Psychological Bulletin

0033-2909/14/$12.00 DOI: 10.1037/a0035438

Do Women's Mate Preferences Change Across the Ovulatory Cycle? A Meta-Analytic Review

AQ: au

Kelly Gildersleeve, Martie G. Haselton, and Melissa R. Fales
University of California, Los Angeles

Scientific interest in whether women experience changes across the ovulatory cycle in mating-related motivations, preferences, cognitions, and behaviors has surged in the past 2 decades. A prominent hypothesis in this area, the ovulatory shift hypothesis, posits that women experience elevated immediate sexual attraction on high- relative to low-fertility days of the cycle to men with characteristics that reflected genetic quality ancestrally. Dozens of published studies have aimed to test this hypothesis, with some reporting null effects. We conducted a meta-analysis to quantitatively evaluate support for the pattern of cycle shifts predicted by the ovulatory shift hypothesis in a total sample of 134 effects from 38 published and 12 unpublished studies. Consistent with the hypothesis, analyses revealed robust cycle shifts that were specific to women's preferences for hypothesized cues of (ancestral) genetic quality (96 effects in 50 studies). Cycle shifts were present when women evaluated men's "short-term" attractiveness and absent when women evaluated men's "long-term" attractiveness. More focused analyses identified specific characteristics for which cycle shifts were or were not robust and revealed areas in need of more research. Finally, we used several methods to assess potential bias due to an underrepresentation of small effects in the meta-analysis sample or to "researcher degrees of freedom" in definitions of high- and low-fertility cycle phases. Neither type of bias appeared to account for the observed cycle shifts. The existence of robust relationship context-dependent cycle shifts in women's mate preferences has implications for understanding the role of evolved psychological mechanisms and the ovulatory cycle in women's attractions and social behavior.

Keywords: mate preferences, ovulation, menstrual cycle, evolution, masculinity

Supplemental materials: http://dx.doi.org/10.1037/a0035438.supp

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 2 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

---

Kelly Gildersleeve, Department of Psychology and Center for Behavior, Evolution, and Culture, University of California, Los Angeles; Martie G. Haselton, Departments of Psychology and Communication Studies, Institute for Society and Genetics, and Center for Behavior, Evolution, and Culture, University of California, Los Angeles; Melissa R. Fales, Department of Psychology and Center for Behavior, Evolution, and Culture, University of California, Los Angeles.

We gratefully acknowledge the University of California, Los Angeles (UCLA), for awarding Kelly Gildersleeve the Graduate Research Mentorship and Dissertation Year Fellowship to support her throughout her completion of this project and for awarding Martie G. Haselton an Academic Senate Grant to support this work. We also thank the National Science Foundation Graduate Research Fellowship Program for providing Melissa R. Fales with fellowship support while working on this project. For this work, Gildersleeve received the Harold H. Kelley Award for best student research paper in social psychology at UCLA and the New Investigator Award (2013) from the Human Behavior and Evolution Society. We enthusiastically thank the many researchers who generously shared data with us or conducted additional analyses to make it possible for us to include their studies in this meta-analysis. We also thank Steve Gangestad, Britt Ahlstrom, Jennifer Hahn-Holbrook, Christina Larson, Aaron Lukaszewski, Damian Murray, David Pinsof, and Shimon Sapphire-Bernstein for their constructive feedback on earlier versions of this article.

Correspondence concerning this article should be addressed to Kelly Gildersleeve, Department of Psychology, University of California, 5437B Franz Hall, Los Angeles, CA 90095. E-mail: kellygildersleeve@gmail.com

(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 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; Pieta, 2008; Stumpf & Boesch, 2005; 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 females 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). Fn1

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

---

[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.

Prediction 2

The second prediction of the ovulatory shift hypothesis is that the proposed cycle shift in women's attraction to characteristics 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 unlikely 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 immediate 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, 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 coparent 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 benefited 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 into 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 females who entered into 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 nongenetic reproductive benefits (e.g., material investment or protection) in the remainder of the cycle might have had greater reproductive success, on average, than 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 (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 pair bonds with men who were both high in genetic quality and highly suitable as a long-term social partner and coparent. 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 pair bonds with men who were relatively high in suitability as long-term partners but relatively low in genetic quality, women who maintained their primary pair bond 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.

Fn2

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" (Clarke, 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 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; see also Carter, Weier, & Houle, 2009, for experimental evidence) and to negative health outcomes in humans (see Thornhill & Møller, 1997; Van Dongen & Gangestad, 2011). Fn3

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 less symmetrical individuals (meta-analyzed by Møller & 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

---

[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 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.

[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.

age than 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 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 Fn4 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 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 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, Hanson, 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 that, 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.

Method

Search Strategy

As shown in Tables 1 and 2, we identified a large number of T1-2, AQ:1 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 (2005–2012) and of the Human Behavior and Evolution Society (2000–2012) to identify researchers who had given talks or presented posters on research related to mating and the ovulatory cycle. We e-mailed 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. Last, we e-mailed 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*"

---

[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 & Møller, 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).

