

Shifts in Methodology and Theory in Menstrual Cycle Research on Attraction

**Christine R. Harris, Aimee Chabot &
Laura Mickes**

Sex Roles

A Journal of Research

ISSN 0360-0025

Volume 69

Combined 9-10

Sex Roles (2013) 69:525-535

DOI 10.1007/s11199-013-0302-3



Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media New York. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Shifts in Methodology and Theory in Menstrual Cycle Research on Attraction

Christine R. Harris · Aimee Chabot · Laura Mickes

Published online: 19 October 2013
© Springer Science+Business Media New York 2013

Abstract This paper critically examines the hypothesis that different phases of the menstrual cycle induce changes in women's mate preferences. Empirically, we show that literature on this topic may be particularly prone to experimenter degrees of freedom, in which experimenters increase their likelihood of finding significant effects through elasticity in methodological and analytical strategies (e.g., flexibility in calculation of fertile and nonfertile phases, exclusion criteria, moderators, and analysis of dependent variables). Theoretically, we address misconceptions presented by Gildersleeve and colleagues (2013a). We reveal inconsistencies in the theoretical foundation for this work and discuss tension between theory and data. In short, there is sound reason to question whether reported menstrual cycle effects in women's mate preferences are indeed real.

Keywords Attractiveness · Evolutionary theory · Mate preferences · Masculinity preferences · Menstrual cycle shifts · Ovulatory cycle · Infidelity · Experimenter degrees of freedom · Replicability

Introduction

As described in detail below, a substantial number of published papers claim that women's attraction to different types of men changes across the menstrual cycle (e.g., Penton-Voak et al. 1999). In particular, these studies (primarily from both the US and the UK) report that when in a fertile phase of the menstrual

cycle, women are attracted to men who have particular traits such as high masculinity, symmetry, and testosterone; however, when in less fertile phases, women are attracted to men who are lower in these traits. This shift in the features women find attractive is proposed to reflect evolved strategies for women to be inseminated by more genetically fit males (i.e., masculine men), and to avoid insemination by relatively less fit males (i.e., feminine men). Original proponents of this hypothesis (which we refer to as the cycle shift hypothesis) argued for a fitness trade-off between choosing different types of men (e.g., Penton-Voak et al. 1999; Penton-Voak and Perrett 2000). Men with better genes purportedly had social characteristics that make them poorer long-term mates while men with worse genes had social characteristics that made them more desirable long-term partners. Therefore, according to this view, women's inclusive fitness could be maximized by pair-bonding with the less genetically fit men (e.g., feminine-faced men) while engaging in extra-pair mating with more genetically fit men (masculine-faced men) when conception is likely (thereby acquiring better genes for offspring). Inherent in this hypothesis is the assumption that changes in what women find attractive will result in changes in actual mating behavior (otherwise this preference change could not have become an adaptation). This assumption is reflected in the literature, which equates changes in attraction with changes in mating preferences. The current paper focuses exclusively on possible menstrual cycle effects on women's mate preferences (as opposed to how a woman's cycle might affect men's mate preferences).

Although findings on changes in women's mate preferences across the menstrual cycle have been widely accepted, these effects and the theory behind them are controversial. Harris (2011, 2012) raised empirical concerns regarding the literature on menstrual cycle shifts in women's preferences for masculinity in particular and also discussed theoretical shortcomings of the more general theory that women's mate preferences change across the menstrual cycle and reflect

C. R. Harris (✉) · A. Chabot
Psychology Department, University of California, San Diego, 9500
Gilman Dr. #0109, La Jolla, CA 92093-0109, USA
e-mail: charris@ucsd.edu

L. Mickes
Royal Holloway, University of London, Egham, Surrey, UK

evolved adaptations. Gildersleeve et al. (this issue) offer a critique of Harris (2011, 2012) and purport to uncover seven “misconceptions” in Harris’ analysis of the theoretical basis of menstrual cycle work on women’s mate preferences. Here we respond to that critique and discuss each purported misconception in turn. The claims of Gildersleeves et al. notwithstanding, we argue that Harris neither misunderstood nor misrepresented the menstrual cycle shift hypothesis. It seems to us that Gildersleeve and colleagues have made some major changes in their theoretical claims following the Harris papers (although they do not acknowledge their revisions as such). Below we raise empirical, theoretical, and logical concerns with their revised view, showing that it too suffers from theoretical inconsistencies and does not comport well with the empirical literature as a whole.

Empirical Shortcomings

The claim made by Gildersleeve et al. that there is strong empirical evidence for cycle shifts in women’s mate preferences is not supported as detailed below. In the case of attraction to masculinity, there are now a number of published studies that fail to replicate menstrual cycle preference shifts (Harris 2011; Peters et al. 2009; Rupp et al. 2009; Morrison et al. 2009). A meta-analysis by Wood et al., *in press*, provides a compelling case that, when the totality of studies on cycle effects on women’s mate preferences is examined, there is no indication of changes in masculinity preferences (Wood et al. *in press*; see also conference presentations Wood et al. 2012a, b; Wood and Joshi 2011). Gildersleeve et al. attempt to dismiss the Wood et al. meta-analysis by saying they have their own meta-analysis (presented at conferences, e.g., Gildersleeve et al. 2013b). Until we read their critique, we were unaware of any such work, and Haselton unfortunately declined our request to provide us with a copy of their meta-analysis (personal communication date, Feb. 25, 2013).

In our view, the Wood et al. *in press* meta-analysis, which provides a broad assessment across six traits, puts to rest the question of whether there are robust effects in this literature. Nevertheless, for the sake of completeness, we extend the points made in the meta-analysis by outlining additional issues that also cast serious doubt on any purported effects. We also discuss possible reasons why such effects sometimes appear in isolated papers and why even someone familiar with this literature might have been misled into thinking such effects were robust.

When Cycle Effects Are Found, Could Researcher Degrees of Freedom Be at Play?

