

Comment: Why Meta-Analyses Rarely Resolve Ideological Debates

Christopher J. Ferguson

Department of Psychology, Stetson University, USA

Abstract

In their meta-analysis Wood, Kressel, Joshi, and Louie (2014) argue little evidence supports shifts in mating preferences across the menstrual cycle. They imply this may represent a critical weakness of evolutionary psychology theories of mating preferences more generally. This report represents a fairly common use of meta-analysis: to assemble data to support or reject a particular proposition over which there is debate. Yet, rarely do meta-analyses succeed at resolving ideological debates. Multiple decision points related to the selection, coding, effect size extraction, and interpretation of studies leaves considerable room for meta-analytic authors to interject their own beliefs. Meta-analyses are typically hailed by those who agreed a priori with their conclusion, and rejected as fatally flawed by those in disagreement. As such, meta-analyses have failed in replacing narrative reviews as more objective.

Keywords

evolutionary psychology, mating preferences, meta-analysis

In their meta-analytic review Wood et al. (2014), argue research evidence refutes that women's mating preferences shift across the menstrual cycle. Further, the authors interpret this as a problematic finding for evolutionary psychology more generally. The meta-analysis by Wood et al. appears to be a competently executed analysis. I will note upfront several things. First, my a priori position on evolution is that evolution is a powerful determinant of our behavior. Second, I served as a reviewer for this meta-analysis and recommended its publication. Third, although the meta-analysis by Wood et al. appears to be well done, I remain skeptical it will resolve academic debates related to evolutionary psychology.

The purpose behind the development of meta-analysis was to replace the subjective nature of narrative reviews which often reflected a researcher's individual biases (Quintana & Minami, 2006). Narrative reviews often were used to advance particular ideological positions and thus competing narrative reviews

typically reflected the individual scholar rather than the actual data. By quantifying the data available across a field of study, it was thought, it would be possible to eliminate the subjectivity of narrative reviews.

Having both conducted meta-analyses and watched debates in other research areas, I struggle to think of an ideological debate that was resolved through the use of meta-analysis. Typically, proponents of the opposing view simply reject a meta-analysis out of hand as fatally flawed. Just a small subset of examples includes debates over the validity of fMRI studies (Vul, Harris, Winkielman, & Pashler, 2009), an alleged narcissism epidemic (Trzesniewski, Donnellan, & Robins, 2008), rejection, frustration and numbness (Gerber & Wheeler, 2009) or even childhood sexual abuse (Lilienfeld, 2002). If researchers truly accepted meta-analyses as objective summaries of a research field, meta-analyses would have more impact in shifting debates than they do.

It may be that meta-analyses are not capable of answering big questions due to issues of methodological flexibility which introduce exactly the type of subjectivity they were designed to avoid. Consider the area of video game violence which experienced a "war of the metas" in which both sides of the debate have produced meta-analyses supporting and critiquing beliefs in harmful effects (e.g., Anderson et al., 2010; Sherry, 2007). Examining the most recent, we can see why meta-analyses do little to resolve academic debates. Several coauthors of the Anderson et al. (2010) meta-analysis are strong advocates for wide searches for unpublished studies (Rothstein & Bushman, 2012). Despite this, one of the coauthors recently acknowledged they made no such effort with their own meta-analysis (Bushman & Huesmann, in press). Instead, the authors only included their own unpublished studies and those of close colleagues, introducing selection bias. This demonstrates how authors can stack a meta-analysis toward a particular outcome.

Part of the problem in using meta-analyses to resolve scholarly debates is misuse of an "average effect size wins" approach

to meta-analysis which has no basis in science. Imagine, 10 studies test the hypothesis that eating rhubarbs causes depression. All of the studies are direct replications of equal quality. Five of the studies find that rhubarbs are associated with depression at the $r = .3$ level. The other five studies are exactly null $r = .00$. A heated rhubarb debate ensues. Advocates of the “rhubarbs are bad” theory conduct a meta-analysis and find that the average rhubarb effect on depression is a statistically significant $r = .15$ and declare victory. This is entirely absurd as the meta-analysis, in effect, washes away a 50% failed replication rate we should be very curious about. And this issue ignores the widespread publication bias of our field (Fanelli, 2010). This is further exacerbated in meta-analyses of correlational data in which spuriously high bivariate correlations are preferred over theoretically superior partial correlations or standardized regression coefficients in the name of statistical purity (Pratt & Cullen, 2000). In this sense, meta-analyses have their own bias and that is to favor the maintenance of hypotheses over their falsification. Many meta-analyses place a reverse burden of proof on null studies or failed replications by creating a scenario in which only a massive number of failed replications would warrant theory rejection. This use of meta-analysis is inherently unscientific.

It is not my intent to rail entirely against meta-analysis. Used properly and with care they can be informative, telling us about methodological and population moderators that can influence research results. But they seldom resolve academic debates. Returning to the issue of the meta-analysis by Wood et al. (2014), I suspect it has accomplished the task of putting the ball in the court of those who would argue that women’s mating preferences vary across the menstrual cycle. But I suspect the debate is not over yet. Certainly, to imply that this is a critical challenge for all of evolutionary psychology is probably unwarranted. Nonetheless, I suspect it will be good theater.

References

- Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., & . . . Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in Eastern and Western countries. *Psychological Bulletin, 136*(2), 151–173. doi:10.1037/a0018251
- Bushman, B., & Huesmann, L. R. (in press). Twenty-five years of research on violence in digital games and aggression: A reply to Elson & Ferguson (2013). *European Psychologist*.
- Fanelli, D. (2010). “Positive” results increase down the hierarchy of the sciences. *PLoS ONE, 5*(4), e10068.
- Gerber, J., & Wheeler, L. (2009). On being rejected: A meta-analysis of experimental research on rejection. *Perspectives on Psychological Science, 4*(5), 468–488. doi:10.1111/j.1745-6924.2009.01158.x
- Lilienfeld, S. O. (2002). When worlds collide: Social science, politics, and the Rind et al. (1998) child sexual abuse meta-analysis. *American Psychologist, 57*(3), 176–188. doi:10.1037/0003-066X.57.3.176
- Pratt, T., & Cullen, C. (2000). The empirical status of Gottfredson and Hirschi’s general theory of crime: A meta-analysis. *Criminology, 38*, 931–964.
- Quintana, S., & Minami, T. (2006). Guidelines for meta-analyses of counseling psychology research. *The Counseling Psychologist, 34*(6), 839–877.
- Rothstein, H. R., & Bushman, B. J. (2012). Publication bias in psychological science: Comment on Ferguson and Brannick (2012). *Psychological Methods, 17*(1), 129–136. doi:10.1037/a0027128
- Sherry, J. (2007). Violent video games and aggression: Why can’t we find links? In R. Preiss, B. Gayle, N. Burrell, M. Allen & J. Bryant (Eds.), *Mass media effects research: Advances through meta-analysis* (pp. 231–248). Mahwah, NJ: L. Erlbaum.
- Trzesniewski, K. H., Donnellan, M., & Robins, R. W. (2008). Do today’s young people really think they are so extraordinary? An examination of secular trends in narcissism and self-enhancement. *Psychological Science, 19*(2), 181–188. doi:10.1111/j.1467-9280.2008.02065.x
- Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science, 4*(3), 274–290. doi:10.1111/j.1745-6924.2009.01125.x
- Wood, W., Kressel, L., Joshi, P., & Louie, B. (2014). Meta-analysis of menstrual cycle effects on women’s mate preferences. *Emotion Review, 6*(3), xxx–xxx.