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*)

Study and effects	Relationship context	*N*	Sample size: High-fertility *n*	Sample size: Low-fertility *n*	Effect Size (*g*)	Variance	Reason for exclusion	Inclusion in analyses: Long-term partner quality: Broad set of measures	Inclusion in analyses: Long-term partner quality: Narrow set of measures	Inclusion in analyses: Genetic quality: Broad set of measures	Genetic quality: Narrow set of measures	Genetic quality: Facial symmetry	Genetic quality: Scent cues of symmetry	Genetic quality: Facial masculinity	Genetic quality: Body masculinity	Genetic quality: Vocal masculinity	Genetic quality: Behavioral dominance	Genetic quality: 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*)

								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 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*)

Study and effects	Relationship context	*N*	Sample size		Effect Size (*g*)	Variance	Reason for exclusion	Inclusion in analyses										
								Long-term partner quality		Genetic quality								
			High-fertility *n*	Low-fertility *n*				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: 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
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		Cycle position estimation method	Stimuli	Method to determine amount of male trait	Stated versus revealed preferences	Task	Number of trials
						High fertility	Low fertility						
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: 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
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		Cycle position estimation method	Stimuli	Method to determine amount of male trait	Stated versus revealed preferences	Task	Number of trials
						High fertility	Low fertility						
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: 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
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 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
Thornhill & Gangestad (1999b)													
Scent cues of body symmetry	U	United States	Col	L	B	Continuous		Forw & Rev	SS	Meas	Revealed	Ratings	78
Thornhill et al. (2013)													
Scent cues of testosterone	U	United States	Col	L	B	Continuous		Forw & Rev	SS	Meas	Revealed	Ratings	46
Scent cues of cortisol	U	United States	Col	L	B	Continuous		Forw & Rev	SS	Meas	Revealed	Ratings	46
Thornhill et al. (2003)													
Scent cues of body symmetry	U	United States	Col	L	B	Continuous		Forw & Rev	SS	Meas	Revealed	Ratings	56
Vaughn et al. (2010)													
Facial masculinity	U	United States	Col	L	B	0.08	0.01	Forw	FP	Manip	Revealed	Slider	1
Welling et al. (2007)													
Facial masculinity	U	M	M	L	W	0.09	0.01	Forw	FP	Manip	Revealed	TOFC	20

Note. ST = short term; LT = long term; U = unspecified; Col = college/university students; Com = community; L = lab; F = field (online unless specified as “Mag” [magazine survey]); W = within participants; B = between participants; Forw = forward counting method; Rev = reverse counting method; Forw & Rev = average from forward and reverse counting methods; VM = verified benchmark date of menstrual onset; Saliv = salivary ferning method to verify ovulation; LH = luteinizing hormone tests to verify ovulation; SR = self-reported preference; FP = facial photos; BP = body photos; BP & FP = body photos and facial photos (ratings averaged); VR = vocal recordings, VB = videotaped behavior; SS = scent samples; FAv = face avatars; BAv = full body avatars; MFO = moving facial outlines; BOD = body outline drawings; D = descriptions of hypothetical men; PLW = point-light walker; Manip = manipulated; Meas = measured; SSR = ratings by a separate sample of participants; N/A = not applicable; PR = ratings by cycle study participants; Ratings = rated stimuli on a scale; TOFC = two-option forced choice; MOFC = multiple-option (3+) forced choice; Mate dollars = allocated “mate dollars” from a fixed budget to purchase more of a characteristic in a hypothetical mate; M = missing data; CPI = California Psychological Inventory.

[a] Degraded to conceal identity of stimulus men.

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, 2008; Gangestad, Thornhill, & Garver-Apgar, 2005a; Jones et al., 2008). We discontinued our literature search in December 2012.[5] Fn5

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] Fn6

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 cycle (e.g., date of last menstrual onset; see supplemental materials 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 on the basis of 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., Caryl et al., 2009; Miller, 2003; Prokosch, Coss, Scheib, & Blozis, 2009); 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 (see, e.g., Miller, 2000; Prokosch, 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. However, 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

---

[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 who 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.

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, AQ: 2 Pillsworth, & Haselton, 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 Fn7 (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 an article or poster did not report the needed information, we contacted study authors to request this information. If 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 trade-offs, 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).

---

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

Excluding this characteristic had a negligible impact on the mean weighted 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 long-term partner quality in ancestral males: relationship skills, parenting skills, nurturance, sympathy, warmth, kindness, trustworthiness, faithfulness (including likelihood of being unfaithful, reverse-coded), 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, chest 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 argued 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 (see, e.g., 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 on the basis of their ratings (see supplemental materials 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). 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 3 months (Nassar-alla 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). Last, 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] Fn8 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 instead to categorize women who participated on forward cycle days 9–15 as high fertility and women who participated on forward cycle days 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.

---

[8] One study (Harris, 2011) 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 et al. (2001).

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 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 supplemental materials 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 article 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, nonindependent 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 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 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 examining cycle shifts in women's preferences for all hypothesized cues of ancestral genetic quality revealed robust cycle shifts. These analyses included large but heterogeneous samples 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 vs. 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 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.