Harris (2012) raised the possibility that flexibility in data analysis strategies (a concept referred to as *researcher degrees of freedom* by Simmons et al. 2011) might be allowing

potentially very high numbers of false-positive findings (i.e., Type 1 errors) to creep into the literature on menstrual cycle shifts in women’s mate preferences. Simmons et al. have shown that practices such as excluding subsets of participants, not reporting all analyses or conditions, including a variety of different covariates, and exploring diverse combinations and transformations of dependent measures can increase false-positive rates far above the nominal 5 % level adopted by most researchers. These practices have come to be colloquially termed “p-hacking”. Through simulations, Simmons et al. revealed that engaging in just four of these types of practices can result in a 60 % likelihood of finding an effect that appears significant at $p=.05$ but is in fact not real.

The elasticity in data analysis practices across studies within the menstrual cycle literature makes this literature a plausible candidate for having high rates of Type 1 errors. For example, as noted by Harris (2012; see also Wood et al. *in press*), there are a number of ways to calculate fertile and nonfertile cycle phases, and the practices vary greatly across studies. Moreover, significant effects have often been revealed *only* when moderators are employed (the effects do not appear in the sample as a whole but only in some subgroups; thus, the moderator is “essential” to find a significant effect). Yet, the use and effects of such moderators are inconsistent even within a single group of researchers, who meanwhile also vary from study to study in their method of determining cycle phase (see details below). Furthermore, some researchers use different transformations of dependent variables across studies without offering any justification of why they altered their practices (e.g., the use of sexual and investment attractiveness variables in Haselton and Gangestad 2006 vs. Pillsworth and Haselton 2006, discussed below).

Gildersleeve and colleagues (this issue) bristle at these suggestions, stating that Harris’ concerns are “very serious and unfounded”, adding that “differences between studies typically reflect constraints of study design and often reflect coherent progression in methodology”. We are glad that they view the issue of experimenter degrees of freedom as a serious one, but as we shall see below, their attempt to rebut concerns about this problem in their own work falls well short of compelling.

Menstrual Cycle Calculations—Flexibility in Days Counted as Fertile, as Not Fertile, and Who Gets Excluded Altogether

As noted above, the most obvious place in which experimenter degrees of freedom might arise is in the choice of how to calculate fertile and nonfertile cycle phases. To determine whether differences across studies are due to a “coherent progression in methodology”, as claimed by Gildersleeve et al., we examine a number of papers published by coauthors of the Gildersleeve et al. piece (specifically, DeBruine, Haselton, Frederick, Penton-Voak, Jones, and Perrett).

We first consider methods (backward vs. forward) for assessing fertility from self-report. In forward calculation, ovulation is estimated by counting forward from the first day of a woman's last menstrual cycle. In backward calculation, ovulation is predicted by counting backwards from the predicted (or actual) day of a woman's next period. Gildersleeve and colleagues claim, "the backward counting method is generally regarded as a more accurate method of estimating cycle position and fertility" (this issue). To use the backward method, researchers only need to ask one additional question (when a woman is likely to start her next period or the length of her typical cycle). Gildersleeve et al. also assert that "because study design typically dictates whether researchers are able to use forward vs. backward counting, there is no "researcher degrees of freedom" problem here" (this issue).

Gildersleeve and colleagues' assertions would lead one to expect a very strong trend over time: more recent studies should use the backward calculation method (given that they regard it as more accurate and their goal is to improve methodology). This is not the case. A sample of papers from this research group is listed in Table 1 to illustrate the inconsistency in methods, variables, and moderators across studies and time. An examination of Column 3 reveals that backward calculation was used early by this research group (e.g., Penton-Voak et al. 1999) but then has been inconsistently used over time. In fact, their most recent studies appear quite frequently to employ forward calculations.

A number of other aspects of cycle calculations are also strikingly variable, further raising concerns regarding experimenter degrees of freedom. The studies displayed in Fig. 1 all come from papers published by the coauthors of the Gildersleeve et al. paper. The top part of the figure displays studies that have employed a backwards calculation to determine the fertile phase while the bottom part displays studies that have employed a forward calculation method. (Note the two types of studies are on slightly different scales because the scale is dictated by the restrictions an experimenter placed on participants regarding cycle length; this accentuates variability across forward and backward calculation methods.)

As the figure shows, even within each type of calculation (forward vs. backward) decisions are inconsistent in terms of which days are counted as fertile, which as infertile, and which are excluded altogether. If such variability occurs just within a single group of collaborators, one expects that there is probably even more variability across the literature as a whole. Despite Harris (2011, 2012) raising concerns about these practices, Gildersleeve and colleagues did not offer any explanation nor did they acknowledge that there is a potential problem.

The Case of Moderators—Shifting Sands

Another way in which false-positives might emerge in the literature is by testing a host of moderators to obtain

significant effects. Clearly some psychological phenomena are moderated and are influenced by different promoting and inhibiting conditions. However, in the menstrual cycle literature, the use of moderators as well as their inconsistent effects prove problematic, raising further concerns about whether there are indeed cycle effects on women's mate preferences.

There are two different ways that moderators could be theorized to play a role in the cycle effect literature. The first possibility is that moderators are necessary for an effect to be revealed (i.e., the effect only occurs under some contexts). If so, then once a moderator is revealed it should be consistently assessed in future studies. The second possibility is that menstrual cycle effects occur regardless of moderators, but the moderators may sometimes amplify effects. If this is the case, then one would expect to see main effects of cycle phase whether or not a moderator is employed. Below we examine moderation in menstrual cycle studies and show that research does not reliably fit either path of moderation.

Relationship Status and Relationship Context—Now We See It, Now We Don't

The application of moderators in the menstrual cycle mate preference literature has been disconcertingly inconsistent. We focus here on two such moderators: 1) participant's relationship status—single or paired and 2) relationship context—judging attractiveness for a short-term relationship (e.g., a one-night stand) vs. a long-term relationship. (As noted previously, this literature assumes that changes in attractiveness ratings serve as a proxy for changes in mating preferences). When Penton-Voak and colleagues' (1999, Experiment 1) initially reported that women's attraction to masculinity changed across the menstrual cycle, no moderators were needed to reveal the cycle phase effect. Relationship status was assessed, but the cycle shift effect was found for both single and paired women, although more strongly so in the latter group. In Experiment 2 of that same paper, cycle shifts in masculinity preferences were again examined. There is no mention of participant's relationship status but this time a different potential moderator was measured: whether attractiveness was being judged for a short-term relationship or a long-term relationship. Masculinity preference shifts were found for the short-term relationship judgments, but *no* shift occurred for the long-term relationship judgments. Given these findings one would expect that future studies, particularly by Penton-Voak's team, would assess what appeared to be an essential moderator—relationship context (or specify that the judgment was for a short-term relationship).

