Journal of Marriage and Family

Volume 76, Issue 1
February 2014
Pages 105–129

Original Article

Gender Differences in Marital Satisfaction: A Meta-analysis

Jeffrey B. Jackson , Richard B. Miller, Megan Oka, Ryan G. Henry

First published:
13 January 2014 Full publication history

DOI:
10.1111/jomf.12077 View/save citation

Cited by:
4 articles Refresh Citing literature

Abstract

The purpose of the present meta-analysis was to empirically test the widely held assumption that women experience lower marital satisfaction than men. A total of 226 independent samples with a combined sum of 101,110 participants were included in the meta-analysis. Overall results indicated statistically significant yet very small gender differences in marital satisfaction between wives and husbands, with wives slightly less satisfied than husbands; moderator analyses, however, indicated that this difference was due to the inclusion of clinical samples, with wives in marital therapy 51% less likely to be satisfied with their marital relationship than their husbands. The effect size for nonclinical community-based samples indicated no significant gender differences among couples in the general population. Additional moderator analyses indicated that there were also no gender differences when the levels of marital satisfaction of husbands and wives in the same relationship

(i.e., dyadic data) were compared .

Enhanced Article Feedback

Marital scholars generally acknowledge that men and women experience marriage differently. In 1972, prominent family scholar Jesse Bernard famously stated, “There are two marriages in every marital union, his and hers. And his … is better than hers” (p. 14). She elaborated by stating,

> There is a very considerable research literature … which shows that: more wives than husbands report marital frustration and dissatisfaction; more report negative feelings; … more wives than husbands consider their marriages unhappy, and have considered separation or divorce. (pp. 26–27)

On the basis of Bernard's compelling argument, family scholars have assumed that women consistently experience significantly less marital satisfaction than men (Connides, **2001**). For instance, one group of scholars reported that “women consistently report lower marital quality than men in national surveys” (Umberson, Williams, Powers, Liu, & Needham, **2006** , p. 3).

Indeed, many studies have found that wives' reports of marital satisfaction are significantly lower than husbands' (Kamp Dush, Taylor, & Kroeger, **2008** ; Myers & Booth, **1999** ; Stevenson & Wolfers, **2009** ; Whiteman, McHale, & Crouter, **2007**). For example, national surveys of married adults in the United States in 1980 and 2000 found that, on average, women reported lower levels of marital quality (Amato, Booth, Johnson, & Rogers, **2007**). Other studies, though, have found no gender differences (Broman, **2005** ; Kurdek, **2005**). For instance, using national probability data from the National Study of Families and Households, Gager and Sanchez (**2003**) found no significant differences in the mean levels of husbands' and wives' marital satisfaction. Moreover, when they compared paired husbands' and wives' scores of marital happiness, they found that a strong majority of husbands and wives had the same score. Of those with different scores, 11.0% of the couples consisted of a very satisfied husband and an unsatisfied wife, whereas 9.9% of the couples consisted of a very satisfied wife and an unsatisfied husband. Likewise, Broman (**2005**) analyzed data from the American Changing Lives survey, a large national data set of married individuals, and found no statistically significant gender difference in marital quality.

In addition, there is a question as to the magnitude of the difference between men's and women's levels of marital satisfaction among those studies that found a significant difference. For example, Whiteman et al. (**2007**) found that women's mean level of marital satisfaction was lower than men's at a statistically significant level; however, the actual difference was very modest. Using a scale of marital quality that ranged from 7 to 63, women's mean score was only 1.38 points lower than men's. Likewise, after reporting that women reported lower levels of marital quality than men in their national studies, Amato et al. (**2007**) qualified their findings by stating that the differences were small.

In the context of these discrepancies in findings as well as questions about the magnitude of potential

differences, there is a need for a comprehensive quantitative review of literature to examine gender differences in marital satisfaction. The objective of this study was to use meta-analytic methods to analyze gender differences in marital satisfaction.

Literature Review

Theoretical Perspectives

Bernard's (**1972**) influential dictum was based on the feminist perspective that marriage is oppressive to women, which is isomorphic of a larger societal environment of male privilege. The devalued and subordinated position of women in society and families is a central theme of feminist theory (Osmond & Thorne, **1993**): “Feminists agree that male dominance within families is part of a wider system of male power, is neither natural nor inevitable, and occurs at women's cost” (Ferree, **1990** , p. 866). Women's subordinate role in marriage is represented by unequal control of family money, higher risk for interpersonal violence, and double standards in regard to sexual behavior (Finlay & Clarke, **2003** ; Walker & Thompson, **1995**). Because an unequal balance of power is associated with lower marital satisfaction (Ball, Cowan, & Cowan, **1995** ; Brezsnyak & Whisman, **2004** ; Gray-Little, Baucom, & Hamby, **1996**), it follows that women would likely experience less satisfaction than men.