F1

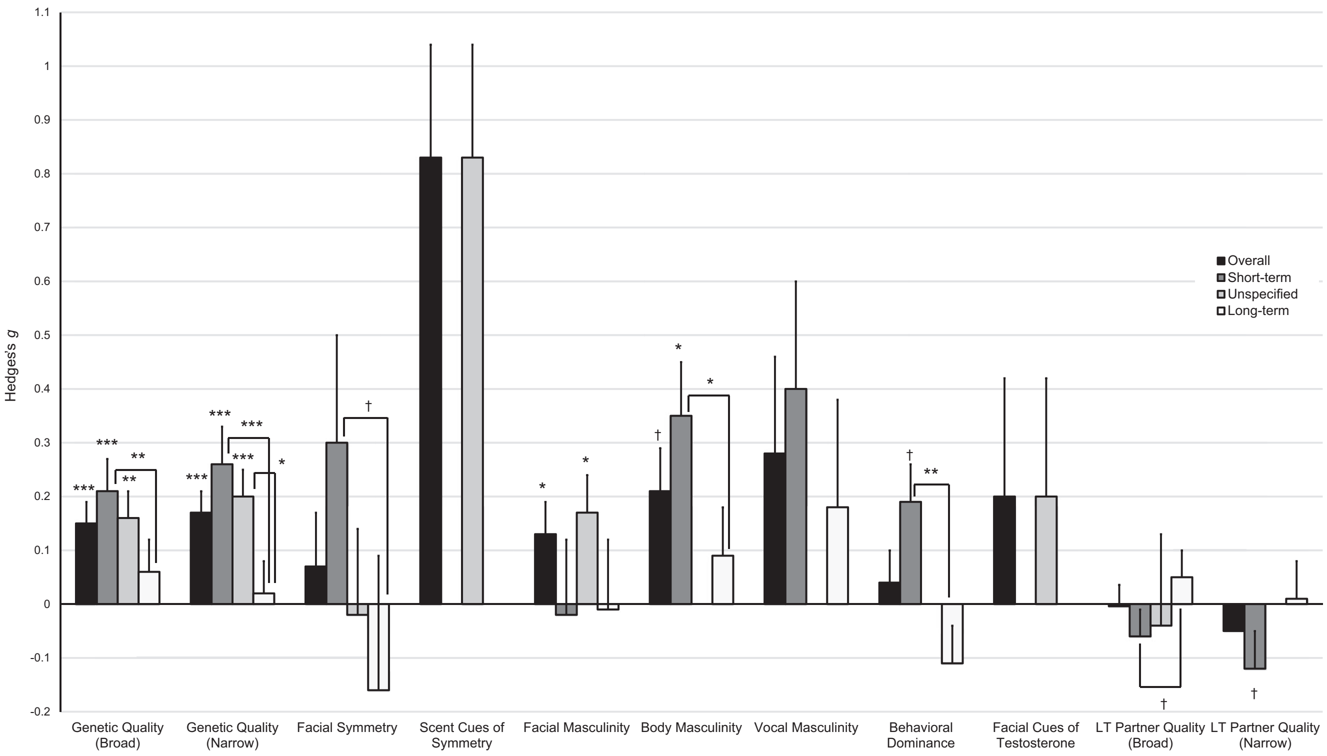
Step 1. In each sample of effects, we first specified an unconditional random-effects model to compute the weighted mean *g* as an estimate of the "true" (population) mean standardized mean difference between high and low fertility in women's preference for a characteristic across all relationship contexts:

$$\text{Level 1 model (effects): } g_{ij} = \delta_{0j} + e_{ij}.$$

$$\text{Level 2 model (studies): } \delta_{0j} = \gamma_{00} + u_{0j}.$$

In the model above, g_{ij} is the observed standardized mean difference i for study j, δ_{0j} is the "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



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

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.

AQ: 4

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 the weighted mean *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 3–13.

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

$$\text{Level 2 model (studies): } \delta_{0j} = \gamma_{00} + u_{0j}$$

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

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

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 analyses 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 vs. unspecified and short-term vs. 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 estimate 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 by 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 attraction to characteristics that reliably indicated genetic quality in ancestral males. Specifically, the ovulatory shift hypothesis predicts that women's attraction to these characteristics is 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., Li & Kenrick, 2006; Regan, 1998); 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 an overall 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.

Last, 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, 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, the following study characteristic was 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, rather than any other task to assess mate preferences ($p = .02$). In contrast, the following was 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 ($p = .04$).

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).

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 ($p = .001$); 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 significantly 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 using 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

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).

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 shift 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 ($g = 0.17$, $SE = 0.07$) but 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 ($p = .96$).

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 statis-

---

[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.

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.

tically 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 ($g = 0.28$, $SE = 0.18$), 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. T9

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. T10

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).

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 shift in women's preference for facial cues of circulating testosterone and to examine possible differences across relationship contexts.

Preference for All Hypothesized Cues of Ancestral Long-Term 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

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).

T12

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 Long-Term 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.

T13

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

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).

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

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).

right. At low precision, only large effects reach statistical significance. Therefore, the gap that forms in the lower left quadrant of the funnel suggests that small effects are missing, perhaps due to publication bias or some other sources of bias preventing the inclusion of nonsignificant effects.

We used funnel plots to assess whether it was likely that small effects were underrepresented in the sample of effects in our analysis 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 ancestral 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.