This assessment, however, was not done. In their next paper on menstrual cycle shifts in women's mate preferences (Penton-Voak and Perrett 2000), relationship context is not

Table 1 Sample of studies from the coauthors of the Gildersleeve et al. paper (e.g., DeBruine, Haselton, Jones, Penton-Voak, & Perrett) (in chronological order) illustrating the inconsistency in the use of variables, methods, and moderators (relationship status and relationship context) to assess menstrual cycle effects on attractiveness and mate preferences across studies and time

Study	Variable	Fertility calculation method ^a	Was relationship status examined? (Single vs. Paired)	Was relationship context specified? (Short-term vs. Long-term)
Penton-Voak et al. (1999)				
Study 1	Masculinity	Backward	Yes	No
Study 2	Masculinity	Not specified	No	Yes ^b
Penton-Voak and Perrett (2000)				
	Masculinity	Forward	No	No
DeBruine et al. (2005)				
	Self-resemblance	Forward	No	No
Jones et al. (2005a)				
Study 1	Masculinity	Forward	Sample included only women in relationships	No
Study 2	Masculinity & apparent health	Backward	Yes, but analysis not reported	No
Jones et al. (2005b)				
Study 1	Apparent health	Backward	Yes, but analysis not reported	No
Study 2	Apparent health	Kits	No	No
Study 3	Apparent health	Kits	No	Yes ^b
Feinberg et al. (2006)				
	Masculinity	Kits	No	No
Haselton and Gangestad (2006)				
	Flirtation/ Attraction	Backward	Yes	No
Little et al. (2007b)				
Study 1	Symmetry	Kits	No	No
Study 2	Symmetry	Forward	Yes ^{b,c}	Yes ^{b,c}
Little et al. (2007a)				
Study 1	Masculinity	Forward	Yes (ns)	Yes ^{b,c}
Study 2	Masculinity	Forward	Yes, but analysis not reported	Yes ^b
Little et al. (2008)				
	Masculinity	Forward	Yes ^{b,c}	No
Morrison et al. (2009)				
	Masculinity & flirtatiousness	Forward	No	Yes short-term only
Little and Jones (2012) ^d				
Study 1	Masculinity	Forward	No	Yes ^b
Study 2	Masculinity	Forward	No	Yes ^b

^a Method used to group women as fertile or not fertile. Forward = ovulation was estimated by counting forward from first day of a woman's last menstrual cycle. Backward = ovulation was predicted by counting backwards from the predicted (or actual) day of a woman's next period. Kits = hormonal assays were used to assess fertility risk

^b Indicates that moderation by the variable is necessary in order to reveal cycle effects (i.e., effect only occurred in women who were in relationships or only occurred when the context was specified as short-term)

^c There was a significant main effect of cycle phase; however, based on interaction terms and follow up analyses, the original authors suggested effects only occurred for partnered women or only occurred for short-term context

^d Despite obtaining information about cycle length, analyses were performed using forward rather than backward calculations (which is at odds with Gildersleeve and colleagues' claims that backward calculation is superior and is a method that has been adopted by later research)

ns not significant

specified, and yet a significant cycle shift was reported. It seems odd that having discovered what appeared to be an essential moderator, Penton-Voak, Perrett, Jones, Little and colleagues simply stopped examining it (or at least, did not report results for this variable) in a number of subsequent

studies. Table 1 lists, in chronological order, a sample of papers from this research team. As can be seen in the last column of Table 1, the relationship context moderator is often notably absent, even in papers that specifically examine masculinity preferences (Jones et al. 2005a; Feinberg

