													
Facial masculinity	U	United States	Col	L	W	0.06	0.02	Rev, VM	FP	Manip	Revealed	Slider	1

(*table continues*)

Table 2 (*continued*)

Study and effects	Relationship context	Country	Sample type	Setting	Design	Average conception probability: High fertility	Average conception probability: Low fertility	Cycle position estimation method	Stimuli	Method to determine amount of male trait	Stated versus revealed preferences	Task	Number of trials
Jones, Little, et al. (2005), Study 2													
Facial masculinity	U	United Kingdom	Com	F	B	0.07	0.02	Rev	FP	Manip	Revealed	TOFC	7
Koehler et al. (2002)													
Facial symmetry	ST, LT	Australia	Col	L	W	0.09	0.01	Forw	FP	Manip	Revealed	Ratings	24
Koehler et al. (2006)													
Facial averageness	U	Australia	Col	L	W	0.09	0.01	Forw	FP	Manip	Revealed	TOFC	24
Facial symmetry	U	Australia	Col	L	W	0.09	0.01	Forw	FP	Manip	Revealed	TOFC	24
Li et al. (2006)													
Multiple traits	ST, LT	United States	Col	L	W	M	M	Forw	SR	N/A	Stated	Mate dollars	M
Little, Jones, et al. (2007), Study 1													
Facial symmetry	U	United Kingdom	M	L	W	0.08	0.01	Rev	FP	Manip	Revealed	TOFC	6
Little, Jones, et al. (2007), Study 2													
Facial symmetry	ST	United Kingdom	Col/Com	L, F	B	0.06	0.02	Forw	FP	Manip	Revealed	TOFC	15
Facial symmetry	LT	United Kingdom	Col/Com	L, F	B	0.06	0.02	Forw	FP	Manip	Revealed	TOFC	15
Little, Jones, & Burriss (2007), Study 1													
Body masculinity	ST	United Kingdom	Com	F	B	0.06	0.02	Forw	BP	Manip	Revealed	TOFC	10
Body masculinity	LT	United Kingdom	Com	F	B	0.06	0.02	Forw	BP	Manip	Revealed	TOFC	10
Little, Jones, & Burriss (2007), Study 2													
Body masculinity	ST	United Kingdom	M	L	W	0.06	0.02	Forw	BP	Manip	Revealed	TOFC	10
Body masculinity	LT	United Kingdom	M	L	W	0.06	0.02	Forw	BP	Manip	Revealed	TOFC	10
Little et al. (2008)													
Facial masculinity	U	United Kingdom	Com	F	B	0.06	0.02	Forw	FP	SSR	Revealed	TOFC	10
Luevano & Zebrowitz (2006)													
Dominant	ST	United States	M	L	W	0.08	0.02	Rev	FP	SSR	Revealed	Ratings	153
Dominant	LT	United States	M	L	W	0.08	0.02	Rev	FP	SSR	Revealed	Ratings	153
Facial masculinity	ST	United States	M	L	W	0.08	0.02	Rev	FP	SSR	Revealed	Ratings	153
Facial masculinity	LT	United States	M	L	W	0.08	0.02	Rev	FP	SSR	Revealed	Ratings	153
Warm	ST	United States	M	L	W	0.08	0.02	Rev	FP	SSR	Revealed	Ratings	153
Warm	LT	United States	M	L	W	0.08	0.02	Rev	FP	SSR	Revealed	Ratings	153
Lukaszewski & Roney (2009)													
Dominant	ST	United States	Col	L	B	Continuous		Forw	SR	N/A	Stated	Ratings	1
Dominant	LT	United States	Col	L	B	Continuous		Forw	SR	N/A	Stated	Ratings	1
Kind	ST	United States	Col	L	B	Continuous		Forw	SR	N/A	Stated	Ratings	1
Kind	LT	United States	Col	L	B	Continuous		Forw	SR	N/A	Stated	Ratings	1
Trustworthy	ST	United States	Col	L	B	Continuous		Forw	SR	N/A	Stated	Ratings	1
Trustworthy	LT	United States	Col	L	B	Continuous		Forw	SR	N/A	Stated	Ratings	1
McClellan et al. (2007)													

(*table continues*)

Table 2 (*continued*)

Study and effects	Relationship context	Country	Sample type	Setting	Design	Average conception probability		Cycle position estimation method	Stimuli	Method to determine amount of male trait	Stated versus revealed preferences	Task	Number of trials
						High fertility	Low fertility						
Body masculinity	ST, LT, U	United States	Col	L	B	0.05	0.02	Forw	BP[a]	Meas, PR	Revealed	Rank	M
Age	ST, LT, U	United States	Col	L	B	0.05	0.02	Forw	BP[a]	Meas, PR	Revealed	Rank	M
McDonald & Navarrete (2012), Sample 1													
Body muscularity	U	United States	Col	L	B	0.09	0.01	Forw	BAv	Manip	Revealed	Ratings	10
Same-race (vs. other-race) face	U	United States	Col	L	B	0.09	0.01	Forw	FAv	Manip	Revealed	Ratings	10
McDonald & Navarrete (2012), Sample 2													
Body muscularity	U	United States	Col	L	B	0.09	0.01	Forw	BAv	Manip	Revealed	Ratings	M
Same-race (vs. other-race) face	U	United States	Col	L	B	0.09	0.01	Forw	FAv	Manip	Revealed	Ratings	M
Miller (2003)													
Intelligent	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Intelligent	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Future kids' intelligence	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Future kids' intelligence	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Mathematical problem-solving ability	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Mathematical problem-solving ability	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Good grades	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Good grades	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Creative/imaginative	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Creative/imaginative	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Future kids' sense of humor	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Future kids' sense of humor	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Social sensitivity	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Social sensitivity	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Adaptable to situations and challenges	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Adaptable to situations and challenges	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Big ego	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Big ego	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Body muscularity	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Body muscularity	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Facial masculinity	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Facial masculinity	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Tall	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3

(*table continues*)

Table 2 (*continued*)

Study and effects	Relationship context	Country	Sample type	Setting	Design	Average conception probability: High fertility	Average conception probability: Low fertility	Cycle position estimation method	Stimuli	Method to determine amount of male trait	Stated versus revealed preferences	Task	Number of trials
Tall	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Happy	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Happy	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Exciting/spontaneous	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Exciting/spontaneous	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Talkative/extraverted	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Talkative/extraverted	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Likelihood of being unfaithful (reverse-coded)	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Likelihood of being unfaithful (reverse-coded)	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Future money making	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Future money making	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Good at playing with and caring for kids	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Good at playing with and caring for kids	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Future career success	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Future career success	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Sympathetic/kind	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Sympathetic/kind	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Constructive in arguments	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Constructive in arguments	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Neat/organized	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Neat/organized	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Moody/irritable	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Moody/irritable	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Fun at sex	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Fun at sex	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Sexually experienced	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Sexually experienced	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Likelihood of using threats to get sex	ST	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Likelihood of using threats to get sex	LT	United States	Col	L	B	Continuous		Forw	D	PR	Revealed	Ratings	3
Moore et al. (2011), Study 2													
Facial cues of testosterone	U	Scotland	Col	M	W	0.07	0.02	Rev	FP	Manip	Revealed	Ratings	16
Facial cues of cortisol	U	Scotland	Col	M	W	0.07	0.02	Rev	FP	Manip	Revealed	Ratings	16