F2 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 (2002) "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

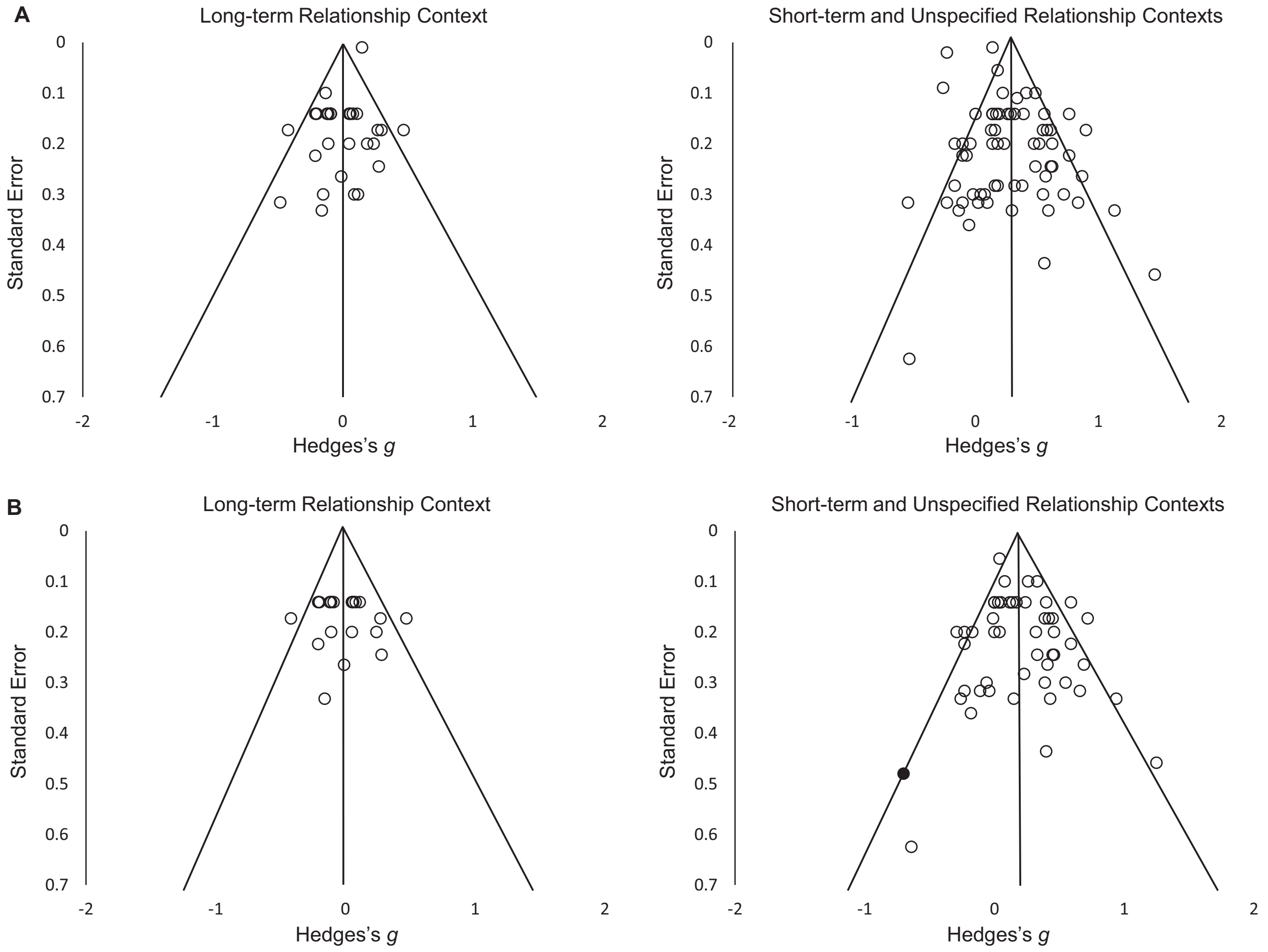


Figure 2. Funnel plots to examine evidence for an underrepresentation of small effects among the sample of effects for which the ovulatory shift hypothesis predicts relationship context-dependent cycle shifts—namely, effects assessing cycle shifts in preferences for characteristics hypothesized to have 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.

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 (thereby dramatically increasing the Type I 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- and 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 supplemental materials).