One of Bernard's (**1972**) central arguments was that men benefit more than women from marriage, with husbands having significantly better health than wives. In support of her argument, there is empirical evidence that men derive more health benefits than women from being married (Waite, **1995**). A nationally representative longitudinal study found that the transition to first marriage for men was associated with a 27.9% increase in the probability of being in excellent or very good health, compared to a 4.8% increase for women (Williams & Umberson, **2004**). Likewise, men's transition to remarriage was associated with a greater health benefit than for women.

Men also benefit more than women from marriage because women shoulder the majority of child care and housework (Bernard, **1972**). Indeed, the inequitable division of household labor and child care has been a central focus of feminist theory and research (Ferree, **2010** ; Osmond & Thorne, **1993**). Since the Industrial Revolution, which moved husbands from working on family farms and businesses with their wives to working at factories and outside businesses, women have had primary responsibility of household duties. Although research suggests that husbands' level of participation in household tasks has increased (Sayer, **2005**), wives still perform a disproportionate amount of household tasks (Baxter, **2000**) and child care (Bianchi & Milkie, **2010**). Evidence suggests that husbands resist their wives' efforts to more equally distribute child care and housework and that wives are generally unhappy with the division of labor in their relationships (Dempsey, **2000**). This lack of equitable division of labor is associated with lower marital satisfaction (Grote & Clarke, 2001; Stevens, Kiger, & Mannon, **2005**).

In addition to having primary child care responsibility, wives also provide an inequitable amount of emotion work by supporting their husbands and managing the emotional climate of the relationship (Loscocco & Walzer, **2013**). Research has found that emotion work is more closely tied to gender than are child care and housework (Erickson, **2005**). Wives are generally more aware of the emotional climate of the relationship (Croyle & Waltz, **2002**), and they are more likely to monitor the

relationship's emotional quality (Loscocco & Walzer, 2013). The result of this discrepancy is that perceptions of imbalance in emotion work are associated with women's lower marital satisfaction (Croyle & Waltz, 2002 ; Duncombe & Marsden, 1995 ; Strazdins & Broom, 2004).

Possible Moderators of Wife–Husband Differences

Although there is inconsistency among studies regarding gender differences in marital quality (Broman, 2005 ; Umberson et al., 2006), there is clear evidence that wives are more likely to initiate marital therapy (Vessey & Howard, 1993), as much as 73.2% of the time (Doss, Atkins, & Christensen, 2003). In addition, wives generally report more relationship problems among those couples who are attending marital therapy (Miller, Yorgason, Sandberg, & White, 2003). Moreover, women are more likely than men to initiate divorce proceedings (Montenegro, 2004 ; Rokach, Cohen, & Dreman, 2004), regardless of whether initiators cite their own characteristics or their partner's characteristics as the reason for the divorce (Hewitt, Western, & Baxter, 2006). Thus, it is important to differentiate studies that examine clinical samples from those that examine community samples.

The extant literature indicates that there has been a substantial shift in the distribution of power in marital relationships over the last few decades. Amato et al.'s (2007) national study of U.S. marriages revealed a significant decrease in patriarchal marriages from 1980 to 2000. Specifically, the authors found that the percentage of wives who reported that they shared an equal part in decision making increased from 47% in 1980 to 64% in 2000. In addition, among wives who reported having an unequal part in decision making in 2000, the majority of them reported that *they* most often had the final say. Therefore, it is important to explore data collection time frames to investigate potential historical and cohort effects in marital satisfaction gender differences.

Another relevant issue regarding gender differences in marital satisfaction is whether the data are dyadic or nondyadic. Many researchers who have examined marital satisfaction have used study designs in which only one partner in a dyad responded; that is, the men and women who participated were not married to each other. For example, national studies of marital relationships conducted in 1980 and 2000 (Amato et al., 2007) used a research design in which either the husband or the wife in each household was randomly selected to respond to the phone survey. Conversely, other studies, such as the National Study of Families and Households, consist of dyadic data that comprise the linked responses of both marital partners to enable within-couple comparison (Gager & Sanchez, 2003). The fact that the differences between men's and women's responses were significant in research with nonrelated married adults (Amato et al., 2007) and nonsignificant in research that analyzed dyadic data (Gager & Sanchez, 2003) raises questions as to the effects of dyadic and nondyadic data on the determination of gender differences.