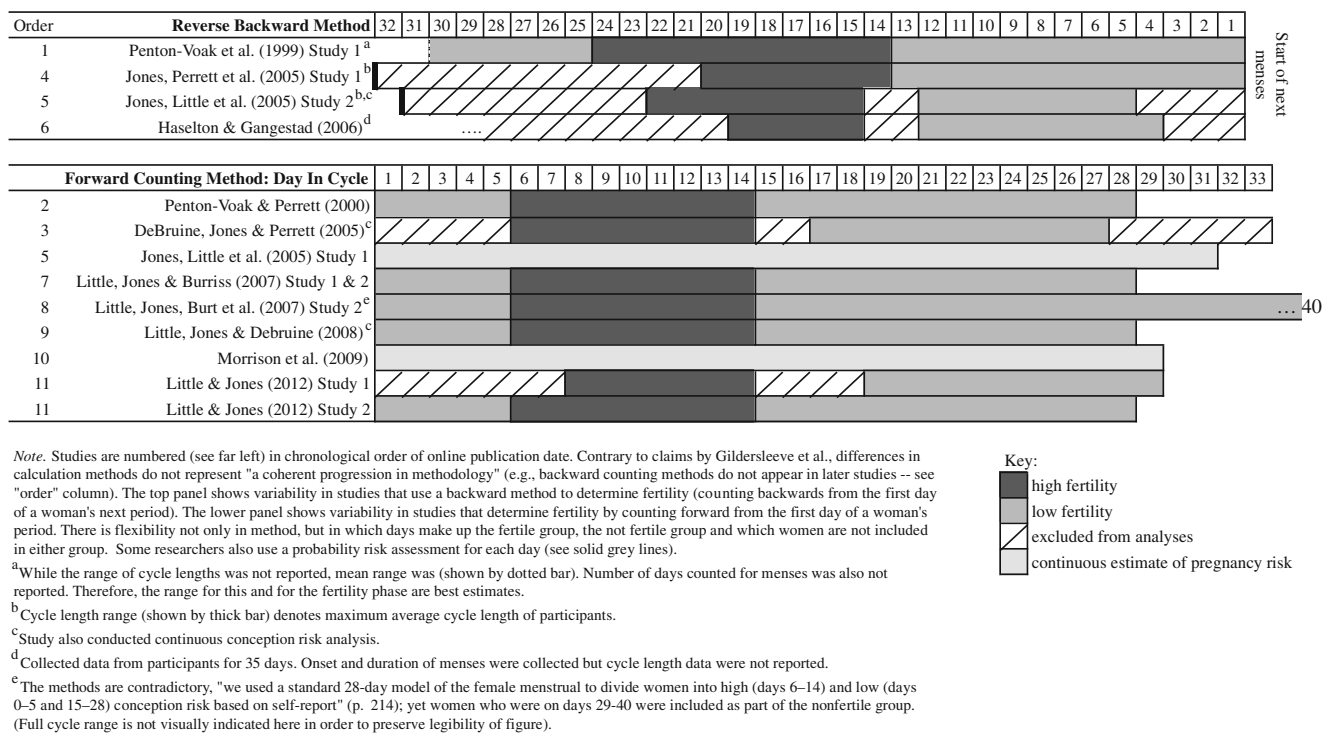


Fig. 1 Demonstration of the variability of fertility classification for self-report methods by a single research Group (Jones, Little, and Colleagues)

et al. 2006; Little et al. 2007b, Study 1). Furthermore, these papers provide no clear justification for why relationship context appears and disappears in this literature in a peek-a-boo fashion. It is clearly *not* the case that after the effect of a given moderator was discovered, the moderator is regularly assessed in subsequent research (see also Wood et al. *in press*, for similar conclusions). At the very least, this illustrates the field's failure to build on itself in a systematic way, but it also raises concerns about both experimenter degrees of freedom (possible "p-hacking") and about whether analyses are being reported in a transparent and complete fashion.

Still another possible moderator, relationship status, also figures in these researchers' analyses in an inconsistent manner, as shown in Column 4 of Table 1. For example, in Study 1 of Jones et al. (2005a) the sample was limited to women who are in relationships. Their Study 2 assessed relationship status, but did not report analyses of this variable. Both Little et al. (2007b) and (2007a) have similar inconsistencies regarding analysis of participants' relationship status. In each of these papers, two studies are reported yet relationship status is only analyzed in one of the two studies. Interestingly, relationship status re-emerged prominently in 2008—this time moderation is *necessary* for a cycle effect to be found (Little et al. 2008). As described above, such effects were only found for women in relationships; there was no hint of a menstrual cycle shift effect for single women. Importantly, the inconsistent use of moderators is not only seen in the

early days of this literature, but persists. For example, Little and Jones (2012) do not report examining participants' relationship status, despite it being an essential moderator in their 2008 paper with DeBruine (Little et al. 2008).

Clearly, in any given study, an experimenter might neglect to include a variable, but the inexplicable appearance and disappearance of key moderator variables from one study to the next, particularly when such variables were originally presented as important, is quite disconcerting. The picture becomes even more convoluted when one takes into account that there are other papers by this research group that report the effects of relationship status and/or relationship context on women's mate preferences but do not include measures of menstrual cycle (e.g., Little et al. 2002a, 2011).

In sum, contrary to Gildersleeve and colleagues' claims, moderation does not explain the disparate findings in the cycle literature. As outlined earlier, there are two possible paths by which moderators could influence effects. The first is that a moderator is required to reveal an effect (i.e., the effect only occurs in certain contexts). If this is the case, then it is baffling why the research by Perrett, Penton—Voak, DeBruine and colleagues has not consistently included (and reported) these key moderators, and why the effects are so inconsistently dependent on "essential" moderators. The second possibility is that menstrual cycle effects occur regardless of moderators, but the moderators magnify effects. If this is the case, then the failures to find menstrual

cycle shift effects in a number of papers cannot logically be attributed to a failure of those papers to assess potential moderator variables (as Little and Jones 2012, have attempted to argue).

Partner Characteristics—Yet Another Peek-a-boo Moderator

A similar set of concerns surround moderators in studies of a woman's interest in extra-pair mating across the menstrual cycle. As discussed in detail in other sections, proponents of cycle shifts in women's mate preferences have focused heavily on women in relationships, and have attempted to use studies on women's desires for men other than their primary partners as supplemental support. Inherent in the cycle hypothesis is the assumption that women should primarily desire (and therefore, engage in) extra-pair mating during the phase of peak fertility (when better genes can be obtained for offspring). As described in Harris (2012), the literature does not support this assumption. Gildersleeve et al. respond by stating that the cycle shift hypothesis does not predict "that women *in general* will experience a high-fertility increase in attractions to men other than their partners (i.e., a main effect of fertility on attractions to other men)" (this issue). They contend that they have proposed, and found evidence for, several a priori moderators that affect whether a woman has increased extra-pair attraction during peak fertility. The moderators they focus on are characteristics of the primary mate.

Unfortunately, the studies cited by Gildersleeve et al. suffer from the same types of problems raised above for other types of moderators, posing questions of whether any effects found due to mate characteristic moderators are the result of Type 1 errors. For example, one would expect that once a factor was predicted and found to be essential in order to reveal cycle shift effects on extra-pair mate attraction, this factor would be similarly assessed in future studies, particularly those conducted within the same lab. However, this is not the case. Haselton and Gangestad (2006) used a difference score of the primary mate's sexual vs. investment attractiveness as a moderator of women's desire for extra-pair mating. Pillsworth and Haselton (2006) also assessed these types of attractiveness, but chose to analyze them separately rather than compute a difference score. These "moderators" of women's extra-pair attraction, however, are not included in Durante et al. (2008). Then in 2012, sexual attractiveness is again assessed by Haselton's group (Larson et al. 2012), but analyses that amount to direct replication attempts of Pillsworth and Haselton (2006) do not reveal significant effects (e.g., women's ratings of partner sexual attractiveness only moderated in-pair attraction, not attraction to other men). Performing a number of analyses that include a variety of moderators and dependent variables easily increases the likelihood that effects are the results of Type 1 errors (cf Simmons et al. 2011).