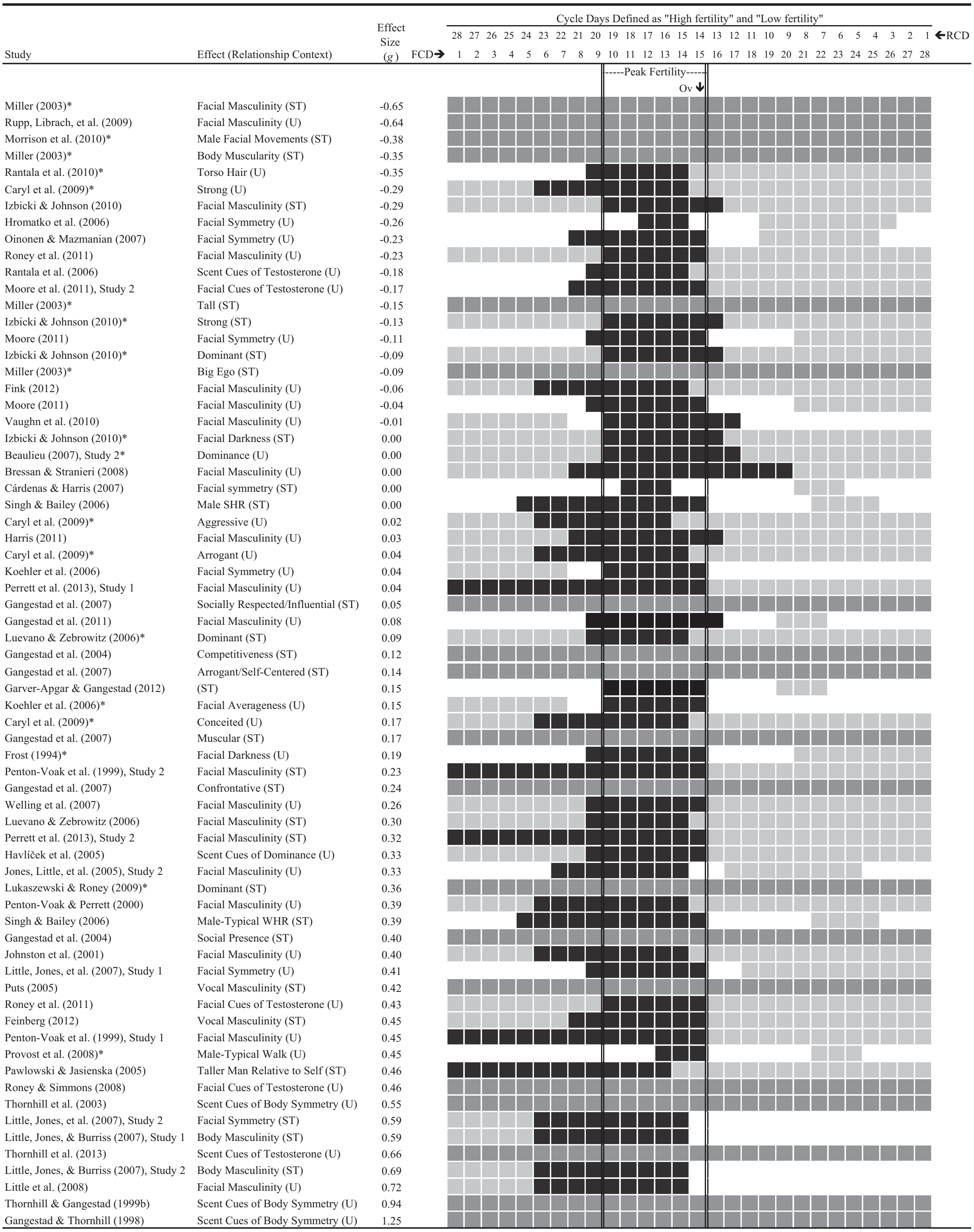
Given the uninformative nature of this null finding, we next attempted to address the issue by visually examining associations between effect size and high- and low-fertility window definitions. F3 Figure 3 presents the high- and low-fertility windows used to measure each effect that was predicted to 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 in this meta-analysis 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 Figure 3 does not reveal obvious evidence for 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 this meta-analysis 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 (though we cannot definitively rule out the possibility that researchers chose, post hoc, to use a continuous fertility variable because doing so produced predicted cycle shifts). 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.



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 et al. (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)	Socially Respected/Influential (ST)	0.05
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)	(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 et al. (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
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 indicate 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-

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 males. 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.

Figure 3 (opposite). and low-fertility windows are displayed in terms of forward cycle day (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.

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 3 times as many effects as 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 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 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, nonmutually 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 articles that report positive findings as more worthy of publication than articles 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 articles that report positive findings as more worthy of publication than articles 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. 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 supplemental materials), 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 all 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 nonindependence 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 article, 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 et al., 2013), preferences for attractive and revealing clothing (Durante, Li, & Haselton, 2008; Haselton, Mortezaie, Pillsworth, Bleske-Rechek, & Frederick, 2007), 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 women whose primary partners are relatively lacking in the characteristics women particularly prefer at high fertility—namely, characteristics thought to have reflected genetic quality in ancestral males—will be particularly likely to 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 males (e.g., partner sexual attractiveness, Pillsworth & Haselton, 2006a; partner sexual attractiveness relative to investment attractiveness, Haselton & Gangestad, 2006; partner facial masculinity, Gangestad, Thornhill, & Garver-Apgar, 2010b; facial masculinity and partner facial attractiveness [marginally significant], Gangestad et al., 2010b; composite partner face and body attractiveness, Larson et al., 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 during 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 pair bonding 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, women in traditional, "natural-fertility" populations experience far fewer ovulatory cycles than 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 Strassmann, 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 supplemental materials), 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 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 2 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 articles 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 mate 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.

References

References marked with an asterisk indicate studies included in the meta-analysis.

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

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, England: Wiley. doi:10.1002/9780470743386

Braude, S., Tang-Martinez, Z., & Taylor, G. T. (1999). Stress, testosterone, and the immunoredistribution hypothesis. *Behavioral Ecology, 10,* 345–350. doi:10.1093/beheco/10.3.345

*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. doi:10.1111/j.1467-9280.2008.02060.x

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

*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. doi:10.1016/j.evolhumbehav.2006.08.003

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. doi:10.1038/hdy.2009.13

*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. doi:10.1016/j.paid.2008.11.024

Chen, B. P., & Parham, P. (1989). Direct binding of influenza peptides to Class I HLA molecules. *Nature, 337,* 743–745. doi:10.1038/337743a0

Clarke, G. M. (1993). The genetic basis of developmental stability, I. Relationships between stability, heterozygosity, and genomic coadaptation. *Genetica, 89,* 15–23. doi:10.1007/BF02424502

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

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 (in press) "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. doi:10.1016/j.yhbeh.2004.11.006