(*table continues*)

Table 2 (*continued*)

Study and effects	Relationship context	Country	Sample type	Setting	Design	Average conception probability		Cycle position estimation method	Stimuli	Method to determine amount of male trait	Stated versus revealed preferences	Task	Number of trials
						High fertility	Low fertility						
Moore (2011)													
Intelligent	U	M	Com	L	B	0.08	0.01	Forw	FP	Manip	Revealed	Ratings	6
Facial masculinity	U	M	Com	L	B	0.08	0.01	Forw	FP	Manip	Revealed	Ratings	16
Facial symmetry	U	M	Com	L	B	0.08	0.01	Forw	FP	Manip	Revealed	Ratings	6
Morrison et al. (2010)													
Male-typical facial movements	ST	United Kingdom	Col	L	B	Continuous		Forw	MFO	SSR	Revealed	Ratings	30
Flirtatious facial movements	ST	United Kingdom	Col	L	B	Continuous		Forw	MFO	SSR	Revealed	Ratings	30
Navarrete et al. (2009)													
Body muscularity	U	United States	Col	L	B	0.09	0.01	Forw	BAv	Manip	Revealed	Ratings	4
Same-race (vs. other-race) face	U	United States	Col	L	B	0.09	0.01	Forw	FAv	Manip	Revealed	Ratings	4
Oinonen et al. (2008)													
Facial symmetry	ST, LT, U	Canada	Col/Com	L	W	0.07	0.01	Forw & Rev	FP	Manip	Revealed	Ratings	80
Oinonen & Mazmanian (2007)													
Facial symmetry	U	Canada	Col/Com	L	W	0.07	0.02	Rev	FP	Manip	Revealed	TOFC	80
Pawlowski (2005)													
Taller man relative to self	ST	Poland	Com	L	B	0.04	0.03	Rev, VM	BOD	Manip	Revealed	TOFC	45
Taller man relative to self	LT	Poland	Com	L	B	0.04	0.03	Rev, VM	BOD	Manip	Revealed	TOFC	45
Penton-Voak & Perrett (2000)													
Facial masculinity	U	United Kingdom	Com	*F*(mag)	B	0.06	0.02	Forw	FP	Manip	Revealed	MOFC	1
Penton-Voak et al. (1999), Study 1													
Facial masculinity	U	Japan	Col/Com	L	W	0.04	0.02	Rev	FP	Manip	Revealed	MOFC	10
Penton-Voak et al. (1999), Study 2													
Facial masculinity	ST	United Kingdom	Col	M	W	0.04	0.02	Rev	FP	Manip	Revealed	MOFC	1
Facial masculinity	LT	United Kingdom	Col	M	W	0.04	0.02	Rev	FP	Manip	Revealed	MOFC	1
Perrett et al. (2013), Study 1													
Facial masculinity	U	M	Com	F	B	0.04	0.02	Rev	FP	Manip	Revealed	TOFC	3
Perrett et al. (2013), Study 2													
Facial masculinity	ST	United Kingdom	Col	L	W	M	M	M	FP	Manip	Revealed	TOFC	6
Facial masculinity	LT	United Kingdom	Col	L	W	M	M	M	FP	Manip	Revealed	TOFC	6
Peters et al. (2008)													
Face and body cues of semen quality	ST	Australia	M	L	W	0.32	M	LH	BP, FP	Meas	Revealed	Ratings	101

(*table continues*)

Table 2 (*continued*)

Study and effects	Relationship context	Country	Sample type	Setting	Design	Average conception probability: High fertility	Average conception probability: Low fertility	Cycle position estimation method	Stimuli	Method to determine amount of male trait	Stated versus revealed preferences	Task	Number of trials
Face and body averageness	ST	Australia	M	L	W	0.32	M	LH	BP, FP	SSR	Revealed	Ratings	116
Face and body masculinity	ST	Australia	M	L	W	0.32	M	LH	BP, FP	SSR	Revealed	Ratings	116
Face and body symmetry	ST	Australia	M	L	W	0.32	M	LH	BP, FP	SSR	Revealed	Ratings	116
Peters et al. (2009)													
Facial masculinity	ST	Australia	M	L	W	0.32	M	LH	FP	SSR	Revealed	Ratings	117
Body masculinity	ST	Australia	M	L	W	0.32	M	LH	BP	SSR	Revealed	Ratings	117
Facial symmetry	ST	Australia	M	L	W	0.32	M	LH	FP	SSR	Revealed	Ratings	117
Body symmetry	ST	Australia	M	L	W	0.32	M	LH	BP	SSR	Revealed	Ratings	117
Prokosch et al. (2009)													
Creativity and intelligence	ST, LT	United States	Col	L	B	Continuous		Rev	VB	PR & Meas	Revealed	Ratings	5
Provost et al. (2008)													
Male-typical walk	U	Canada	Col	L	W	0.08	0.01	Forw, Saliv	PLW	Manip	Revealed	Slider	1
Puts (2005)													
Vocal masculinity	ST	United States	Col	L	B	Continuous		Rev	VR	Manip	Revealed	Ratings	30
Vocal masculinity	LT	United States	Col	L	B	Continuous		Rev	VR	Manip	Revealed	Ratings	30
Vocal cues of perceived physical dominance	ST	United States	Col	L	B	0.07	0.02	Rev	VR	SSR	Revealed	Ratings	33
Vocal cues of perceived physical dominance	LT	United States	Col	L	B	0.07	0.02	Rev	VR	SSR	Revealed	Ratings	31
Vocal cues of perceived social dominance	ST	United States	Col	L	B	0.07	0.02	Rev	VR	SSR	Revealed	Ratings	34
Vocal cues of perceived social dominance	LT	United States	Col	L	B	0.07	0.02	Rev	VR	SSR	Revealed	Ratings	32
Rantala et al. (2006)													
Scent cues of testosterone	U	Finland	Col	L	B	0.08	0.02	LH	SS	Meas	Revealed	Ratings	19
Rantala et al. (2010)													
Torso hair	U	Finland	Com	L	B	0.08	0.02	Rev	BP	Manip	Revealed	TOFC	20
Rikowski & Grammer (1999)													
Scent cues of body symmetry	U	Austria	Col	L	B	0.06	0.01	Forw	SS	Meas	Revealed	Ratings	8
Scent cues of facial symmetry	U	Austria	Col	L	B	0.06	0.01	Forw	SS	Meas	Revealed	Ratings	8

(*table continues*)

Table 2 (*continued*)

