Men Still Value Physical Attractiveness in a Long-Term Mate More Than Women: Rejoinder to Eastwick, Neff, Finkel, Luchies, and Hunt (2014)

Andrea L. Meltzer
Southern Methodist University

James K. McNulty
Florida State University

Grace L. Jackson and Benjamin R. Karney
University of California, Los Angeles

Sexual selection theory and parental investment theory suggest that partner physical attractiveness should more strongly affect men’s relationship outcomes than women’s relationship outcomes. Nevertheless, the contextual nature of this prediction makes serious methodological demands on studies designed to evaluate it. Given these theories suggest that men value observable aspects of partner attractiveness more than women do only in the context of long-term and reproductively viable relationships, they require that studies testing this sex difference involve (a) participants in long-term relationships, (b) women of child-bearing age, and (c) measures of physical attractiveness that assess observable aspects of appearance. In our original article (Meltzer, McNulty, Jackson, & Karney, 2014), we applied 7 methodological standards that allowed us to meet these 3 criteria and demonstrated that partner physical attractiveness is more strongly associated with men’s long-term relationship satisfaction than women’s long-term relationship satisfaction. Eastwick, Neff, Finkel, Luchies, and Hunt (2014), in contrast, described an unfocused meta-analysis, a refocused meta-analysis, and new data that all failed to meet these criteria and, not surprisingly, failed to demonstrate such a sex difference. We continue to believe that men value physical attractiveness more than women do, that such preferences have implications for their evaluations of long-term relationships, and that studies properly calibrated to detect such differences will do so.

Keywords: sex differences, physical attractiveness, evolutionary psychology, mate preferences, meta-analysis

In response to our article demonstrating that partner physical attractiveness is more strongly associated with marital satisfaction among men than women (Meltzer, McNulty, Jackson, & Karney, 2014), Eastwick, Neff, Finkel, Luchies, and Hunt (2014) stated that “the field is best served by focusing on the larger foundation of data built by multiple laboratories rather than any one constituent brick” (p. 429). We could not agree more.

The specific hypothesis that men and women are differentially affected by partner physical attractiveness is derived from two theories that have been supported by thousands of studies across dozens of disciples observing hundreds of species: sexual selection theory (Darwin, 1871) and parental investment theory (Trivers, 1972). According to sexual selection theory, the mate preferences and mating behaviors that characterize different species evolved as adaptations to various reproductive challenges. One reproductive challenge faced by most sexual species was the need to identify mates capable of successful reproduction. Human males faced the challenge of identifying fertile partners. Although female fertility is not directly observable, it is steeply age graded, such that younger women have much higher reproductive value than older women. Thus, males resolved the challenge of identifying fertile partners by evolving a preference for observable aspects of female physical appearance reflective of youth (e.g., full lips, smooth skin, lustrous hair). Although human females also evolved to prefer observable aspects of male physical appearance that are indicators of men’s reproductive capacity (e.g., facial symmetry, strong jaw, broad chin), the extent of this preference varies across short- versus long-term mating contexts (see Buss &
Schmidt, 1993). Whereas women pursuing a short-term mating strategy would have benefited to the extent that they also prioritized such physically attractive features, women pursuing a long-term mating strategy needed to overcome an additional reproductive challenge. Specifically, because human females are required to invest more time and energy into bearing and rearing their offspring than are human males, women pursuing long-term relationships needed to identify partners who were willing and able to provide resources to support child-rearing. According to parental investment theory (Trivers, 1972), they resolved this challenge by evolving preferences for male characteristics and behaviors indicative of the ability and willingness to provide such resources. Given that physically attractive men tended to be less willing to invest resources in long-term relationships, this preference necessarily involved a trade-off that weakened women’s preference for physical attractiveness in a long-term partner relative to men’s (see Gangestad & Simpson, 2000).

A large body of evidence from multiple laboratories across the globe supports this prediction (for a review, see Schmitt et al., 2012). For example, Buss (1989) demonstrated that such sex differences emerge in data collected in 37 different cultures by mostly native researchers blind to study hypotheses. Further, Feingold (1990, 1992) conducted two meta-analyses based on these and other data from independent scholars, and both demonstrated that men value the physical attractiveness of a long-term partner more than women.