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. doi:10.1177/0146167208323103

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. doi:10.1111/j.0006-341X.2000.00455.x

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,* 245–264. doi:10.1037/0022-3514.94.2.245

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. doi:10.1037/a0032432

Fehring, R. J., & Gaska, N. (1998). Evaluation of the Lady Free Biotester® in determining the fertile period. *Contraception, 57,* 325–328. doi:10.1016/S0010-7824(98)00039-0

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. doi:10.1111/j.1552-6909.2006.00051.x

*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., . . . Perrett, D. I. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behavior, 49,* 215–222. doi:10.1016/j.yhbeh.2005.07.004

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

Fiske, P., Rintamaki, P. T., & Karvonen, E. (1998). Mating success in lekking males: A meta-analysis. *Behavioral Ecology, 9,* 328–338. doi:10.1093/beheco/9.4.328

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. doi:10.1016/j.evolhumbehav.2011.10.006

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,* 507–514. doi:10.2466/pms.1994.79.1.507

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 close relationships* (pp. 374–399). New York, NY: Oxford University Press. doi:10.1093/oxfordhb/9780195398694.013.0017

*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 and Social Psychology, 92,* 151–163. doi:10.1037/0022-3514.92.1.151

Gangestad, S. W., & Simpson, J. A. (2000). The evolution of human mating: Trade-offs and strategic pluralism. *Behavioral and Brain Sciences, 23,* 573–587. doi:10.1017/S0140525X0000337X

*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. doi:10.1111/j.0956-7976.2004.01503010.x

*Gangestad, S. W., & Thornhill, R. (1998). Menstrual cycle variation in women's preferences for the scent of symmetrical men. *Proceedings of the Royal Society B: Biological Sciences, 265,* 927–933. doi:10.1098/rspb.1998.0380

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

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 B: Biological Sciences, 269,* 975–982. doi:10.1098/rspb.2001.1952

Gangestad, S. W., Thornhill, R., & Garver-Apgar, C. E. (2005a). Adaptations to ovulation: Implications for sexual and social behavior. *Current Directions in Psychological Science, 14,* 312–316. doi:10.1111/j.0963-7214.2005.00388.x

Gangestad, S. W., Thornhill, R., & Garver-Apgar, C. E. (2005b). Women's sexual interests across the ovulatory cycle depend on primary partner developmental instability. *Proceedings of the Royal Society B: Biological Sciences, 272,* 2023–2027. doi:10.1098/rspb.2005.3112

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. doi:10.1016/j.evolhumbehav.2010.05.003

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. doi:10.1016/j.evolhumbehav.2010.06.001

*Gangestad, S. W., Thornhill, R., & Garver-Apgar, C. E. (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. (2008). Hormonal correlates of women's midcycle preference for the scent of symmetry. *Evolution and Human Behavior, 29,* 223–232. doi:10.1016/j.evolhumbehav.2007.12.007

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. doi:10.1111/j.1467-9280.2006.01789.x

Geary, D. C. (2000). Evolution and proximate expression of human paternal investment. *Psychological Bulletin, 126,* 55–77. doi:10.1037/0033-2909.126.1.55

Guéguen, N. (2009a). Menstrual cycle phases and female receptivity to a courtship solicitation: An evaluation in a nightclub. *Evolution and Human Behavior, 30,* 351–355. doi:10.1016/j.evolhumbehav.2009.03.004

Guéguen, N. (2009b). The receptivity of women to courtship solicitation across the menstrual cycle: A field experiment. *Biological Psychology, 80,* 321–324. doi:10.1016/j.biopsycho.2008.11.004

Guermandi, E., Vegetti, W., Bianchi, M. M., Uglietti, A., Ragni, G., & Crosignani, P. (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 Workshop (STRAW) staging system. *Journal of Clinical Endocrinology & Metabolism, 92,* 3060–3067. doi:10.1210/jc.2007-0066

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

*Harris, C. R. (2011). Menstrual cycle and facial preferences reconsidered. *Sex Roles, 64,* 669–681. doi:10.1007/s11199-010-9772-8

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. doi:10.1016/j.yhbeh.2005.10.006

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. doi:10.1007/s12110-006-1020-0

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. doi:10.1016/j.yhbeh.2006.07.007

Havlíček, J., Roberts, S. C., & Flegr, J. (2005). Women's preference for dominant male odour: Effects of menstrual cycle and relationship status. *Biology Letters, 1,* 256–259. doi:10.1098/rsbl.2005.0332

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world. *Behavioral and Brain Sciences, 33,* 61–83. doi:10.1017/S0140525X0999152X

*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. doi:10.1007/s12110-010-9101-5

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

Hughes, A. L., & Nei, M. (1988). Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature, 335,* 167–170. doi:10.1038/335167a0

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 of the United States of America, 86,* 958–962.