Study and effects	Relationship context	Country	Sample type	Setting	Design	Average conception probability		Cycle position estimation method	Stimuli	Method to determine amount of male trait	Stated versus revealed preferences	Task	Number of trials
						High fertility	Low fertility						
Roney & Simmons (2008)													
Facial cues of testosterone	U	United States	Col	L	B	Continuous		Forw	FP	Meas	Revealed	Ratings	37
Roney et al. (2011)													
Facial cues of testosterone	U	United States	Col	L	W	0.08	0.02	Rev, VM	FP	Manip	Revealed	TOFC	7
Facial masculinity	U	United States	Col	L	W	0.08	0.02	Rev, VM	FP	Manip	Revealed	TOFC	7
Rupp, Librach, et al. (2009)													
Facial masculinity	U	United States	Col	F	B	Continuous		Forw	FP	SSR	Revealed	Ratings	510
Rupp, James, et al. (2009)													
Facial masculinity	U	United States	Col	L	W	0.08	0.02	Forw	FP	Manip	Revealed	Ratings	224
Singh & Bailey (2006)													
Male-typical shoulder-to-hip ratio	ST	United States	M	M	B	0.06	0.01	Forw	BOD	Manip	Revealed	Ratings	6
Male-typical shoulder-to-hip ratio	LT	United States	M	M	B	0.06	0.01	Forw	BOD	Manip	Revealed	Ratings	6
Male-typical waist-to-hip ratio	ST	United States	M	M	B	0.06	0.01	Forw	BOD	Manip	Revealed	Ratings	6
Male-typical waist-to-hip ratio	LT	United States	M	M	B	0.06	0.01	Forw	BOD	Manip	Revealed	Ratings	6
Soler et al. (2003), Study 1													
Facial cues of semen quality	LT	Spain	Col	L	B	M	M	Forw	FP	Meas	Revealed	Ratings	66
Soler et al. (2003), Study 2													
Facial cues of semen quality	LT	Spain	Col	L	B	M	M	Forw	FP	Meas	Revealed	Ratings	12
Teatero (2009)													
Kindness	ST	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A		Ratings	1
Kindness	LT	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1
Faithfulness	ST	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1
Faithfulness	LT	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1
Social status	ST	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1
Social status	LT	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1
Financial resources	ST	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1
Financial resources	LT	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1
Sense of humor	ST	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1
Sense of humor	LT	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1
Good parent	ST	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1
Good parent	LT	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1
Intelligence	ST	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1
Intelligence	LT	Canada	Col/Com	F	W	0.09	0.01	Rev, VM	SR	N/A	Stated	Ratings	1

(*table continues*)

Table 3

All Hypothesized Cues of Ancestral Genetic Quality: Broad Set of Measures

Effect	Coefficient	*SE*	Variance component	*SD*	*t* ratio	*df*	χ^2	*p*
Step 1								
Fixed								
Overall weighted mean effect size, γ_{00}	0.15	0.04			4.13	49		<.001
Random								
True mean effect size, δ_{0j}			0.03	0.18		49	141.32	<.001
Step 2								
Fixed								
Weighted mean effect size in a short-term context, γ_{00}	0.21	0.06			3.54	49		.001
Difference between a long-term and short-term context, γ_{10}	−0.15	0.04			−3.28	93		.002
Difference between an unspecified and short-term context, γ_{20}	−0.05	0.08			−0.62	93		.54
Random								
True mean effect size in a short-term context, δ_{0j}			0.03	0.18		49	138.33	<.001

Note. Step 1: Results from unconditional multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for all hypothesized cues of ancestral genetic quality in the sample of effects selected using relatively relaxed inclusion criteria. Step 2: Results from multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for all hypothesized cues of ancestral genetic quality (relaxed inclusion criteria) in a short-term relationship context (compared to a long-term or unspecified relationship context).

Table 4

All Hypothesized Cues of Ancestral Genetic Quality: Narrow Set of Measures

Effect	Coefficient	*SE*	Variance component	*SD*	*t* ratio	*df*	χ^2	*p*
	Step 1							
Fixed								
Overall weighted mean effect size, γ_{00}	0.17	0.04			4.33	41		<.001
Random								
True mean effect size, δ_{0j}			0.03	0.18		42	111.48	<.001
	Step 2							
Fixed								
Weighted mean effect size in a short-term context, γ_{00}	0.26	0.07			4.07	41		<.001
Difference between a long-term and short-term context, γ_{10}	−0.24	0.05			−4.52	65		<.001
Difference between an unspecified and short-term context, γ_{20}	−0.07	0.08			−0.82	65		.42
Random								
True mean effect size in a short-term context, δ_{0j}			0.03	0.18		41	108.23	<.001

Note. Step 1: Results from unconditional multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for all hypothesized cues of ancestral genetic quality in the sample of effects selected using relatively relaxed inclusion criteria. Step 2: Results from multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for all hypothesized cues of ancestral genetic quality (relaxed inclusion criteria) in a short-term relationship context (compared to a long-term or unspecified relationship context).

Table 5

Facial Symmetry

Effect	Coefficient	*SE*	Variance component	*SD*	*t* ratio	*df*	χ^2	*p*
			Step 1					
Fixed								
Overall weighted mean effect size, γ_{00}	0.07	0.1			0.75	6		.48
Random								
True mean effect size, δ_{0j}			0.03	0.16		6	10.55	.1
			Step 2					
Fixed								
Weighted mean effect size in a short-term context, γ_{00}	0.3	0.2			1.47	6		.19
Difference between a long-term and short-term context, γ_{10}	−0.46	0.21			−2.2	5		.08
Difference between an unspecified and short-term context, γ_{20}	−0.32	0.26			−1.24	5		.27
Random								
True mean effect size in a short-term context, δ_{0j}			0.06	0.24		6	13.41	.04

Note. Step 1: Results from unconditional multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for facial symmetry in the sample of effects selected using relatively strict inclusion criteria. Step 2: Results from multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for facial symmetry (strict inclusion criteria) in a short-term relationship context (compared to a long-term or unspecified relationship context).

Table 6

Scent Cues of Face and Body Symmetry

Effect	Coefficient	*SE*	Variance component	*SD*	*t* ratio	*df*	χ^2	*p*
			Step 1					
Fixed								
Overall weighted mean effect size, γ_{00}	0.83	0.20			4.15	2		.14
Random								
True mean effect size, δ_{0j}			0.0003	0.02		2	1.82	>.50

Note. Step 1: Results from unconditional multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for scent cues of symmetry in the sample of effects selected using relatively strict inclusion criteria.