Some demographic characteristics may moderate gender differences in marital satisfaction. Although no studies have directly examined the effects of demographic characteristics on gender differences in marital satisfaction, there is evidence that marital power and division of household labor differ by socioeconomic status (e.g., education, income; Amato et al., 2007), nationality (Davis & Greenstein, 2004 ; Geist & Cohen, 2011 ; Warner, Lee, & Lee, 1986), and race (Amato et al., 2007). Also, there is evidence that division of labor and child care responsibilities change over the course of marriage, as children enter and leave the home and as the partners retire (Miller & Yorgason, 2009). According to feminist theory (Bernard, 1972) and evidence from empirical research (Grote & Clarke,

2001; Stevens et al., 2005), these variations in marital power and division of household labor that are associated with demographic factors may affect gender differences in marital satisfaction.

In the context of these previous studies, a comprehensive quantitative review of the literature is needed so that credible conclusions (Shadish & Baldwin, 2003) about gender differences in marital satisfaction can be made. The use of meta-analysis in marriage research (Wampler, Reifman, & Serovich, 2005) and gender difference research (Hyde, 2005) has been encouraged. We used meta-analytic statistical techniques (Borenstein, Hedges, Higgins, & Rothstein, 2009) to examine gender differences in marital satisfaction to determine the magnitude, precision, and significance of gender differences across all included samples. Specifically, our objective was to answer the following three questions. First, do gender differences in marital satisfaction exist between husbands and wives? Second, if gender differences in marital satisfaction exist, what is the magnitude of these differences? Third and finally, what are the effects of moderating variables on gender differences in marital satisfaction?

Method

Inclusion and Exclusion Criteria

Inclusion criteria

We designed the selection criteria to include as many samples as possible in an effort to minimize the limitations associated with excluding relevant effect sizes in meta-analysis (i.e., increased sampling error, increased result bias, decreased generalizability of results; R. A. Peterson & Brown, 2005). Accordingly, we established only two sample inclusion criteria: Sample reports had to contain (a) a measure of marital relationship satisfaction (e.g., adjustment, dissatisfaction, satisfaction) and (b) statistics necessary to calculate sample estimate standardized mean difference effect sizes (*d*) for marital satisfaction by gender.

Exclusion criteria

We excluded samples for which the reports were completed before 1970 because pre-1970s relationships may vary significantly from post-1970 relationships: Marital relationships have undergone significant changes since the 1960s, such as increased gender role flexibility and fluidity, increased acceptance of cohabitation, increased acceptance of divorce on grounds of unhappiness, decreased social stigma of people who are not married, and decreased pressure to have children (F. D. Cox, 2006). We also excluded samples for which the reports were in languages other than English, if English translations were not readily available. In addition, we excluded samples with mixed participant marital status (i.e., some participants were married and some participants were not married) if the nonmarried participants (e.g., cohabiting) constituted 50% or more of the sample; we decided to include samples with a minority of nonmarried participants in an effort to (a) maximize inclusivity with regard to effect sizes from extant research and (b) determine whether gender differences in marital satisfaction varied systematically among samples with married participants only and samples with both married and nonmarried participants. Given that the purpose of this study was to identify gender differences in marital satisfaction, we excluded samples of participants in same-sex marriages or

relationships.

Search Strategies

In order to conduct a comprehensive review of the gender and marital satisfaction literature, we implemented multiple search methods. We limited our search to reports from 1970 through the first few months of 2009. The following electronic citation and reference databases were searched using the key words *marital quality*, *marital satisfaction*, and *marital adjustment* and *gender*: Academic Search Premier, Dissertation Abstracts International, Google Scholar, PsycARTICLES, PsycBOOKS, PsycCRITIQES, PsycEXTRA, PsycINFO, Socfile, SocINDEX, and Social Sciences Citation Index. We also searched each issue of relevant major journals from 1970 to the first few months of 2009. In addition, separate searches were conducted on prominent researchers in the field of marital relationships. Furthermore, we used references cited in reports we evaluated for inclusion to locate additional potentially relevant articles.

After we had identified and obtained the sample reports, we evaluated them in terms of our inclusion and exclusion criteria. We excluded a substantial number of samples because the associated reports did not contain separate marital satisfaction statistics for husbands and wives.