In short, in the 13 years since the much-cited publication of claims about cycle shift effects on women's masculinity preferences (Penton-Voak et al. 1999), a host of moderators have been argued to be essential for the effect to be revealed, but then have either vanished from published studies (either not assessed or not reported) or have failed to be important, only to sometimes remerge in other studies as essential. This shifting pattern in a literature would be consistent with the inadvertent exploitation of "experimenter degrees of freedom" (p-hacking) to obtain significant effects (Simmons et al. 2011). For example, rather than reflecting the effect of a true moderator, cycle effects under such circumstances might reflect false-positives due to the analysis of many potential moderators, only some of which get reported. When an effect inconsistently requires various moderators in order to be revealed, it is reasonable to lose faith in the "effect".

Other Odd Features of the Empirical Literature

There are a number of other aspects of the menstrual cycle literature on women's mate preferences that makes one wonder about the reliability and meaningfulness of significant effects. For one, a large number of studies published by Jones, Little, and colleagues seem to have relied on participants who were recruited in potentially problematic ways including through the researchers' own website, which also appear to have contained links to their published papers (e.g., Little and Jones 2012; Little et al. 2007a), and through Science websites and magazines, which also may have been running stories on these topics (e.g., Jones et al. 2005a, b; Penton-Voak and Perrett 2000). Thus, it is possible that at least some participants may have been alerted to the hypotheses.

Second, Gildersleeve et al. and DeBruine et al. raise a number of objections to Harris (2011, 2012), which they suggest could account for her failure to find menstrual cycle effects on women's mate preferences. Yet, the methods these researchers are objecting to are used in a preponderance of their own papers making it unlikely that these factors alone could be responsible for null findings. For example, DeBruine et al. (2005) employed a between-subjects design *without* hormonal assays and included women up to 35 years of age. They reported a significant cycle effect with a sample of only 43 women (21 in the fertile group). Yet, they take issue with Harris's sample size of 248 (80 in the fertile group, 19 of whom were under 30 years of age). Similarly, Little et al. (2007a) reported between-subject cycle effects on masculinity preferences in an Internet sample of 97 women (36 high fertility). It seems odd that effects would be seen with small, underpowered studies such as these but not in larger studies that use similar types of methodology (cf Pashler and Harris 2012).

Two Additional Failures to Replicate Menstrual Cycle Shift Effects on Women's Masculinity Preferences in Mates

Gildersleeve et al. and DeBruine et al. (2010) expressed concern that Harris' failure to replicate cycle effects on women's mate preferences might be due to participant age. They note the small number of women in Harris' sample who were in their 20s, but fail to mention that even in this subsample, the effect was in the *opposite* direction of the cycle hypothesis: Young women during peak fertility preferred more feminized faces; and there was no hint of an interaction of age and mate preferences. Gildersleeve et al. also suggest that a cycle effect might have been revealed had the relationship context been specified as short-term, e.g., a one-night stand, (although as noted above, there is no consistent effect, or even use, of this moderator in the literature).

To address these possibilities, we conducted two additional studies (Mickes & Harris, in preparation) that examined the effect of cycle changes on women's preferences for masculinity in mates. Combined these new studies had over 800 participants, including a large sample of young women (age 30 or younger). In both studies, participants were premenopausal, not pregnant, not on *any* form of hormonal birth control, and reported experiencing regular menstrual cycles. Moreover, one study specified that participants make attractiveness judgments with regard to a short-term relationship. Neither study found evidence of any hint of a menstrual cycle shift effect.

In Study 1, our participants ($n=563$) judged the attractiveness of two sets of faces (using stimuli from Penton-Voak et al. 1999). Women were classified in the high fertility group if they completed the study on current cycle day 6–14, and in the low fertility group on cycle days 0–5 and 15–28 (following methods of Penton-Voak and Perrett 2000; Little and Jones 2012, Study 2). Analyses examined effects of cycle phase, relationship status (single or paired), and their interaction. No significant effects of cycle phase were found. The same analyses were also run including only those who were 30 years of age or less ($n=303$). The results again showed no effects of cycle phase, relationship status, or an interaction.

In Study 2, we performed yet another replication attempt with these same facial masculinity stimuli using the same method for fertility classification. Here we focused only on younger women and stipulated that the attractiveness judgments should be for a short-term relationship—two factors that Gildersleeve et al. and DeBruine et al. argue should help reveal an effect, if there is indeed one. Participants were between 18 and 30 years old ($n=266$; average age=21.7). As in Study 1, we did not find a significant effect of cycle phase on masculinity preferences. Thus, even after directly addressing the concerns of Gildersleeve et al. and DeBruine et al. by accounting for moderators and other issues they claim are critical (relationship context, increased sample size,

excluding women over 30 years of age, excluding any type of hormonal birth control, not just birth control pills), we still do not find any signs of a cycle shift on masculinity preferences.

These new data, therefore, suggest that Harris' previous replication failure was not due to overlooking important moderators or using the wrong exclusionary criteria, but instead add additional weight to the growing list of failed replications and the conclusion from the Wood et al. meta-analysis.

Theoretical Issues and Throwing Sand?

In the previous section, we have presented a number of lines of evidence to support the claim that there is not robust empirical evidence for menstrual cycle shifts in women's mate preferences, particularly in the case of masculinity preferences. Our paper could close on that note. However, we feel it is important to set the theoretical record straight given many of the claims made by Gildersleeve and colleagues. In their critique, they lay out several issues that they misleadingly imply are misunderstood by Harris (2012). Below we address these.

Infidelity

Most of what Gildersleeve and colleagues take issue with centers around infidelity. Harris (2012, this issue), drawing on writings of several evolutionary psychologists (e.g., Penton-Voak et al. 1999; Little et al. 2002b; Haselton and Gangestad 2006), writes:

At its core, the cycle shift hypothesis is about female infidelity and inherent in it are several key underlying assumptions about the nature of such infidelities (see also Harris 2011). For one, the theory requires that women in the ancestral environment engaged in infidelity at high rates, and reaped inclusive fitness benefit from doing so; otherwise the purported preference shift would not have become an adaptation. Proponents further argue that these extra-pair matings were short-term and specifically occurred during the follicular phase (e.g., Penton-Voak et al. 1999). Logically, all these assumptions would need to hold in order for the cycle shift hypothesis to make sense. This is because at its core the cycle preference shift hypothesis presumes that women will be mating with different men when they are fertile (more masculine men) vs. not fertile (less masculine men). (Changes in attractiveness preferences need to translate into changes in behavior in order for such preferences to have emerged as evolutionary adaptations.) As discussed in Harris (2011), if

women were engaging in sex with short-term partners for longer than the small window of time when conception was likely, then there would be no need for a preference shift during the menstrual cycle. (this issue)

In other words, if a woman has sex with her lover throughout her cycle, the exact day in which the affair began would be irrelevant.