Of course, Eastwick et al. (2014) acknowledged that there is ample evidence that these sex differences emerge in people’s reports of their preferences for a mate. They raise questions, however, regarding the extent to which such differences emerge in evaluations of ongoing relationships. Given that a robust body of theoretical and empirical work indicates that partner preferences predict relationship evaluations (e.g., Campbell, Simpson, Kashy, & Fletcher, 2001; Fletcher, Simpson, & Thomas, 2000; Fletcher, Simpson, Thomas, & Giles, 1999; Thibaut & Kelley, 1959), existing work strongly suggests that partner attractiveness should more strongly predict men’s relationship evaluations than it predicts women’s relationship evaluations. Nevertheless, Eastwick, Luchies, Finkel, and Hunt (2013), prompted at least partially by a null finding that emerged in a speed-dating study conducted by their laboratory (Eastwick & Finkel, 2008), challenged this idea. Specifically, they conducted a meta-analysis of 97 studies that reveals no sex differences in the implications of partner attractiveness for relationship outcomes. Given the large body of theoretical and empirical evidence in favor of such sex differences, however, such a challenge must be solidly constructed.

As we indicated in our original article (Meltzer et al., 2014), we believe that Eastwick, Luchies, et al.’s (2013) meta-analysis is not solidly constructed. Specifically, we stated that the strongest test of sex differences in the implications of partner physical attractiveness for relationship evaluations would meet a number of methodological standards that Eastwick et al. (2014) challenged on theoretical and empirical grounds. We first review the theoretical rationale for these standards and then address why Eastwick, Hunt, and Neff (2013) found no empirical support for them in their meta-analysis.

Criteria for Testing Whether Partner Physical Attractiveness Differentially Affects Men and Women

The methodological standards we recommended were meant to help researchers meet three crucial criteria for examining sex differences in the implications of partner physical attractiveness for relationship outcomes. First, the appropriate population within which to study these issues is long-term relationships. According to sexual strategies theory (Buss & Schmitt, 1993), males and females can shift between using a short-term or a long-term mating strategy, and men and women do so according to the advantages afforded by their current context. As noted earlier, women who adopt a long-term mating strategy demonstrate a weakened preference for physically attractive partners because physically attractive men are less likely to invest the resources necessary for successfully bearing and rearing offspring. But men’s willingness to provide resources is irrelevant in a short-term mating context. Without the constraint of identifying a partner willing to invest resources, women pursuing short-term mates demonstrate preferences for a physically attractive partner that are very similar to men’s preferences. Thus, sex differences in preferences for partner physical attractiveness should be most likely to emerge in the context of long-term relationships compared with short-term relationships.

Evidence from multiple laboratories supports this prediction. For example, partner physical attractiveness is equally important to men and women seeking short-term relationship partners (Gangestad & Simpson, 2000; Gangestad & Thornhill, 1997; Kenrick, Groth, Trost, & Sadalla, 1993; Kenrick, Sadalla, Groth, & Trost, 1990; Li & Kenrick, 2006). Moreover, direct tests demonstrate that whether men and women are focused on short-term versus long-term relationships moderates the size of the sex differences in their preferences for an attractive partner (Kenrick et al., 1990). Specifically, when male and female college students were asked to rate the importance of various characteristics of a potential partner at four levels of commitment (i.e., a single date, sexual relations, steady dating, and marriage), sex differences in the importance of partner physical attractiveness only emerged in the ultimate long-term relationship—marriage. Given that there are no sex differences in preferences for a physically attractive partner in short-term relationships, and no theoretical reason even to expect such sex differences, there is no reason to expect sex differences to emerge in the implications of partner physical attractiveness for evaluations of short-term relationships.

It is worth noting that Eastwick et al. (2014) claimed that we argued “only married” couples offer a relevant test of this sex difference (p. 431) and then went on to argue why such a standard is not justified. In fact, we never suggested that married couples are the only relevant population for this research; we stated that the strongest test of sex differences in the implications of partner attractiveness would “use samples of young couples involved in long-term relationships” (p. 420). We only raised the issue of married samples to address the Eastwick, Luchies, et al. (2013) meta-analysis that relied heavily on samples of dating couples. Although dating couples in long-term relationships would certainly provide a fair test of sex differences, the problem with most existing studies of dating couples is that it is unclear whether such couples are pursuing long-term or short-term relationships. Although it is likely that some couples in such studies may be
pursuing long-term relationships, many of them are likely to be pursuing short-term relationships. Accordingly, the null sex difference that should emerge in short-term relationships will “wash out” any significant sex differences that emerge in the implications of physical attractiveness among those couples in long-term relationships. Such a problem does not emerge in studies of married couples, which makes them ideal, but not essential.