Coding Procedures

We used meta-analytic coding procedures (Hunter & Schmidt, **2004**) to manage and prepare data from included samples for analysis. We developed a 43-item codebook (available on request) to facilitate the systematic coding of the samples. Two primary types of data were coded: (a) statistical information regarding gender differences in marital satisfaction and (b) relevant moderator variables.

The coders consisted of the first author, one marriage and family therapy graduate student, and three upper division social science undergraduate students. The student coders were trained by the first author. Two mixed-gender coding pairs were formed by assigning each student a permanent coding partner. The coders met in pairs to reason through the most appropriate way to code each item in the codebook for each included sample. Instances in which the coders were uncertain how to code for an item were resolved through consultation with the first author. The coders systematically created notes to document situations that required consultation and rationales for resolving those situations. Therefore, intercoder reliability was not calculated because coder consensus was used to identify the most appropriate coding decision for each codebook item (Hawkins, Blanchard, Baldwin, & Fawcett, **2008**). We did not identify any statistically significant differences between the samples coded by the first mixed-gender coding pair and the samples coded by the second mixed-gender coding pair. Pursuant to the completion of coding by the mixed-gender coding pairs, the first author and an opposite-gender doctoral student conjointly reviewed the coding and made corrections where necessary.

We managed missing data in several ways. In an effort to maximize sample inclusion, we used methods for imputing statistics necessary for computing *d* (Lipsey & Wilson, **2001**). In an attempt to remediate situations in which statistical information could not be imputed for recently completed sample reports, we contacted authors to request statistics and clarifying information for moderator variables. In total, we contacted 11 researchers, with a 100% response rate. We managed situations in

which samples lacked information for specific moderator variables by coding the relevant moderator variables as *not reported*, allowing for analyses to determine whether the samples without information were statistically different from samples with information.

Moderator Variables

Sample report variables

We coded sample reports according to publication status (published or unpublished), type (journal article, book or chapter in an edited book, doctoral dissertation, master's thesis, other), and year (publication year for published reports or completion year for unpublished reports; we coded both interval data and nominal data by decade). We did not code samples for quality or design features.

Data variables

We coded samples according to whether they contained dyadic data or nondyadic data. We also coded samples according to data collection year (we coded both interval data and nominal data by decade).

Descriptive participant variables

Marital status

We coded samples as *married* if all sample participants were married and as *mixed* if some sample participants were married and some were not married (e.g., cohabiting). In addition, we coded the percentage of married sample participants; we coded samples in which all participants were married as 100% and samples with participants of mixed marital status between 49% and 1%, according to the percentage of married sample participants.

Marriage type

We coded samples according to whether the participants were in first marriages or remarriages (*second + marriages*). We coded samples with both first marriages and remarriages as *mixed*.

Marriage duration

We coded samples according to the average length of participant marriage at the time of data collection in years. We assigned both interval codes (e.g., 2 years, 13 years, 37 years) and nominal codes (e.g., 0–4 years, 5–9 years, 10–14 years).

Race

We coded samples according to participant race (Asian, Black, Latino, Middle Eastern, White, or other). We coded samples with more than one race as *mixed*. We also coded samples according to the amount of racial diversity present based on the percentage of participants who were members of ethnic/racial minority groups: (a) *significant diversity* (>33%), (b) *some diversity* (10%–33%), or (c) *virtually no diversity* (<10%).

Nationality

We coded samples according to nationality. We also coded samples as either domestic (United States) or international.

Education

We coded samples according to average attained education level (no high school diploma, high school diploma, some college, college graduate, postgraduate education) for husbands and wives. We also coded samples to compare husband and wife average education level (similar level of education, husband higher level of education, and wife higher level of education).

Annual household income

We coded samples according to average participant annual household income. We converted the average annual household income from each sample into 2008 inflation-adjusted dollars. Given the median household income in 2008—$52,029 (U.S. Census Bureau, **2009**)—we subsequently coded samples as (a) *primarily low income* (<$25,000 annual household income), (b) *primarily middle income* ($25,000–$80,000 annual household income), or (c) *primarily high income* (>$80,000 annual household income).

Sample type

We coded samples as either *clinical* (participants were in marital therapy at the time of data collection), *nonclinical* (participants were not in marital therapy at the time of data collection; a community sample), or *mixed* (some participants were in marital therapy at the time of data collection and some were not).