Table 7

Structural Facial Masculinity

Effect	Coefficient	*SE*	Variance component	*SD*	*t* ratio	*df*	χ^2	*p*
			Step 1					
Fixed								
Overall weighted mean effect size, γ_{00}	0.13	0.06			2.09	18		.05
Random								
True mean effect size, δ_{0j}			0.04	0.2		18	51.25	<.001
			Step 2					
Fixed								
Weighted mean effect size in a short-term context, γ_{00}	−0.02	0.14			−0.11	18		.91
Difference between a long-term and short-term context, γ_{10}	0.01	0.14			0.05	20		.96
Difference between an unspecified and short-term context, γ_{20}	0.19	0.16			1.2	20		.24
Random								
True mean effect size in a short-term context, δ_{0j}			0.04	0.19		18	46.32	<.001

Note. Step 1: Results from unconditional multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for facial masculinity in the sample of effects selected using relatively strict inclusion criteria. Step 2: Results from multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for facial masculinity (strict inclusion criteria) in a short-term relationship context (compared to a long-term or unspecified relationship context).

Table 8

Structural Body Masculinity

Effect	Coefficient	*SE*	Variance component	*SD*	*t* ratio	*df*	χ^2	*p*
	Step 1							
Fixed								
Overall weighted mean effect size, γ_{00}	0.21	0.08			2.45	4		.07
Random								
True mean effect size, δ_{0j}			0.02	0.14		4	9.52	.05
	Step 2							
Fixed								
Weighted mean effect size in a short-term context, γ_{00}	0.35	0.1			3.49	4		.04
Difference between a long-term and short-term context, γ_{10}	−0.27	0.11			−2.46	10		.03
Random								
True mean effect size in a short-term context, δ_{0j}			0.02	0.13		4	8.68	.07

Note. Step 1: Results from unconditional multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for body masculinity in the sample of effects selected using relatively strict inclusion criteria. Step 2: Results from multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for body masculinity (strict inclusion criteria) in a short-term relationship context (compared to a long-term relationship context).

Table 9

Vocal Masculinity

Effect	Coefficient	*SE*	Variance component	*SD*	*t* ratio	*df*	χ^2	*p*
Step 1								
Fixed								
Overall weighted mean effect size, γ_{00}	0.28	0.18			1.61		(Unable to compute)	
Random								
True mean effect size, δ_{0j}			0.04	0.2		1	3.03	.08
Step 2								
Fixed								
Weighted mean effect size in a short-term context, γ_{00}	0.4	0.2			1.98		(Unable to compute)	
Difference between a long-term and short-term context, γ_{10}	−0.21	0.2			−1.08	2		.39
Random								
True mean effect size in a short-term context, δ_{0j}			0.04	0.19		1	2.86	.09

Note. Step 1: Results from unconditional multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for vocal masculinity in the sample of effects selected using relatively strict inclusion criteria. Step 2: Results from multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for vocal masculinity (strict inclusion criteria) in a short-term relationship context (compared to a long-term relationship context).

Table 10

Behavioral Dominance and Felt Superiority Over Other Men

Effect	Coefficient	*SE*	Variance component	*SD*	*t* ratio	*df*	χ^2	*p*
	Step 1							
Fixed								
Overall weighted mean effect size, γ_{00}	0.04	0.06			0.71	2		.55
Random								
True mean effect size, δ_{0j}			0.004	0.06		2	2.49	.29
	Step 2							
Fixed								
Weighted mean effect size in a short-term context, γ_{00}	0.19	0.07			2.65	2		.09
Difference between a long-term and short-term context, γ_{10}	−0.3	0.08			−3.65	10		.01
Random								
True mean effect size in a short-term context, δ_{0j}			0.0004	0.06		2	2.57	.28

Note. Step 1: Results from unconditional multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for behavioral dominance and felt superiority over other men in the sample of effects selected using relatively strict inclusion criteria. Step 2: Results from multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for behavioral dominance and felt superiority over other men (strict inclusion criteria) in a short-term relationship context (compared to a long-term relationship context).

Table 11

Facial Cues of Testosterone

Effect	Coefficient	*SE*	Variance component	*SD*	*t* ratio	*df*	χ^2	*p*
			Step 1					
Fixed								
Overall weighted mean effect size, γ_{00}	0.20	0.22			0.9	2		.46
Random								
True mean effect size, δ_{0j}			0.08	0.29		2	4.91	.08

Note. Step 1: Results from unconditional multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for facial cues of testosterone in the sample of effects selected using relatively strict inclusion criteria.

Table 12

All Hypothesized Cues of Ancestral Long-Term Partner Quality: Broad Set of Measures

Effect	Coefficient	*SE*	Variance component	*SD*	*t* ratio	*df*	χ^2	*p*
			Step 1					
Fixed								
Overall weighted mean effect size, γ_{00}	−0.004	0.04			−0.11	7		.91
Random								
True mean effect size, δ_{0j}			0.002	0.04		7	6.97	>.50
			Step 2					
Fixed								
Weighted mean effect size in a short-term context, γ_{00}	−0.06	0.05			−1.13	7		.3
Difference between a long-term and short-term context, γ_{10}	0.11	0.06			1.76	35		.09
Difference between an unspecified and short-term context, γ_{20}	0.02	0.18			0.1	35		.92
Random								
True mean effect size in a short-term context, δ_{0j}			0.002	0.05		7	7.17	.41

Note. Step 1: Results from unconditional multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for all hypothesized cues of ancestral long-term partner quality in the sample of effects selected using relatively relaxed inclusion criteria. Step 2: Results from multilevel model estimating the true mean standardized mean difference between high and low fertility in women's preference for all hypothesized cues of ancestral long-term partner quality (strict inclusion criteria) in a short-term relationship context (compared to a long-term or unspecified relationship context).

Table 13

All Hypothesized Cues of Ancestral Long-Term Partner Quality: Narrow Set of Measures

Effect	Coefficient	*SE*	*t* ratio	*df*	*p*
		Step 1			
Fixed					
Overall weighted mean effect size, γ_{00}	−0.05	0.05	−1.174	7	.28
		Step 2			
Fixed					
Weighted mean effect size in a short-term context, γ_{00}	−0.12	0.07	−1.89	6	.11
Difference between a long-term and short-term context, γ_{10}	0.14	0.09	1.49	6	.19

Note. Step 1: Results from unconditional multilevel model estimating the true mean standardized mean difference (*g*) between high and low fertility in women's preference for all hypothesized cues of ancestral long-term partner quality in the sample of effects selected using relatively strict inclusion criteria. Because this sample consisted of eight effects from a single study, these are least squares estimates. Step 2: Results from multilevel model estimating the true mean *g* between high and low fertility in women's preference for all hypothesized cues of ancestral long-term partner quality (strict inclusion criteria) in a short-term relationship context (compared to a long-term relationship context).