Gildersleeve et al. attempt to refute this logic, claiming that each of the following is a misconception: the cycle shift hypothesis is only about female infidelity, requires that female infidelity occurred primarily at high fertility and more frequently than sex with the primary partner, and requires that female infidelity occurred at a high rate [in the ancestral environment].

Gildersleeve et al. concede that infidelity has been emphasized in their previous explanations of menstrual cycle effects but now claim that the “cycle shift hypothesis does not require any female infidelity” (this issue). They further propose that, “One possibility is that psychological mechanisms that produce cycle shifts evolved in ancestral species that did not typically pair-bond because these shifts increased females’ probability of conceiving with males of relatively high genetic quality” (this issue). Although they do not frame it so, this is a major alteration of their previous theorizing regarding origins of possible menstrual cycle shifts in women’s masculinity preferences (e.g., see Penton-Voak et al. 1999; Haselton and Gangestad 2006).

The contention that menstrual cycle shifts evolved in the context of nonpair bonding makes a very clear prediction: Single women should show the cycle shift in mate preferences every bit as strongly as paired women. Moreover, one could argue that the *strongest* effects should occur in single women not in paired women. This is because, according to Gildersleeve and colleagues, moderators involving the primary partners’ characteristics (e.g., his sexual attractiveness) can reduce menstrual cycle effects in women with partners. Interestingly, past research does not comport well with their revised position.

As noted earlier, evolutionary psychologists have examined relationship status but their arguments, when this moderator has been used, have always taken the form that the effect should appear in paired, not single women (hence the focus on infidelity). For example, in a relatively recent paper, Little et al. (2008) report,

... women preferred more masculine faces in the follicular phase than in the luteal phase only when they had a partner. Splitting the sample on partner confirmed a significant effect of fertility for women with a partner ($F(2,75)=9.00$, $p=.002$) but not for those without a partner ($F(2,69)=0.04$, $p=.845$). (p. 480).

Moreover, Penton-Voak et al. (1999) reported a trend towards a stronger cycle shift effect in women with partners relative to those without. Presumably, these authors were conducting analyses on partner status for a theoretical reason. In other work, Little et al. (2002b) provide that reasoning,

Once women have secured a partner, those seeking extra-pair partners may be more attentive to cues to possible heritable quality (as signaled by androgenic effects on face shape) than those not in a relationship, cuckolded her current partner and providing better genes for their children. For females not in a relationship, maximizing offspring fitness and raising the child alone may not be as successful a strategy as securing a long-term investment from a male. (p. 72).

Thus, despite Gildersleeve and colleagues’ claims to the contrary, Harris (2012) accurately portrayed previous claims regarding infidelity advanced by theorists proposing cycle effects on women’s mate preferences. Moreover, Gildersleeve and colleagues’ critique exposes a major revision of their views on the cycle shift hypothesis and infidelity, although they do not acknowledge this. Finally, their revision does not square with a number of aspects in the literature.

Gildersleeve and colleagues’ revised view of the origin of cycle preference shifts is not only inconsistent with most previous research but even statements within their critique are inconsistent. For example, they repeatedly state that the menstrual cycle view is not a view about infidelity and yet they then write, “Harris (2012) correctly notes that studies examining shifts in women’s attraction to men other than their primary partners are particularly relevant to the hypothesis that women’s mate preferences shift across the cycle.” (this issue). If the cycle shift hypothesis is not at its core about infidelity, then it does not make sense to focus on women in relationships. Not only is this revised hypothesis of Gildersleeve et al. inconsistent with their previous choices to use relationship status as a moderator in statistical analyses, but it is also not in keeping with their previous findings that when relationship status is a significant moderator, paired women, not single women, are the ones who showed the cycle shift.

In sum, Gildersleeve and colleagues seem confused about their own past and present theoretical arguments, and about how their current theorizing relates to their previous findings and choices in analytic practices. Theoretical revisions are fine and are how a field grows, but they need to be based on cumulative empirical data. Gildersleeve and colleagues appear to alter their theory from paper to paper without acknowledging that such alterations are not in keeping with their own prior findings.

Masculinity, Immunocompetence, and Good Genes

There is no doubt that in previous papers the coauthors of the Gildersleeve et al. proposed the immunocompetence handicap hypothesis (ICHH) as an explanation for purported menstrual cycle effects on women's mate preferences. According to the ICHH, originally proposed by Folstad and Karter (1992) for nonhuman animals, testosterone not only promotes secondary sexual characteristics in males but also suppresses the immune system, creating greater vulnerability to parasitic illness. Folstad and Karter (1992) suggested that males who appear to be flourishing despite the immunological handicap must be genetically fit and that females should therefore prefer to mate with these males to obtain good genes for offspring.

Proponents of cycle shifts in female mate preferences have clearly relied on the ICHH to explain women's mate preferences. For example, Penton-Voak et al. (1999) write

A female might choose a primary partner whose low masculine appearance suggests cooperation in parental care ('long-term' preferences are unchanged across the menstrual cycle) but occasionally copulate with a male with a more masculine appearance (indicating good immunocompetence) when conception is most likely. (p. 741).

Moreover in a recent summary of the field, Penton-Voak and colleagues (Scott et al. 2013) are very clear on what the literature has claimed:

The central role of heritable immunocompetence signaling properties of masculinity have increasingly been treated as established, rather than hypothetical. Research papers and textbooks on evolutionary psychology present the immunocompetence hypothesis of masculinity preferences as plausible, well established, or even as factual despite the absence of direct tests (Rossano 2003; Bressan and Stranieri 2008; Cartwright 2008; Little et al. 2010; DeBruine et al. 2010a; DeBruine et al. 2010b) (p. 580).

Therefore, we are at a loss to understand why Gildersleeve et al. claim that it is a misconception to state that the ICHH has been integral to the cycle shift hypothesis of women's mate preferences.