Statistical Methods

We used meta-analytic statistical methods to (a) determine whether gender differences in marital satisfaction exist between husbands and wives; (b) determine the degree to which gender differences in marital satisfaction exist, if gender differences do indeed exist; and (c) explore the effects of moderating variables on gender differences in marital satisfaction. We used Comprehensive Meta-analysis (Borenstein, Hedges, Higgins, & Rothstein, **2005**) to conduct the statistical analyses.

We implemented dyadic data analysis techniques to address nonindependence and dyadic interaction effects for samples of husband and wife pairs (Kenny, Kashy, & Cook, **2006**). Because we identified a wide range of study designs and participant characteristics across the included sample reports, we determined that random-effects models were more appropriate than fixed-effect models for analyzing the data. Because small sample sizes tend to yield overestimated effect sizes (Hunter & Schmidt, **2004**), we converted all of the sample estimate standardized mean difference effect sizes (d) to unbiased sample estimate standardized mean difference effect sizes (g, also known as Hedges's g) to correct for overestimated bias (Hedges, **1981**). To ensure sample independence, we identified the data source for each included sample and then grouped samples with the same data source, yielding only one mean effect size per sample (Lipsey & Wilson, **2001**).

After we converted the observed gender difference in marital satisfaction for each sample into a weighted common metric (*g*) using a random-effects model, we averaged them to yield an aggregated effect size representing the cumulative sum of gender differences in marital satisfaction across all included samples. We calculated effect sizes for the following aggregated summary effects: magnitude (*g*), precision (SE_g), and significance (*p*). The statistical power of the meta-analysis was 1.0. We used a mixed-effects model (a fixed-effect model across subgroups and a random-effects model within subgroups) to calculate between-subgroup differences for potential moderating variables (Borenstein et al., **2009**). For additional details, see our online supporting method description on the *Journal of Marriage and Family* website: http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1741-3737.

Results

Included Samples Summary

Samples

A total of 226 independent samples (*K*) from 173 reports were included (included report references are available on request). We originally coded 258 samples. We identified 41 nonindependent samples derived from 10 independent data sources. After we recoded each of the 41 nonindependent samples as the relevant data source, there were 227 independent samples. Then we excluded the smallest sample ($n = 4$) because of uncorrectable issues relating to small sample bias, resulting in the 226 included independent samples. Because many samples had multiple gender difference effect sizes (e.g., samples with multiple waves, samples with multiple marital satisfaction measures), there were 451 raw gender difference effect sizes between the 226 included samples. We treated multiple raw gender difference effect sizes as within-sample subgroups by computing a mean effect size for each sample, yielding only one effect size per sample prior to aggregation. The zero-order correlation coefficient for husband marital satisfaction and wife marital satisfaction (r_{hw}) was reported for 63 dyadic data samples and not reported for the remaining 115 dyadic data samples; the average reported r_{hw} was .51.

Descriptive summary data for the included samples are provided in Table 1. The included sample reports were published between 1975 and 2009 (*Mdn* = 1999). Most of the included sample reports (93%) were published journal articles. The data for the samples were collected between 1969 and 2009 (*Mdn* = 1996). The median sample size was 172, and the combined total participants for all included samples was 101,110 (50,502 husbands and 50,608 wives).

Table 1. Descriptive Summary of Sample Characteristics

Characteristic		**%**		***K***
Sample report characteristics				
Publication status				227

Published	93	
Unpublished	7	
Report type		228
Book or chapter in an edited book	1	
Journal article	93	
Unpublished (i.e., dissertation and thesis)	6	
Report decade		229
1970s	14	
1980s	12	
1990s	27	
2000s	47	
Marital satisfaction measures		
Type		236
Marital dissatisfaction	4	
Marital satisfaction	96	
Construction		228
Multiple-item	95	
Single-item	5	
Measure		244
DAS	27	
ENRICH	4	
KMS	6	
MAT	27	
MSI	2	
QMI	5	
QMR	2	

SMD	3	
Other multiple-item measures	21	
Single-item measures	5	
Type of data collection		226
Prospective self-report	100	
Retrospective self-report	0	
Third-party observation	0	
Data		
Dyadic/nondyadic data		228
Dyadic (i.e., data for husband–wife pairs)	76	
Nondyadic (i.e., data for only one spouse)	24	
Collection year		230
1970s	19	
1980s	14	
1990s	35	
2000s	32	
Husband–wife marital satisfaction correlation (r_{hw})		232
Not reported ($r_{imputed}$)	50	
Reported	27	
Not applicable (i.e., independent groups)	23	
Descriptive participant variables		
Marital status		227
Married	96	
Mixed	4	
Marriage type		232
First marriages	21	