Table 14

Studies Examining High- versus Low-Fertility Attractiveness and Related Qualities: Effect Size, Publication Status, Sample Size, and Analyses in which Effects were Included

				Women		Analyses																			
						Other-rated Attractiveness and Related Features						Attractiveness and Related Features					Attractiveness-enhancing Behavior						Men's Responses		
Study and Effect(s)	*g*	Var	Pub	High-fertility *n*	Low-fertility *n*	Overall Attr.	Overall Attr. with Correlates	Scent Attr.	Scent Intensity	Facial Attr.	Voice Pitch (vowels)	Overall Attr.	Facial Attr.	Body Attr.	Sexiness	ST Mate Value	Overall Attr. Enhancement	Hypothetical Clothing Task	Walk Attr.	Sexy Clothing	Voice Pitch (social)	Flirtatiousness	Overall Men's Responses	Men's T Response	Partner Jealousy
Adams unpublished Master's thesis			U	33 (within-participants)																					
Self-perceived Mating Success	M	M																							
Self-perceived Mating Effort (Flirtation)	0.13	0.02															Y					Y			
Self-perceived Mate Value (MV-2)	0.47	0.01										Y													
Self-perceived Long-term Mate Value (MV-7)	0.18	0.01																							
Self-perceived Attractive Face (MV-7 Item 1)	0.53	0.02											Y												
Self-perceived Attractive Body (MV-7 Item 2)	0.31	0.01												Y											
Self-perceived Ambition (MV-7 Item 3)	-0.08	0.00																							
Self-perceived Desires Children (MV-7 Item 4)	0.28	0.00																							
Self-perceived Enthusiastic about Sex (MV-7 Item 5)	-0.03	0.01																							
Self-perceived Faithful (MV-7 Item 6)	0.45	0.01																							
Self-perceived Financially Secure (MV-7 Item 7)	-0.07	0.01																							
Beall & Tracy (2013) - Sample A			P	51	49																				
Red/pink clothing	0.74	0.11															Y								
Beall & Tracy (2013) - Sample B			P	10	14																				
Red/pink clothing	1.15	0.42															Y								
Beaulieu (2007)			P	33	59																				
Combined effect across a, b, c, d	0.19	0.03										Y													
Composite of a, b, c	0.16	0.05																							
a. Self-rated facial attractiveness	0.31	0.05											Y												
b. Self-rated body attractiveness	-0.05	0.05												Y											
c. Self-rated sexiness	0.17	0.05													Y										
d. Self-rated short-term mate value	0.22	0.05														Y									
Self-rated long-term mate value	0.21	0.05																							
Bleske-Recheck (2011)			P	17 (within-participants)																					
Combined effect across a, b, c	0.30	0.03										Y													
Combined effect across a, b	0.37	0.03																							
a. Self-rated attractiveness	0.33	0.06																							
b. Self-rated ugliness	0.41	0.03																							
c. Self-rated sexiness	0.14	0.07													Y										
Self-rated flirtatiousness	0.00	0.05																				Y			
WHR	-0.12	0.05					Y																		
Breast size (bust circumference)	-0.33	0.06																							
Other-rated facial attractiveness	Percentage high chosen over low: 43.64% (SD = 18.07%)									Y															
Bobst & Lobmaier (2012)			P	20 (within-participants)																					
Flirtatious-looking facial appearance	Percentage high chosen over low: 63.7% (SD = 4.1%)																								
Receptive-looking facial appearance	Percentage high chosen over low: 62.0% (SD = 3.8%)																								
Caring-looking facial appearance	Percentage high chosen over low: 63.1% (SD = 3.7%)																								
Other-rated facial attractiveness	Percentage high chosen over low: 59.0% (SD = 5.0%)									Y															
Bryant & Haselton (2009)			P	69 (within-participants)																					
Average vocal pitch across a, b, c	0.03	0.002																							
a. Vocal pitch - sentence	0.18	0.003															Y				Y				
b. Vocal pitch - monopthong vowels	0.02	0.01					Y																		
c. Vocal pitch - sustained vowels	-0.13	0.005					Y																		
Vocal pitch variability - spoken sentence	0.08	0.01																							
Noise to harmonics ratio - sustained vowels	0.11	0.02																							
Duration - spoken sentence	0.07	0.01																							
Formant Dispersion - monopthong vowels	-0.07	0.01																							
Jitter - sustained vowels	0.05	0.03																							
Shimmer - sustained vowels	0.08	0.02																							
Harmonics to noise ratio - sustained vowels	-0.08	0.01																							
Cantu et al. (2013)			P	25 (within-participants)																					
Average across a, b, c, d	0.07	0.02															Y					Y			
a. Verbal flirting with "Cads"	0.34	0.03																							
b. Verbal flirting with "Dads"	-0.12	0.03																							
c. Nonverbal flirting with "Cads"	0.21	0.03																							
d. Nonverbal flirting with "Dads"	-0.15	0.02																							
Cobey, Buunk, Pollet, Klipping, & Roberts (2013)			P	14 (within-participants)																					
Self-rated attractiveness	-0.14	0.03										Y													
Partner-rated attractiveness	0.45	0.04				Y	Y																		
Doty, Ford, Preti, & Huggins (1975)			P	4 (within-participants)																					
Vaginal odor pleasantness	0.11	0.09				Y	Y	Y																	
Vaginal odor intensity	-0.13	0.09							Y		Y														
Durante unpublished dissertation - Study 1			U																						
Composite of self-rated face attractiveness, body attractiveness, & sexiness	0.10	0.004		49 (within-participants)								Y													
Preference for sexy clothes/accessories on shopping task	0.11	0.01		51 (within-participants)													Y	Y							
Durante unpublished dissertation - Study 2			U																						
Combined effect across a, b, c	-0.19	0.04		57 (within-participants)													Y	Y							
a. Preference for sexy clothes/accessories on shopping task (primed with attractive women)	0.49	0.12		18 (within-participants)																					
b. Preference for sexy clothes/accessories on shopping task (primed with average-looking women)	-0.45	0.07		26 (within-participants)																					
c. Preference for sexy clothes/accessories on shopping task (viewed landscapes)	-0.73	0.15		13 (within-participants)																					