*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, 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. doi:10.1016/S1090-5138(01)00066-6

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. doi:10.1007/s10508-007-9268-y

*Jones, B. C., Little, A. C., Boothroyd, L., DeBruine, L. M., Feinberg, D. R., Law Smith, M. J., . . . Perrett, D. I. (2005). 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. doi:10.1016/j.yhbeh.2005.03.010

Jones, B. C., Perrett, D. I., Little, A. C., Boothroyd, L., Cornwell, R. E., Feinberg, D. R., & Moore, F. R. (2005). Menstrual cycle, pregnancy and oral contraceptive use alter attraction to apparent health in faces. *Proceedings of the Royal Society B: Biological Sciences, 272,* 347–354. doi:10.1098/rspb.2004.2962

Keddy-Hector, A. C. (1992). Mate choice in non-human primates. *American Zoologist, 32,* 62–70. doi:10.1093/icb/32.1.62

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

Knott, C. D., Thompson, E. M., & Stumpf, R. M. (2007). Sexual coercion and mating strategies of wild Bornean orangutans [Abstract]. *American Journal of Physical Anthropology, 44*(Suppl.), 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. doi:10.1006/anbe.2002.3063

*Koehler, N., Rhodes, G., & Simmons, L. W. (2006). Do cyclic changes in women's face preferences target cues to long-term health? *Social Cognition, 24,* 641–656. doi:10.1521/soco.2006.24.5.641

Laeng, B., & Falkenberg, L. (2007). Women's pupillary responses to sexually significant others during the hormone cycle. *Hormones and Behavior, 52,* 520–530. doi:10.1016/j.yhbeh.2007.07.013

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

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. doi:10.1016/j.yhbeh.2012.10.005

Larson, C. M., Pillsworth, E. G., & Haselton, M. G. (2012). Ovulatory shifts in women's attractions to primary partners and other men: Further evidence of the importance of primary partner sexual attractiveness. *PLoS ONE, 7,* e44456. doi:10.1371/journal.pone.0044456

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. doi:10.1037/0022-3514.82.6.947

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. doi:10.1037/0022-3514.90.3.468

Li, N. P., Pillsworth, E. G., & Haselton, M. G. (2006). [Cycle shifts in women's mate preferences assessed using a mate dollars paradigm]. Unpublished raw data. AQ: 3

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. doi:10.1093/beheco/arr061

*Little, A. C., Jones, B. C., & Burriss, R. P. (2007). Preferences for masculinity in male bodies change across the ovulatory cycle. *Hormones and Behavior, 51,* 633–639. doi:10.1016/j.yhbeh.2007.03.006

*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. doi:10.1016/j.biopsycho.2007.08.003

*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. doi:10.1016/j.paid.2008.05.024

*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, 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. doi:10.1016/j.paid.2009.02.019

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, Williamsburg, VA.

*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.

Miller, G. F. (2000). Sexual selection for indicators of intelligence. In G. R. Bock, J. A. Goode, & K. Webb (Eds.), *The nature of intelligence* (pp. 260–275). New York, NY: Oxford University Press. doi:10.1002/0470870850.ch16

*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, Lincoln, NE.

Møller, A. P., & Thornhill, R. (1998). Bilateral symmetry and sexual selection: A meta-analysis. *American Naturalist, 151,* 174–192. doi:10.1086/286110

*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: Biological Sciences, 278,* 774–780. doi:10.1098/rspb.2010.1678

*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. doi:10.1007/s10508-009-9527-1

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. doi:10.1089/jwh.2010.2001

Navarrete, C. D., Fessler, D. M. T., Santos Fleischman, D., & Geyer, J. (2009). Race bias tracks conception risk across the menstrual cycle. *Psychological Science, 20,* 661–665. doi:10.1111/j.1467-9280.2009.02352.x

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

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). Hauppauge, NY: Nova Science.

*Oinonen, K. A., & Mazmanian, D. (2007). Facial symmetry detection ability changes across the ovulatory cycle. *Biological Psychology, 75,* 136–145. doi:10.1016/j.biopsycho.2007.01.003

*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. doi:10.1016/j.biopsycho.2005.02.002

Penn, D. J., Damjanovich, K., & Potts, W. K. (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences of the United States of America, 99,* 11260–11264. doi:10.1073/pnas.162006499

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

*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. doi:10.1038/21557

*Perrett, D. I., Jones, B. C., Little, A. C., Boothroyd, L., Cornwell, R. E., Feinberg, D. R., . . . 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. doi:10.1111/j.1420-9101.2007.01477.x

Peters, M., Simmons, L. W., & Rhodes, G. (2009). Preferences across the menstrual cycle for masculinity and symmetry in photographs of male faces and bodies. *PLoS ONE, 4,* e4138. doi:10.1371/journal.pone.0004138

Pieta, K. (2008). Female mate preferences among *Pan troglodytes schweinfurthii* of Kanyawara, Kibale National Park, Uganda. *International Journal of Primatology, 29,* 845–864. doi:10.1007/s10764-008-9282-5

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. doi:10.1016/j.evolhumbehav.2005.10.002

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. doi:10.1080/10532528.2006.10559837

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. doi:10.1016/j.evolhumbehav.2008.07.004

Provost, M. P., Troje, N. F., & Quinsey, V. L. (2008). Short-term mating strategies and attraction to masculinity in point-light walkers. *Evolution and Human Behavior, 29,* 65–69. doi:10.1016/j.evolhumbehav.2007.07.007

*Puts, D. A. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evolution and Human Behavior, 26,* 388–397. doi:10.1016/j.evolhumbehav.2005.03.001

Puts, D. A. (2010). Beauty and the beast: Mechanisms of sexual selection in humans. *Evolution and Human Behavior, 31,* 157–175. doi:10.1016/j.evolhumbehav.2010.02.005

*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. doi:10.1016/j.evolhumbehav.2005.11.002