The second criterion for studies testing sex differences in the implications of partner physical attractiveness for relationship outcomes is that they examine younger couples. Sexual selection and parental investment theories suggest preferences that should have been adaptive for successful reproduction. Given that women’s reproductive potential significantly decreases after the age of 35 (Rothman et al., 2013), these theories may not apply to women older than 35. If so, sex differences in the implications of partner physical attractiveness may be less pronounced in older couples. Indeed, not only have several studies of older couples failed to document the sex difference that is typically observed in studies of younger couples, some studies have documented trends in the opposite direction (e.g., Barelæs & Dijkstra, 2009; Eastwick, Finkel, & Eagly, 2011; Lucas et al., 2004; Murstein & Christy, 1976). Although such findings are certainly interesting and deserve explanation, they do not challenge theories meant to apply to reproductive behaviors. Indeed, in older couples, physical appearance may be less strongly associated with partner qualities that are differentially preferred across the sexes, such as reproductive capacity, and more strongly associated with partner qualities that are equally important to both sexes, such as health and vitality.

The third criterion for studies of sex differences in the implications of partner physical attractiveness is that they use measures of physical attractiveness that are valid indicators of physical appearance. Both sexual selection and parental investment theories strongly suggest that sex differences should emerge in reaction to observable indicators of physical attractiveness, such as clear skin, symmetry, feminine features (for women), and masculine features (for men). But partners are motivated to see themselves and each other as more attractive than they actually are, and such motivations can lead to perceptions of own and partner attractiveness that are independent of observable features of attractiveness (e.g., Epley & Whitcomb, 2008; Montoya, 2008). In one interesting documentation of such effects, Epley and Whitcomb (2008) presented participants with photographs of themselves, a close other, and a stranger, and then asked them to identify the unaltered image of each target from among other images that had been altered to appear more and less physically attractive. Participants were more likely to incorrectly identify the attractively enhanced photograph of themselves and the close other, but not the stranger. Given that evolutionary-based predictions regarding physical attractiveness are based on the idea that observable qualities of men and women’s appearance were linked to their genetic fitness and fertility, rather than such subjective self- and partner-perceptions of attractiveness, sexual selection and parental investment theories suggest that men and women should have evolved to respond differentially to observable features of partner physical attractiveness, not to their own (or their partners’) perceptions of partner attractiveness. Accordingly, it is crucial that studies testing for such sex differences obtain measures of physical attractiveness that isolate the effects of objective features of partner physical appearance, unbiased by other variables that can influence attractiveness ratings.

This third criterion explains why we also controlled for related confounds in our original study (e.g., age, income, extraversion, own attractiveness) and why we argued it would be ideal if other studies did the same. Given that parental investment theory suggests that women should prioritize partner income, age, and perhaps even extraversion but not necessarily partner attractiveness, a woman may select a mate who is physically attractive, socially skilled, and wealthy not because he is physically attractive, but because he is socially skilled and/or wealthy. Likewise, partners’ levels of physical attractiveness within a relationship tend to be positively correlated (Feingold, 1988), and one’s own level of attractiveness is associated with numerous factors that influence relationship outcomes. For example, attractive men are less satisfied with their relationships, on average (McNulty, Neff, & Karney, 2008). Without partialing out the variables associated with own and partner attractiveness, it is impossible to determine whether partner physical appearance or these other factors (e.g., age, income, extraversion, own attractiveness) are responsible for relationship satisfaction.

These three criteria subsume most of the methodological standards we set forth in our original article, as well as those Eastwick et al. (2014) attributed to us. As can be seen in Table 1, the first two criteria align perfectly with our original standards and those outlined by Eastwick et al., with the exception that we do not specify that participants be married. The third criterion subsumes three more of our original standards and three of those that Eastwick et al. attributed to us. This leaves two additional standards we originally proposed: that studies of sex differences in the implications of physical attractiveness for relationship outcomes (a) involve measures of relationship satisfaction and (b) longitudinal data. Although these standards are not essential, they continue to seem ideal. Evolutionary perspectives suggest relationship satisfaction may have evolved to gauge the extent to which people’s preferences are being met (see Shackelford & Buss, 1997), and it is possible that such preferences may affect either absolute levels of or changes in satisfaction over time (see Karney & Bradbury, 1997).