Table 14 *Continued*

						Analyses																				
				Women		Other-rated Attractiveness and Related Features						Attractiveness and Related Features					Attractiveness-enhancing Behavior							Men's Responses		
Study and Effect(s)	Hedges's *g*	Var	Pub	High-fertility *n*	Low-fertility *n*	Overall Attr.	Overall Attr. with Correlates	Scent Attr.	Scent Intensity	Facial Attr.	Voice Pitch (vowels)	Overall Attr.	Facial Attr.	Body Attr.	Sexiness	ST Mate Value	Overall Attr. Enhancement	Hypothetical Clothing Task	Walk Attr.	Sexy Clothing	Voice Pitch (social)	Flirtatiousness		Overall Men's Responses	Men's T Response	Partner Jealousy
Durante unpublished dissertation - Study 3			U																							
Combined effect across a, b, c	0.11	0.01		52 (within-participants)													Y	Y								
a. Preference for sexy clothes/accessories on shopping task (primed with attractive men)	0.15	0.02		22 (within-participants)																						
b. Preference for sexy clothes/accessories on shopping task (primed with average-looking men)	0.06	0.02		19 (within-participants)																						
c. Preference for sexy clothes/accessories on shopping task (viewed landscapes)	0.09	0.002		11 (within-participants)																						
Durante, Griskevicius, Hill, Perilloux, & Li (2010) - Study 1			P	60 (within-participants)																						
Preference for sexy clothes/accessories on shopping task	0.38	0.02															Y	Y								
Durante, Griskevicius, Hill, Perilloux, & Li (2010) - Study 2			P																							
Combined effect across a, b, c, d	0.03	0.02		48 (within-participants)													Y	Y								
a. Preference for sexy clothes/accessories on shopping task (primed with attractive women)	0.92	0.21		11 (within-participants)																						
b. Preference for sexy clothes/accessories on shopping task (primed with average-looking women)	-0.36	0.20		13 (within-participants)																						
c. Preference for sexy clothes/accessories on shopping task (primed with attractive men)	-0.22	0.10		11 (within-participants)																						
d. Preference for sexy clothes/accessories on shopping task (primed with average-looking men)	0.04	0.04		13 (within-participants)																						
Durante, Griskevicius, Hill, Perilloux, & Li (2010) - Study3			P																							
Combined effect across a, b, c, d	0.27	0.03		73	87												Y	Y								
a. Preference for sexy clothes/accessories on shopping task (primed with attractive local women)	1.13	0.11		19	22																					
b. Preference for sexy clothes/accessories on shopping task (primed with attractive distant women)	0.04	0.10		17	21																					
c. Preference for sexy clothes/accessories on shopping task (primed with average-looking local women)	0.06	0.10		17	24																					
d. Preference for sexy clothes/accessories on shopping task (primed with average-looking distant women)	-0.04	0.10		20	20																					
Durante, Li, & Haselton (2008)			P																							
Average across a, b	0.07	0.02		72 (within-participants)													Y									
Average across c, d, e, f	-0.03	0.003		86 (within-participants)								Y														
a. Average across 1, 2	0.01	0.02		72 (within-participants)																Y						
1. Other-rated sexiness of clothes worn to lab	0.11	0.02		72 (within-participants)																						
2. Other-rated revealingness of clothes worn to lab	-0.07	0.02		72 (within-participants)																						
b. Average across 3, 4, 5	0.05	0.02		88 (within-participants)														Y								
3. Other-rated sexiness of outfit drawing	0.18	0.01		88 (within-participants)																						
4. Other-rated revealingness of outfit drawing	0.12	0.01		88 (within-participants)																						
5. Measured skin exposure of outfit drawing	0.05	0.02		88 (within-participants)																						
c. Self-rated facial attractiveness	0.02	0.01		86 (within-participants)									Y													
d. Self-rated body attractiveness	-0.04	0.00		86 (within-participants)										Y												
e. Self-rated sexiness	0.02	0.00		86 (within-participants)											Y											
f. Self-rated short-term mate value	-0.09	0.01		86 (within-participants)												Y										
Self-rated long-term mate value	0.06	0.01		86 (within-participants)																						
Durante & Rae Arsena (unpublished raw data)			U																							
Composite of a, b, c, d	0.05	0.01		63 (within-participants)								Y														
a. Self-rated facial attractiveness	0.10	0.01		65 (within-participants)									Y													
b. Self-rated body attractiveness	-0.09	0.01		66 (within-participants)										Y												
c. Self-rated sexiness	0.07	0.01		64 (within-participants)											Y											
d. Self-rated short-term mate value	0.03	0.01		66 (within-participants)												Y										
Self-rated long-term mate value	-0.03	0.01		65 (within-participants)																						
Fink, Hugill, & Lange (2012)			P	48 (within-participants)																						
Other-rated dance attractiveness	0.32	0.02															Y									
Other-rated walk attractiveness	0.69	0.03															Y		Y							
Fischer et al. (2011) Samp 1			P	23 (within-participants)																						
Vocal pitch - sustained vowels	-0.01	0.03					Y																			
Maximum vocal pitch - sustained vowels	-0.09	0.01																								
Minimum vocal pitch - sustained vowels	-0.11	0.01																								
Degree of unvoiceness - sustained vowels	0.14	0.01																								
Noise to harmonic ratio - sustained vowels	0.00	0.04																								
Jitter - sustained vowels	0.19	0.01																								
Shimmer - sustained vowels	0.16	0.01																								
Other-rated vocal attractiveness - sustained vowel	M	M																								
Vocal pitch - free speech	0.13	0.02															Y				Y					
Vocal pitch variability - free speech	0.22	0.08																								
Other-rated vocal attractiveness - free speech (raters spoke a different language)	M	M																								
Flowe, Swords, & Rockey (2012)			P																							
Average across a, b	0.20	0.04		45	61												Y					Y				
a. Behavioral engagement with more masculine game show host	0.74	0.04		25	32																					
b. Behavioral engagement with less masculine game show host	-0.36	0.04		20	29																					