*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. doi:10.1093/beheco/arp206

Raudenbush, S. W., & Bryk, A. S. (1985). Empirical Bayes meta-analysis. *Journal of Educational and Behavioral Statistics, 10,* 75–98. doi:10.3102/10769986010002075

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

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. doi:10.1177/01461672982412004

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57,* 199–226. doi:10.1146/annurev.psych.57.102904.190208

Rikowski, A., & Grammer, K. (1999). Human body odour, symmetry, and attractiveness. *Proceedings of the Royal Society B: Biological Sciences, 266,* 869–874. doi:10.1098/rspb.1999.0717

Roberts, M. L., Buchanan, K. L., & Evans, M. R. (2004). Testing the immunocompetence handicap hypothesis: A review of the evidence. *Animal Behaviour, 68,* 227–239. doi:10.1016/j.anbehav.2004.05.001

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: Biological Sciences, 273,* 2169–2175. doi:10.1098/rspb.2006.3569

*Roney, J. R., & Simmons, Z. L. (2008). Women's estradiol predicts preference for facial cues of men's testosterone. *Hormones and Behavior, 53,* 14–19. doi:10.1016/j.yhbeh.2007.09.008

*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. doi:10.1016/j.psyneuen.2010.10.010

Rowe, L., & Houle, D. (1996). The lek paradox and the capture of genetic variance by condition dependent traits. *Proceedings of the Royal Society B: Biological Sciences, 263,* 1415–1421. doi:10.1098/rspb.1996.0207

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. doi:10.1016/j.evolhumbehav.2008.08.006

*Rupp, H., 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. doi:10.1007/s12110-009-9056-6

Said, C. P., & Todorov, A. (2011). A statistical model of facial attractiveness. *Psychological Science, 22,* 1183–1190. doi:10.1177/0956797611419169

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. doi:10.1016/j.evolhumbehav.2007.10.001

Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society B: Biological Sciences, 276,* 575–584. doi:10.1098/rspb.2008.1177

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. doi:10.1177/0956797611417632

*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, Philadelphia, PA.

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. doi:10.1016/S1090-5138(03)00013-8

Strassmann, B. I. (1997). The biology of menstruation in *Homo sapiens*: Total lifetime menses, fecundity, and nonsynchrony in a natural-fertility population. *Current Anthropology, 38,* 123–129. doi:10.1086/204592

Stumpf, R. M., & Boesch, C. (2005). Does promiscuous mating preclude female choice? Female sexual strategies in chimpanzees (*Pan troglo-*

dytes verus) of the Taï National Park, Côte d'Ivoire. *Behavioral Ecology and Sociobiology, 57,* 511–524. doi:10.1007/s00265-004-0868-4

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

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.

*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. doi:10.1016/j.evolhumbehav.2013.01.003

Thornhill, R., & Gangestad, S. W. (1994). Human fluctuating asymmetry and sexual behavior. *Psychological Science, 5,* 297–302. doi:10.1111/j.1467-9280.1994.tb00629.x

Thornhill, R., & Gangestad, S. W. (1999a). Facial attractiveness. *Trends in Cognitive Sciences, 3,* 452–460. doi:10.1016/S1364-6613(99)01403-5

*Thornhill, R., & Gangestad, S. W. (1999b). The scent of symmetry: A human sex pheromone that signals fitness? *Evolution and Human Behavior, 20,* 175–201. doi:10.1016/S1090-5138(99)00005-7

Thornhill, R., & Gangestad, S. W. (2008). *The evolutionary biology of human female sexuality*. New York, NY: 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. doi:10.1093/beheco/arg043

Thornhill, R., & Møller, A. P. (1997). Developmental stability, disease, and medicine. *Biological Reviews, 72,* 497–548. doi:10.1111/j.1469-185X.1997.tb00022.x

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 of the United States of America, 104,* 15011–15016. doi:10.1073/pnas.0705290104

Van Dongen, S. (2006). Fluctuating asymmetry and developmental instability in evolutionary biology: Past, present, and future. *Journal of Evolutionary Biology, 19,* 1727–1743. doi:10.1111/j.1420-9101.2006.01175.x

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. doi:10.1016/j.evolhumbehav.2011.03.002

*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: Biological Sciences, 278,* 2223–2232. doi:10.1098/rspb.2010.2145

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. doi:10.1089/jwh.2005.14.248

*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. doi:10.1016/j.yhbeh.2007.01.010

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. doi:10.1016/S0010-7824(01)00191-3

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. *New England Journal of Medicine, 333,* 1517–1521. doi:10.1056/NEJM199512073332301

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. doi:10.1037/a0015300

Received March 2, 2013
Revision received October 16, 2013
Accepted October 23, 2013 ■

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

AQau—Please confirm the given-names and surnames are identified properly by the colors.
🟥= Given-Name, 🟦= Surname
The colors are for proofing purposes only. The colors will not appear online or in print.

AQ1—Author: In Tables 1 and 2, the Pillsworth et al. (unpublished) citation has no corresponding reference. Please provide reference, and give the year. Should citation be Li et al. instead?

AQ2—Author: Please confirm year of Li et al. citation. You had “unpublished.”

AQ3—Author: Please confirm year of Li et al. reference. You had “unpublished.”

AQ4—Author: In caption to Figure 1, please note edit defining error bars: “Errors bars represent standard error.” OK?