Harris (2011, 2012) argued that using the ICHH to account for cycle effects in women's mate preferences is flawed (see also Scott et al. 2013, for a similar conclusion). It appears that Gildersleeve et al. now also have doubts about the immunocompetence hypothesis; they write, "recent work has, by and large, not supported this particular mechanism" (p.15). Gildersleeve and colleagues' reconsidered stance on this issue is a step in the right direction to our minds; it is nonetheless puzzling that rather than acknowledge their

updated thinking, they inaccurately accused Harris of misunderstanding the literature.

Having abandoned the immunocompetence hypothesis, Gildersleeve and coauthors, are left with two serious theoretical/empirical weaknesses, which must be overcome in order for the cycle shift hypothesis of women's mate preferences to make any sense. Their theory requires 1) evidence that masculine men have better genes and 2) that there be some negative aspect to masculine men, otherwise women would choose them as mates across all phases of the menstrual cycle. Gildersleeve and colleagues simply fail to provide compelling evidence for these propositions. Results, cited by Gildersleeve et al., from a few studies that report that men with masculine faces tend to have more uncommitted and short-term sexual relationships does not speak to whether there was an inclusive fitness cost of having a long-term relationship with these types of men in the ancestral past. It is a large logical leap to assume that such findings can be taken as evidence that ancestral women would have received less investment from these men and to such a degree that it would have impacted the female's inclusive fitness. Moreover, even if there were some inclusive fitness outcomes, there is no evidence that this formed the basis for a psychological adaptation that drives modern women's mate preferences.

Other "Misconceptions"—Much Ado About Nothing

The final two misconceptions that Gildersleeve and colleagues attribute to Harris do not appear to be misunderstandings at all.

First, Gildersleeve and colleagues claim that it is a misconception to state, "The cycle shift hypothesis posits "hard-wired" psychological mechanisms". We find it difficult to understand how it can be contended that evolved predispositions are *not* hard-wired; and Gildersleeve et al. fail to expand on this point. In fact, their whole section on this particular "misconception" focuses on flexibility of mechanisms rather than explaining why they do not believe that evolved psychological mechanisms are "hard-wired". While it is fine to emphasize flexibility, to refer to Harris' use of the term "hard-wired" as a misconception is misleading.

Lastly, we address Gildersleeve and colleagues' unwarranted assertion that Harris believes, "The cycle shift hypothesis posits that these mechanisms evolved in and are currently fitness-enhancing in humans." In summarizing predictions of the cycle shift hypothesis, Harris (2012) wrote,

Therefore, to maximize inclusive fitness, women pair-bond with more feminine-faced men and thus reap the benefits of having them as permanent partners, but should seek to mate with masculine-faced men when conception is likely in order to obtain the best genes for their offspring.

In quoting this line, Gildersleeve and coauthors misconstrue Harris as saying that evolutionary psychologists are making “recommendations for what women should do” (this issue) in some prescriptive way. We hope that no reader of Harris (2012) would come away with the same misimpression, but just to be absolutely clear; Harris does not think, nor did she ever think, that the theory of menstrual cycle effects was providing relationship advice to women. Harris used the term “should” in the quoted sentence to mean “the theory predicts that this behavior will occur”. Thus, the use of “should” in this context was used in the same manner that Haselton and Miller (2006) presumably used the term when they wrote, “a hungry fertile female should care about luck versus skill in resource acquisition, and about heritable fitness indicators that have no immediate resource payoffs.” (p. 54) and “women near peak fertility mid-cycle should prefer creativity over wealth, especially in short-term mating” (p. 50).

The idea that cycle effects did not initially evolve in humans seems to be a new proposition for most of the theorists on the Gildersleeve paper, and, as detailed above, raises an additional set of problems for the cycle shift view.

Summary—Misconceptions is a Misnomer

In summary, Gildersleeve and colleagues characterized the writings of Harris (2011, 2012) as presenting seven theoretical misconceptions. We have discussed each purported misconception; and have provided evidence, including quotes from coauthors of the Gildersleeve et al. piece, to show that Harris accurately describe the theoretical foundations of the cycle shift hypothesis. The original theory rested on a number of assumptions about female infidelity in the ancestral past, and proposed the ICHH as a foundation for claiming that masculine men had better genes. Proponents of the cycle shift hypothesis of women’s mate preferences still posit “hard-wired” psychological mechanisms and still fail to provide data for the propositions that masculine men have better genes and make worse long-term mates. Finally, Harris never made any claims about cycle effects being currently fitness-enhancing in humans.

Concluding Remarks

Despite the large number of published studies on cycle shifts in women’s mate preferences, there may be good reason to doubt the reality of these effects (at least those reviewed here). Recently, there has been growing awareness that a number of methodological and analytical practices, which are routinely engaged in by many researchers, can greatly distort a literature (e.g., see Pashler and Wagenmakers 2012). As Simmons et al. (2011) demonstrate, such practices can lead to apparent effects where there are, in fact, none. There

are a number of aspects of the menstrual cycle literature on women’s mate preferences that make it particularly suspect. It is our hope that this discussion will inspire researchers who want to study topics such as these to join the growing movement within psychology, and science more broadly, toward employing tighter research methodology. In the present case, what would be particularly useful would be well-powered studies that register all methods and predictions in advance, especially relating to the method of distinguishing high fertile from low fertile periods and the forms of moderation to be expected. This will undoubtedly lead to fewer empirical papers on the topic, but could greatly increase the credibility of the literature in this area.

Acknowledgments We thank Wendy Wood and Hal Pashler for helpful comments and discussion and appreciate Noriko Coburn’s assistance with manuscript preparation.