Table 14 *Continued*

				Women		Analyses																			
						Other-rated Attractiveness and Related Features						Attractiveness and Related Features					Attractiveness-enhancing Behavior						Men's Responses		
Study and Effect(s)	Hedges's *g*	Var	Pub	High-fertility *n*	Low-fertility *n*	Overall Attr.	Overall Attr. with Correlates	Scent Attr.	Scent Intensity	Facial Attr.	Voice Pitch (vowels)	Overall Attr.	Facial Attr.	Body Attr.	Sexiness	ST Mate Value	Overall Attr. Enhancement	Hypothetical Clothing Task	Walk Attr.	Sexy Clothing	Voice Pitch (social)	Flirtatiousness	Overall Men's Responses	Men's T Response	Partner Jealousy
Gangestad, Thornhill, & Garver (2002)			P	27 (within-women)																					
Total across a, b	0.23	0.01																					Y		
a. Women's reports of male partner's proprietariness	0.13	0.01																							Y
b. Women's reports of male partner's attentiveness	0.33	0.02																							
Gildersleeve et al. (2012)			P	41 (within-participants)																					
Average across a, b, c	0.30	0.02				Y	Y	Y																	
a. Axillary odor pleasantness	0.28	0.02																							
b. Axillary odor sexiness	0.29	0.03																							
c. Physical attractiveness inferred from axillary odor	0.31	0.03																							
Axillary odor intensity	-0.20	0.02							Y		Y														
Gonzales unpublished Master's thesis			U	10 (within-participants)																					
Red/pink clothing or accessories	0.74	0.11															Y								
Self-rated attractiveness	0.40	0.09										Y													
Gueguen (2009, *EHB*)			P	51	78																				
Receptivity to dance invitation at night club	0.45	0.04															Y					Y			
Gueguen (2009, *Biological Psychology*)			P	60	89																				
Receptivity to request for phone number on public street	0.64	0.08															Y					Y			
Gueguen (2012, *Gait & Posture*)			P	14	67																				
Time spent walking down the hall (slower walk) in front of attractive male confederate	1.02	0.09																							
Other-rated walk sexiness	1.15	0.09															Y		Y						
Gueguen (2012, *The Psychological Record*) - Study 1			P	14	50																				
Makeup-artist-rated attractiveness of cosmetics	0.62	0.09															Y								
Makeup-artist-rated amount of cosmetics	0.81	0.09															Y								
Self-reported time spent applying makeup	0.93	0.10															Y								
Gueguen (2012, *Psychological Record*) - Study 2			P	114	789																				
RA-rated amount of cosmetics	0.53	0.01															Y								
Haselton, Mortezaie, Pillsworth, Bleske-Recheck, & Frederick (2007)			P	30 (within-participants)																					
Other-rated "trying to look attractive" from body photos (clothes worn to lab)	Percentage high chosen over low: 60% (SD = 30%)																Y								
Other-rated "trying to look attractive" from facial photos	Percentage high chosen over low: 50% (SD = 25%)																Y								
Haselton & Gangestad (2006)			P																						
Self-rated attractiveness	0.52	0.04		38 (within-participants)								Y													
Self-reported flirtation	0.32	0.05		38 (within-participants)													Y					Y			
Women's reports of desire to go out to meet men	0.46	0.03		38 (within-participants)																					
Women's reports of male partner's jealousy and possessiveness	0.66	0.11		23 (within-participants)																			Y		Y
Havlicek, Dvorakova, Bartos, & Flegr (2006)			P	12 (within-participants)																					
Average across a, b	0.09	0.05				Y	Y																		
a. Axillary odor pleasantness	0.09	0.05																							
b. Axillary odor sexual attractiveness	0.04	0.05						Y																	
Axillary odor femininity	0.19	0.09																							
Axillary odor intensity	0.05	0.05							Y		Y														
Hill & Durante (2009) - Study 1			P	52 (within-participants)																					
Self-esteem	-0.29	0.004																							
Self-rated body, face, and overall sexiness	0.11	0.004										Y			Y										
Hill & Durante (2009) - Study 2			P	59 (within-participants)																					
Composite of self-rated body, face, and overall sexiness	0.08	0.003										Y			Y										
Self-esteem	-0.13	0.002																							
Preference for sexy clothes/accessories/products on shopping task	-0.07	0.01															Y	Y							
Hromatko, Tadinac, & Vranic (2008) - Study 1			P	14	56																				
Self-rated "femininity *over last couple days*" (single item adapted from BSRI)	0.17	0.09																							
Hromatko, Tadinac, & Vranic (2008) - Study 2			P	12	10																				
Self-rated femininity across all items from BSRI	0.36	0.17																							
Hromatko, Tadinac, & Vranic (2008) - Study 3			P	11	52																				
Self-rated "femininity" (single item from BSRI)	-0.53	0.11																							
Self-rated long-term mate value	0.15	0.11																							
Self-perceived Attractive Face (MV-7 Item 1)	-0.18	0.11										Y	Y												
Self-perceived Attractive Body (MV-7 Item 2)	-0.24	0.11										Y		Y											
Self-perceived Ambition (MV-7 Item 3)	-0.25	0.11																							
Self-perceived Desires Children (MV-7 Item 4)	0.22	0.11																							
Self-perceived Enthusiastic about Sex (MV-7 Item 5)	-0.12	0.11																							
Self-perceived Faithful (MV-7 Item 6)	0.57	0.11																							
Self-perceived Financially Secure (MV-7 Item 7)	0.14	0.11																							
Self-perceived "Generous"	0.18	0.11																							
Self-perceived "Good sense of humor"	-0.25	0.11																							
Self-perceived "Healthy"	0.12	0.11																							
Self-perceived "Independent"	-0.24	0.11																							
Self-perceived "Intelligent"	-0.07	0.11																							
Self-perceived "Kind and Understanding"	0.58	0.11																							
Self-perceived "Loyal"	0.60	0.11																							
Self-perceived "Responsible"	0.28	0.11																							
Self-perceived "Sociable"	-0.75	0.11																							
Self-perceived "Emotionally stable"	0.34	0.11																							

Table 14 *Continued*

						Analyses																			
				Women		Other-rated Attractiveness and Related Features						Attractiveness and Related Features					Attractiveness-enhancing Behavior						Men's Responses		
Study and Effect(s)	Hedge s's *g*	Var	Pub	High-fertility *n*	Low-fertility *n*	Overall Attr.	Overall Attr. with Correlates	Scent Attr.	Scent Intensity	Facial Attr.	Voice Pitch (vowels)	Overall Attr.	Facial Attr.	Body Attr.	Sexiness	ST Mate Value	Overall Attr. Enhancement	Hypothetical Clothing Task	Walk Attr.	Sexy Clothing	Voice Pitch (social)	Flirtatiousness	Overall Men's Responses	Men's T Response	Partner Jealousy
Kirchengast & Gartner (2002)			P	M	M																				
Waist-to-hip ratio	M	M																							
Kuukasjarvi et al. (2004)			P	40 total																					
T-shirt odor sexual attractiveness	0.41	0.11				Y	Y	Y																	
T-shirt odor intensity	-0.13	0.10							Y		Y														
Maner & McNulty (2013)			P	4 (within-participants)																					
Women's testosterone response to other women's T-shirt odors (same T-shirt donors as Miller & Maner 2010a, Study 1; 2010b, Study 2)	M	M																							
Manning, Scutt, Whitehouse, Leinster, & Walton (1996) - Sample 1			P	33 total																					
Mean fluctuating asymmetry across three traits	M	M																							
Manning et al. (1996) - Sample 2			P	15 (within-participants)																					
Mean fluctuating asymmetry across four traits	M	M																							
Manning et al. (1996) - Sample 3			P	7 (within-participants)																					
Mean fluctuating asymmetry across four traits	M	M																							
Manning et al. (1996) - Sample 4			P	3 (within-participants)																					
Mean fluctuating asymmetry across four traits	M	M																							
Mikac (unpublished manuscript)			U	19 (within-participants)																					
Skin exposure on female figure dressed by participants	0.46	0.05															Y	Y							
Clothing width (sign reversed to measure of clothing tightness) on female figure dressed by participants	0.27	0.07															Y	Y							
Miller, Tybur, & Jordan (2007)				11 (within-participants)																					
Lap-dance tip earnings	0.69	0.03	P																				Y		
Miller & Maner (2010a) - Study 1				4 (within-participants)																					
Men's testosterone response to women's T-shirt odors	2.04	0.96																					Y	Y	
Miller & Maner (2010a) - Study 2			P																						
Men's testosterone response to women's T-shirt odors	0.53	0.27		9 (within-participants)																			Y	Y	
T-shirt odor pleasantness	0.09	0.20		10 (within-participants)		Y	Y	Y																	
T-shirt odor intensity	0.12	0.15		10 (within-participants)					Y		Y														
Miller & Maner (2010b) - Study 1			P	10 (within-participants)																					
Men's implicit cognitive access to sexual concepts in response to women's T-shirt odors (same T-shirt donors as 2010a, Study 2)	M	M																							
Miller & Maner (2010b) - Study 2			P	4 (within-participants)																					
Men's perception of sexual arousal on behalf of Women whose T-shirts they smelled (same T-shirt donors as 2010a, Study 1)	0.36	0.50																							
Miller & Maner (2010, *JPSP*) - Study 3			P	1																					
Men's behavioral mimicry of woman confederate	M	M																							
Men's risky decisions in blackjack in front of woman confederate	M	M																							
Men's ratings of woman confederate's flirtatiousness	M	M																							
Men's ratings of woman confederate's outgoingness	M	M																							
Men's ratings of woman confederate's attractiveness	M	M																							
Pillsworth & Haselton (2006)				43 (within-participants)																					
Women's reports of male partner's love/attention	M	M																							
Women's reports of male partner's jealousy/possessiveness	M	M																							
Pipitone & Gallup (2008)			P	17 (within-participants)																					
Other-rated voice attractiveness	M	M																							
Poran (1994)																									
Scent attractiveness	M	M																							
Provost, Quinsey, & Troje (2008)			P	19 (within-participants)																					
Knee movement	M	M																							
Ankle movement	M	M																							
Hip movement	M	M																							
Other-rated walk attractiveness	M	M																							
Puts et al. (2013)			P	M																					
Female-rated facial photo flirtatiousness	M	M																							
Female-rated facial photo attractiveness to men	M	M																							
Male-rated facial photo short-term attractiveness	M	M																							
Male-rated facial photo long-term attractiveness	M	M																							
Female-rated voice flirtatiousness	M	M																							
Female-rated voice attractiveness to men	M	M																							
Male-rated voice short-term attractiveness	M	M																							
Male-rated voice long-term attractiveness	M	M																							
Vocal pitch	M	M																							
Jitter	M	M																							
Shimmer	M	M																							
Duration	M	M																							
Formant dispersion	M	M																							
Harmonics	M	M																							
Number of voice breaks	M	M																							
Roberts et al. (2004) - Newcastle & Prague samples			P	48 (within-participants)																					
Male-rated facial attractiveness	Percentage high chosen over low: 53.9% (SD = 16%)									Y															
Female-rated facial attractiveness	Percentage high chosen over low: 53.6% (SD = 18%)																								