**Foundation of Bricks, or House of Cards?**

How does Eastwick, Luchies, et al.’s (2013) meta-analysis stand up to these three criteria? Not well. A total of 79 of their 97 studies did not meet the first criterion that couples be involved in long-term relationships. Of these 79 studies, 56 were at least partially comprised of individuals that were in speed-dating sessions or dating relationships, and 23 were comprised of individuals not even involved in relationships (e.g., studies that utilized zero acquaintance or confederate paradigms). Although it is likely that some of the individuals in the studies involving actual relationships were involved in long-term relationships, there is no way to differentiate the effects of these individuals from effects of the individuals involved in purely short-term relationships.

A total of 10 of the remaining 18 studies (13 of the original 97 studies) did not clearly meet our second criterion that the relationships be comprised of younger couples. As noted earlier, women’s fecundability significantly decreases after the age of 35 (Rothman et al., 2013), and all 10 of these studies involved couples in which the wives were over the age of 35 on average.
Sex-Differences Hypothesis With Those Set Forth by Meltzer et al. (2014) and Eastwick et al. (2014)

Table 1
Comparison of the Three Criteria Suggested by Sexual Selection and Parental Investment Theories Required to Properly Examine the Sex-Differences Hypothesis With Those Set Forth by Meltzer et al. (2014) and Eastwick et al. (2014)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Participants must be in a long-term relationship.</td>
<td>1. Participants must be in a long-term relationship.</td>
<td>1. Participants must be in long-term relationships (preferably marriage).</td>
</tr>
<tr>
<td>2. Participants must be young.</td>
<td>2. Participants must be young.</td>
<td>2. Participants must be young.</td>
</tr>
<tr>
<td>3. Measures of physical attractiveness must exclusively assess observable aspects of appearance.</td>
<td>3. Measures of partner physical attractiveness must be based on objective ratings.</td>
<td>3. Physical attractiveness must be assessed objectively.</td>
</tr>
<tr>
<td></td>
<td>4. Factors confounded with objective ratings of attractiveness must be controlled.</td>
<td>4. Physical attractiveness ratings must be provided by judges who did not meet the target face-to-face.</td>
</tr>
<tr>
<td></td>
<td>5. Own attractiveness must be controlled.</td>
<td>5. Statistical models must include a variety of control variables (about the self and partner).</td>
</tr>
<tr>
<td>Ideal, nonessential standards</td>
<td>6. Relationship satisfaction should be the outcome measure.</td>
<td>6. Relationship satisfaction must be the (only) dependent variable.</td>
</tr>
<tr>
<td></td>
<td>7. Study should provide longitudinal data.</td>
<td>7. Data must be longitudinal.</td>
</tr>
</tbody>
</table>

Finally, a total of seven of the remaining eight studies (67 of the original 97 studies) did not meet our third criterion that measures of physical attractiveness rely on observable features of physical appearance. Although Eastwick, Luchies, et al. (2013) described 32 of their studies as providing “objective” measures of physical attractiveness, closer examination reveals that in at least two of these studies (i.e., Marks, Huston, Johnson, & MacDermid, 2001; Neff, & Karney, 2005), the experimenter made these “objective” ratings after interacting with the couple. As Eastwick, Hunt, and Neff (2013) themselves argued, live interactions can lead to evaluations of targets that are very different from evaluations that are based on static information, like images, because live interactions provide additional information that gets processed dynamically, such as personality (see Jensen-Campbell, Graziano, & West, 1995). Indeed, our study (2014) relied on five independent judges’ ratings of the attractiveness of the 169 couples described by Neff and Karney (2005), and the experimenter ratings they used in their meta-analysis accounted for only about one third of the variance in our judges’ ratings of wives’ attractiveness and less than 20% of the variance in our judges’ ratings of husbands’ attractiveness.

Even though applying all three of our criteria to the meta-analysis eliminates 96 of their 97 studies, Eastwick et al. (2014) argued that the null effect of their meta-analysis held even if they applied our criteria. They presented two types of analyses to support this argument. First, they reported that the moderation analyses described by Eastwick, Luchies, et al. (2013) revealed no evidence that any of our criteria moderated their null sex difference. It is worth noting that this one study (McNulty et al., 2008) does provide evidence for the predicted sex difference. It is also worth noting that our analysis that documented the predicted sex difference (2014) included the sample used in that study as well as three other samples that also met the three criteria.