Remarriages (second + marriages)	2	
Mixed	12	
Not reported	65	
Marriage duration (average)		242
0–4 years	23	
5–9 years	17	
10–14 years	17	
15–19 years	11	
20–29 years	7	
30+ years	4	
Not reported	21	
Race		229
Asian	2	
Black	0	
Latino	1	
Middle Eastern	3	
White	23	
Mixed	23	
Not reported	48	
Racial diversity		231
Significant diversity (> 33% Non-European)	10	
Some diversity (10%–33% Non-European)	19	
Virtually no diversity (< 10% Non-European)	23	
Not reported	48	
Nationality		226

Domestic	80	
International	20	
Canadian (16 samples)		
Israeli (6 samples)		
Australian (4 samples)		
German, Palestinian (3 samples each)		
Chinese, Japanese, Russian (2 samples each)		
Belgian, Czech, Dutch, Finnish, Italian, Lithuanian, Swiss, Taiwanese, Turkic (1 sample each)		
Education of husbands (average)		233
No high school diploma	3	
High school diploma	2	
Some college	44	
College graduate	17	
Postgraduate education	0	
Not reported	34	
Education of wives (average)		233
No high school diploma	2	
High school diploma	2	
Some college	45	
College graduate	17	
Postgraduate education	0	
Not reported	34	
Education level similarity		231
Similar level of education	60	
Husband higher level of education	2	

Wife higher level of education	3	
Not reported	35	
Annual household income (average in 2008 USD)		228
Primarily low income (< $25,000)	6	
Primarily middle income ($25,000–$80,000)	29	
Primarily high income (> $80,000)	5	
Not reported	60	
Sample type (%)		226
Clinical	9	
Nonclinical	89	
Mixed	2	

Note : K = the number of relevant samples for each characteristic. Because samples from reports using the same data source were treated as one sample with multiple within-sample subgroups, the K values for the specified characteristics vary, reflecting within-sample subgroups that represent more than one category for the same characteristic. DAS = Dyadic Adjustment Scale; ENRICH = ENRICH Marital Satisfaction Scale; KMS = Kansas Marital Satisfaction Scale; MAT = Short Marital Adjustment Test; MSI = Marital Satisfaction Inventory; QMI = Quality of Marriage Index; SMD = Semantic Differential Scale.

Participants

The participants in the included samples were predominately White, middle-class Americans with some level of college education. Based on the 137 samples for which age at the time of data collection was reported (61%), the average age for husbands was 38.5 years (SD = 9.1 years), and the average age for wives was 36.5 years (SD = 8.8 years). Based on the 147 samples for which marriage duration at the time of data collection was reported (65%), the participants had been married, on average, between 0 and 37 years (Mdn = 10.0 years, M = 10.8 years, SD = 8.0 years). Although 10 samples included both married and nonmarried participants (i.e., mixed marital status, k = 10; 4%), 99% of the participants in the included samples were married at the time of data collection.

Aggregated standardized mean gender difference in marital satisfaction

The results indicated heterogeneous effect size distributions across the included samples (Q = 1,724.82, p = .000, K = 226), signifying that the variability across the sample effect sizes was more than what would be expected from participant-level sampling error alone and supporting our initial decision to use random-effects models instead of fixed-effect models. The aggregated summary

results for gender differences in marital satisfaction are presented in Table 2. The following ranges provide a guideline specifically for interpreting gender difference effect size magnitude: very small ($g \leq 0.10$), small ($0.11 \leq g \geq 0.35$), moderate ($0.36 \leq g \geq 0.65$), large ($0.66 \leq g \geq 1.00$), and very large ($g > 1.00$; Hyde, **2005**). A positive effect size indicates that marital satisfaction was, on average, greater for husbands, whereas a negative effect size indicates that marital satisfaction was, on average, higher for wives.