References

- 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.
- Cartwright, J. (2008). *Evolution and human behavior: Darwinian perspectives on human nature*. 2nd ed. Cambridge, MA: MIT Press.
- 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.
- DeBruine, L. M., Jones, B. C., Frederick, D. A., Haselton, H. 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.
- 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.
- Feinberg, D. R., Jones, B. C., Law Smith, M. J., Moore, F. R., DeBruine, L. M., Cornwell, R. E., Hillier, S. G., & Perrett, D. I. (2006). Menstrual cycle, trait estrogen level and masculinity preferences in the human voice. *Hormones and Behavior*, *49*, 215–222. doi:10.1016/j.yhbeh.2005.07.004.
- Folstad, I., & Karter, A. J. (1992). Parasites, bright males, and the immunocompetence handicap. *The American Naturalist*, *139*, 603–622.
- Gildersleeve, K., DeBruine, L., Haselton, M. G., Frederick, D. A., Penton-Voak, I. S., Jones, B. C., & Perrett, D. I. (2013a). Shifts in women’s mate preferences across the ovulatory cycle: A critique of Harris (2011) and Harris (2012). *Sex Roles*, this issue doi:10.1007/s11199-013-0273-4.
- Gildersleeve, K., Haselton, M. G., Fales, M. (2013b, January). Do women’s mate preferences change across the ovulatory cycle? A meta-analytic review. Poster presented at the conference of the Society for Personality and Social Psychology, New Orleans, LA.
- Harris, C. R. (2011). Menstrual cycle and facial preferences reconsidered. *Sex Roles*, *64*, 669–681. doi:10.1007/s11199-010-9772-8.

- Harris, C. R. (2012). Shifts in masculinity preferences across the menstrual cycle: Still not there. *Sex Roles*, this issue. doi:10.1007/s11199-012-0229-0.
- 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.
- Jones, B. C., Little, A. C., Boothroyd, L., DeBruine, L. M., Feinberg, D. R., Smith, M. J., ... Perrett, D. I. (2005a). Commitment to relationships and preferences for femininity and apparent health in faces are strongest on days of the menstrual cycle when progesterone level is high. *Hormones and Behavior*, 48, 283–290. doi:10.1016/j.yhbeh.2005.03.010
- Jones, B. C., Perrett, D. I., Little, A. C., Boothroyd, L. G., Cornwell, R. E., Feinberg, D. R., ... Moore, F. R. (2005b). Menstrual cycle, pregnancy and oral contraceptive use alter attraction to apparent health in faces. *Proceedings of the Royal Society of London B*, 272, 347–354. doi:10.1098/rspb.2004.2962
- 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.
- Little, A. C., & Jones, B. C. (2012). Variation in facial masculinity and symmetry preferences across the menstrual cycle is moderated by relationship context. *Psychoneuroendocrinology*, 37, 999–1008. doi:10.1016/j.psyneuen.2011.11.007.
- Little, A. C., Jones, B. C., Penton-Voak, I. S., Burt, D. M., & Perrett, D. I. (2002a). Partnership status and the temporal context of relationships influence human female preferences for sexual dimorphism in male face shape. *Proceedings of the Royal Society of London B*, 269, 1095–1100. doi:10.1098/rspb.2002.1984.
- Little, A. C., Penton-Voak, I. S., Burt, D. M., & Perrett, D. I. (2002b). Evolution and individual differences in the perception of attractiveness: How cyclic hormonal changes and self-perceived attractiveness influence female preferences for male faces. In G. Rhodes & L. Zebrowitz (Eds.), *Advances in social cognition, volume 1: Facial Attractiveness* (pp. 59–90). Westport: Ablex.
- Little, A. C., Jones, B. C., & Burriss, R. P. (2007a). Preferences for masculinity in male bodies change across the menstrual cycle. *Hormones and Behavior*, 31, 633–639. doi:10.1016/j.yhbeh.2007.03.006.
- Little, A. C., Jones, B. C., Burt, D. M., & Perrett, D. I. (2007b). 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 male 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.
- Little, A. C., Saxton, T. K., Roberts, S. C., Jones, B. C., DeBruine, L. M., Vukovic, J., ... Chenore, T. (2010). Women's preferences for masculinity in male faces are highest during reproductive age range and lower around puberty and post-menopause. *Psychoneuroendocrinology*, 35, 912–920.
- 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.
- Morrison, E. R., Clark, A. P., Galewski, L., Campbell, N., & Penton-Voak, I. S. (2009). Women's probability of conception is associated with their preferences for flirtatious but not masculine facial movement. *Archives of Sexual Behavior*, 39, 1297–1304. doi:10.1007/s10508009-9527-1.
- Pashler, H. E., & Harris, C. R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science*, 7, 531–536. doi:10.1177/1745691612463401.
- Pashler, H. E., & Wagenmakers, E. J. (Eds.) (2012). Special section on replicability in psychological science: A crisis of confidence? [Special section]. *Perspectives on Psychological Science*, 7, 528–689.
- Penton-Voak, I. S., & Perrett, D. I. (2000). Female preference for male faces changes 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, L. K., & Minamisawa, R. (1999). Menstrual cycle alters face preference. *Nature*, 399, 741–742. doi:10.1038/21557.
- 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.
- Pillsworth, E. G., & Haselton, M. G. (2006). Male sexual attractiveness predicts differential ovulatory shifts in female extra-pair attraction and mate retention. *Evolution and Human Behavior*, 27, 247–258. doi:10.1016/j.evolhumbehav.2005.10.002.
- Rossano, M. J. (2003). *Evolutionary psychology: The science of human behaviour and evolution*. New York: Wiley.
- 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.
- Scott, I. M. L., Clark, A. P., Boothroyd, L. B., & Penton-Voak, I. S. (2013). Do men's faces really signal heritable immunocompetence? *Behavioral Ecology*, 24, 579–589. doi:10.1093/beheco/ars092.
- 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.
- Wood, W., & Joshi, P. D. (2011, October). A meta-analysis of women's mate preferences across the menstrual cycle. Paper presented at the Annual Conference of the Society of Experimental Social Psychology, Washington, DC.
- Wood, W., Kressel, L., Joshi, P. D., Louie, B. (2012a, January). Empirical findings of menstrual cycle effects on mate preferences: A meta-analytic review. Paper presented at the Annual Conference of the Society for Personality and Social Psychology, San Diego, California.
- Wood, W., Kressel, L., Joshi, P. D., Louie, B. (2012b, March). Meta-analytic review of menstrual cycle effects on mate preferences. Paper presented at the Annual Conference of the Association for Psychological Science, Chicago, Illinois.
- Wood, W., Kressel, L., Joshi, P. D., Louie, B. (in press). Meta-analysis of menstrual cycle effects on women's mate preferences. *Emotion Review*.