The second way Eastwick et al. (2014) challenged the importance of the theoretically derived criteria we outlined was by conducting a “new” meta-analysis of the studies they argued meet these criteria. Specifically, Eastwick et al. (2014) reported that they find no sex difference even when they include our four studies of 916 participants (one of which is the one study in their meta-analysis that met all three criteria) with three studies of 444 participants from their meta-analysis that they claim meet our criteria. But these studies do not meet our criteria. Two of them are studies comprised of couples in which the wives were 35 years or older on average (i.e., Murstein & Christy, 1976; Peterson & Miller, 1980)—one in which the wives were in their late 30s and the other in which the wives were in their 70s. The third study (i.e., Marks et al., 2001) was based on younger couples, but the ratings of attractiveness were made by the experimenter who interacted...
with the couple, and thus, as we argued earlier, likely reflect nonappearance-related information.

Eastwick et al. (2014) argued that the Eastwick, Luchies, et al. (2013) meta-analysis is like a foundation of bricks that should weigh more heavily than one constituent brick, like our four studies. Given that only one of the 97 studies in their meta-analysis met the three criteria necessary for appropriately testing sex differences in the implications of partner physical attractiveness, we contend that their meta-analysis is more accurately described as a house of cards that cannot withstand the weight of even one brick.

Evaluating the New Data Presented by Eastwick and Colleagues

Eastwick et al. (2014) also obtained two new samples of data that they claimed meet the criteria necessary for testing the sex-difference hypothesis. Within these samples, they failed to replicate the sex difference that we observed. Although independent studies can obtain different results for numerous reasons, we believe one reason they failed to replicate our findings is that these two new studies did not actually meet the third criterion that measures of physical attractiveness exclusively assess observable aspects of appearance. Here is how Eastwick et al. (2014) describe their ratings of physical attractiveness: “Judges who had never met the couples rated the attractiveness of both members of the couple by viewing video clips of discussion tasks from an initial intake session” (p. 433). Although this method is indeed better than using partner, self-, or experimenter ratings of attractiveness, there are two ways in which their procedure differed from ours, both of which may have obscured the objectivity of these measures. First, Eastwick et al. did not state that they evaluated partners’ attractiveness independently. Typically, dyadic discussions such as the ones they coded involve a split-screen display of both partners, and research indicates that ratings of individuals’ attractiveness can be confounded with evaluations of close others’ attractiveness (Kernis & Wheeler, 1981). As we noted in our article, when we coded the couples in our split-screen discussions, we were careful to cover up the face of each partner while rating each individual. Second, Eastwick et al. reported that they used “clips” of the discussions. As we have noted several times, and as Eastwick, Hunt, and Neff (2013) themselves stressed, people process information from dynamic interactions differently than the way they process static information. The personal information available from even a “thin slice” of a person’s behavior can have powerful effects on perceptions (Funder, 1987; Kenny, Albright, Malloy, & Kashy, 1994), including ratings of physical attractiveness (Jensen-Campbell et al., 1995).

Conclusions

Based on their old meta-analysis, their new, refocused meta-analysis, and their failed replication of our effects, Eastwick et al. (2014) claimed that they “hesitate to conclude” specific effects because of a “sea of data where the size of this sex difference shifts in small but unpredictable ways” (p. 433). We argue that the data shift in exactly the ways predicted by sexual selection and parental investment theories: when research tests for sex differences using measures of physical attractiveness that assess observable aspects of partner appearance in samples of young couples involved in long-term relationships, as we did using the four samples we describe in our article (Meltzer et al., 2014), the sex difference emerges; when one uses measures of partner physical attractiveness that are confounded with other information that may be more important to women, older couples, and/or couples involved in short-term relationships, the sex difference disappears. Nevertheless, we also do not want to rely simply on our four samples to make broad generalizations. Thus, we hope our work inspires others to examine the sex-differences hypothesis directly using samples that meet all theoretically driven criteria, as well as to devote greater attention to the internal validity of the studies involved in meta-analyses.

References


Received November 5, 2013
Revision received November 5, 2013
Accepted November 7, 2013

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.