Table 2. Aggregated Standardized Mean Gender Difference in Marital Satisfaction Summary Results

Sample and effects model	Subgroup point estimate summary information				
	Type	Value	*SE*	95% CI	*p*
Clinical and nonclinical ($K = 226$)					
Random	*d*	0.04	.01	[0.01, 0.06]	.002
Random	*g*	0.04	.01	[0.01, 0.06]	.002
Fixed	*g*	0.04	.00	[0.04, 0.05]	.000
Random	OR	1.07		[1.03, 1.11]	.002
Nonclinical ($k = 201$)					
Random	*d*	0.02	.01	[−0.01, 0.04]	.154
Random	*g*	0.02	.01	[−0.01, 0.04]	.155
Fixed	*g*	0.01	.00	[0.00, 0.01]	.150
Random	OR	1.03		[0.99, 1.08]	.154
Clinical ($k = 21$)					
Random	*d*	0.23	.06	[0.11, 0.35]	.000
Random	*g*	0.22	.06	[0.11, 0.34]	.000
Fixed	*g*	0.18	.03	[0.12, 0.24]	.000
Random	OR	1.51		[1.23, 1.87]	.000

Note : Fixed-effect models involve an analytic approach to estimating the mean of distributed effects for homogeneous effect size distributions within aggregated variables (i.e., variation attributable to sampling error); random-effects models involve an analytic approach to estimating the mean of distributed effects for heterogeneous effect size distributions within aggregated variables (i.e., variation beyond that attributable to sampling error; Hunter & Schmidt, **2004** ; Lipsey & Wilson, **2001**). K = the total number of independent samples; k = the number of

Schmidt, 2004 ; Lipsey & Wilson, 2001). K = the total number of independent samples; k = the number of independent subsamples; d = aggregated biased standardized mean difference; g = aggregated unbiased standardized mean difference (Hedges's g); OR = odds ratio; SE = standard error; CI = confidence interval. Positive values for *g* and *d* indicate that marital satisfaction was, on average, greater for husbands; negative values for *g* and *d* indicate that marital satisfaction was, on average, higher for wives. Values for OR greater than 1.00 indicate that marital satisfaction was, on average, greater for husbands; values for OR less than 1.00 indicate that marital satisfaction was, on average, higher for wives. Mixed clinical and community samples ($k = 4$) were excluded from both the nonclinical sample analyses and the clinical sample analyses.

The aggregated unbiased standardized mean gender difference in marital satisfaction (Hedges's *g*) was 0.04 ($SE_g = .01$, 95% confidence interval [CI] [0.01, 0.06], $p = .002$, $K = 226$). Therefore, the magnitude of average gender differences in marital satisfaction ($g = 0.04$) can be interpreted as a very small effect (Hyde, 2005). Because the difference effect is positive, the results indicate that wives' average level of marital satisfaction was slightly lower than husbands'. We are 95% confident that the true population marital satisfaction gender mean difference (μ) is between 0.01 and 0.06. The two-tailed test for statistical significance ($p = .002$) rejects the null hypothesis that the true mean effect size (μ) is zero.

We converted the Hedges's *g* to an odds ratio to allow for additional interpretation. The resultant odds ratio was 1.07, indicating the odds that husbands were more satisfied with their marriage were 1.07 times higher than the odds that wives were more satisfied with their marriage. Stated more simply, on average, wives were only slightly less likely (7%) to be satisfied with their marital relationship than husbands.

Another way of appraising the magnitude of Hedges's *g* is to convert it to a binomial effect size display (Lipsey & Wilson, 2001 ; Rosenthal & Rubin, 1983). The correlational (*r*) equivalent for the aggregated unbiased mean difference ($g = 0.04$) was .02. The binomial effect size for husbands was .51, and the binomial effect size for wives was .49. The results of this binomial transformation were a 51% success rate for husbands and a 49% success rate for wives, with *success* defined as reporting higher levels of marital satisfaction, on average. In other words, there was only a 2% difference in the success rate between husbands and wives.

Data censoring

We determined, by conducting several tests for data censoring, that the overall impact of publication bias on the results of this study was trivial; therefore, no adjustments for data censoring were necessary. A *trim and fill* test for data censoring (Duval & Tweedie, 2000), which imputed 26 potentially missing samples yielding nonsubstantively different results, suggested the absence of significant publication bias. We also conducted several Orwin's *fail-safe N* tests for data censoring (Orwin & Boruch, 1983) and found there would need to be 9,173 missing samples with a Hedges's *g* of 0.00 added to the included samples to nullify the statistical significance of the observed aggregated effect size ($g = 0.04$, $K = 226$). Furthermore, there would need to be 155 missing samples with a Hedges's *g* of 0.21 added to the included samples to move the very small observed aggregated effect size ($g = 0.04$, $K = 226$) to a small effect size ($g = 0.11$), 723 missing samples with a Hedges's *g* of 0.46 to move the aggregated effect size to a moderate effect size ($g = 0.36$), 1,404 missing samples

with a Hedges's g of 0.76 to move the aggregated effect size to a large effect size ($g = 0.66$), and 2,176 missing samples with a Hedges's g of 1.10 to move the aggregated effect size to a very large effect size ($g = 1.00$).