Table 14 *Continued*

				Women		Analyses																				
						Other-rated Attractiveness and Related Features						Attractiveness and Related Features					Attractiveness-enhancing Behavior							Men's Responses		
Study and Effect(s)	Hedge s's *g*	Var	Pub	High-fertility *n*	Low-fertility *n*	Overall Attr.	Overall Attr. with Correlates	Scent Attr.	Scent Intensity	Facial Attr.	Voice Pitch (vowels)	Overall Attr.	Facial Attr.	Body Attr.	Sexiness	ST Mate Value	Overall Attr. Enhancement	Hypothetical Clothing Task	Walk Attr.	Sexy Clothing	Voice Pitch (social)	Flirtatiousness		Overall Men's Responses	Men's T Response	Partner Jealousy
Röder, Brewer, & Fink (2009)			P	25 (within-participants)																						
"Interest in shopping" original composite from paper: included b, d	0.46	0.01																								
"Interest in styling" original composite from paper: included a, c, e, g, h, i	1.02	0.02																								
" Attractiveness" original composite from paper: included j, k, l, f	1.90	0.07																								
Combined effect across a, b, c, d, e, f, g, h, i	0.16	0.01															Y									
a. Self-reported time spent styling hair	0.30	0.02																								
b. Self-reported desire to go to hairdresser	0.11	0.02																								
c. Self-reported fashionableness of clothes	0.28	0.03																								
d. Self-reported desire to shop	0.02	0.01																								
e. Self-reported time spent on hygiene	0.03	0.02																								
f. Self-reported diet and exercise to enhance appearance	0.29	0.02																								
g. Self-reported time spent deciding what to wear/getting dressed	0.09	0.02																								
h. Self-reported sexiness of clothes worn that day	0.30	0.02																		Y						
i. Self-reported complexity of makeup	0.06	0.03																								
Combined effect across j, k, l	0.29	0.01										Y														
j. Self-rated physical attractiveness (face and body)	0.06	0.02																								
Combined effect across k, l	0.40	0.02													Y											
k. Self-rated sexiness	0.32	0.05																								
l. Self-rated sexiness to men	0.48	0.04																								
Self-reported flirtatiousness	0.30	0.03																				Y				
Self-reported desire to change appearance	-0.02	0.01																								
Saad & Stenstrom (2012)			P	17 (within-participants)																						
Combined effect across a, b, c, d,e	0.53	0.03															Y									
a. Self-reported wearing clothes that attract attention	0.74	0.06																								
b. Self-reported wearing nice clothes	0.29	0.05																								
c. Self-reported time spent "making myself beautiful"	0.41	0.04																								
d. Self-reported wearing sexy clothes that day	0.55	0.06																		Y	Y					
e. Self-report wore clothes that day that show lots of skin	0.67	0.08																								
Self-reported dollars spent on clothing	0.25	0.12																								
Self-reported wearing a skirt	0.43	0.06																								
Self-reported wore makeup	-0.23	0.10																								
Self-reported sun tanning	0.58	0.10																								
Samson et al. (2011)			P	9 (within-participants)																						
Other-rated facial attractiveness	M	M																								
Facial coloration	M	M																								
Schwarz & Hassebrauck (2008)			P																							
Average of a, b	0.25	0.01		37 (within-participants)													Y			Y						
a. Self-rated provocativeness of clothes worn that day	0.30	0.01		37 (within-participants)																						
Self-rated attractiveness	0.20	0.02		40 (within-participants)								Y														
Other-rated facial attractiveness	0.13	0.004		37 (within-participants)		Y	Y																			
b. Other-rated provocativeness of clothes worn that day	0.20	0.01		37 (within-participants)																						
Other-rated skin exposure of clothes worn that day	0.30	0.02		37 (within-participants(																						
Singh & Bronstad (2001) - Primary Study			P																							
T-shirt odor "attractiveness" (average of a, b)	0.35	0.03		17 (within-participants)		Y	Y	Y																		
a. T-shirt odor pleasantness	0.09	0.26																								
b. T-shirt odor sexiness	0.07	0.26																								
T-shirt odor intensity	-0.01	0.26							Y		Y															
Singh & Bronstad (2001) - Mini-replication Study			P	4 (within-participants)																						
T-shirt odor "attractiveness" (average of pleasantness and sexiness)	0.47	0.07				Y	Y	Y																		
Thornhill & Gangestad (1999)			P	55 (continuous fertility variable)																						
Average of T-shirt odor pleasantness and sexiness (subsample of women who claimed to use no fragrances)	-0.04	0.07				Y	Y	Y																		
T-shirt odor intensity (subsample of women who claimed to use no fragrances)	-0.32	0.08							Y		Y															
Trouton et al.			P	17 (within-participants)																						
Scent attractiveness	M	M																								

Note. M = Missing data; Y = included in that analysis.