Moderator analyses

The most significant moderator was the type of sample (clinical or nonclinical). Hedges's g was 0.22 ($SE_g = .06$, $p = .000$, $k = 21$) for participants who were in marital therapy and 0.02 ($SE_g = .01$, $p = .155$, $k = 201$) for participants who were not in marital therapy ($Q = 20.73$, $p = .000$; see Table 2). Thus, wives in marital therapy were significantly less satisfied with their marriage than husbands, and wives not in marital therapy were not significantly less satisfied with their marriage than husbands. The magnitude of average gender differences in marital satisfaction for clinical couples can be interpreted as a small effect. The odds ratio for clinical samples was 1.51, indicating that, among couples in marital therapy, the odds that husbands were more satisfied with their marriage were 1.51 times the odds that wives were more satisfied with their marriage; the odds ratio for nonclinical samples was 1.03 and not statistically different from zero. Thus, wives in marital therapy were 51% less likely to be satisfied with their marital relationship than husbands in marital therapy, whereas wives not in marital therapy were likely to be just as satisfied with their marital relationship as husbands not in marital therapy. The binomial transformation for clinical samples yielded a 55.5% rate for husbands reporting higher levels of marital satisfaction and 44.5% rate for wives reporting higher levels of marital satisfaction, an 11% difference in the rate of reporting higher levels of marital satisfaction between husbands and wives in marital therapy.

Because sample type (clinical vs. nonclinical) was the strongest moderator, we excluded the clinical samples and then reanalyzed all of the moderator variables to control for the variance in marital satisfaction by gender explained by sample type (see Table 3 for the results of significant moderators; a table containing the nonsignificant moderator findings is available on request). Although there is no standard for determining effect size reliability on the basis of the number of aggregated studies (k) beyond the general guideline that the larger the number of aggregated studies, the greater the confidence in the reliability of the associated aggregated effect sizes, moderators with a small number of aggregated studies should be interpreted descriptively.

Table 3. Moderator Variable Analysis Significant Results for Gender Differences in Marital Satisfaction

Moderator	Nonclinical samples only							
	Subgroup summary information					Heterogeneity		
	g	SE_g	95% CI	p	k	Q	p	K
Data								
Dyadic/nondyadic data						14.44	.000	202
Dyadic	−0.01	.01	[−0.03, 0.02]	.521	150			

Nondyadic	**0.10**	**.03**	**[0.05, 0.15]**	**.000**	**52**			
Descriptive participant variables								
Marriage type						8.25	.041	207
First marriages	−0.04	.02	[−0.09, 0.01]	.114	48			
Remarriages (second + marriages)	0.00	.38	[−0.74, 0.74]	.994	4			
Mixed	**0.05**	**.02**	**[0.01, 0.09]**	**.010**	**25**			
Not reported	0.03	.02	[0.00, 0.06]	.097	130			
Marriage duration (average)						17.50	.008	217
0–4 years	**−0.04**	**.02**	**[−0.08, −0.01]**	**.013**	**55**			
5–9 years	0.03	.03	[−0.03, 0.09]	.295	31			
10–14 years	−0.01	.03	[−0.06, 0.05]	.771	35			
15–19 years	0.07	.04	[−0.02, 0.16]	.126	28			
20–29 years	0.06	.05	[−0.04, 0.15]	.227	16			
30+ years	0.01	.06	[−0.11, 0.13]	.883	9			
Not reported	**0.06**	**.02**	**[0.02, 0.10]**	**.009**	**43**			
Nationality								
Domestic (United States)	0.00	.01	[−0.02, 0.02]	.978	158	4.07	.044	201
International	**0.07**	**.03**	**[0.01, 0.14]**	**.032**	**43**			
Racial diversity (domestic)						12.41	.006	183
Significant diversity	0.00	.03	[−0.07, 0.06]	.905	12			
Some diversity	0.02	.02	[−0.02, 0.06]	.323	42			
Virtually no diversity	**−0.05**	**.02**	**[−0.09, 0.00]**	**.041**	**46**			
Not reported	**0.06**	**.02**	**[0.02, 0.10]**	**.003**	**83**			
Race (domestic & international)						25.04	.000	204
Asian	**0.26**	**.06**	**[0.15, 0.37]**	**.000**	**5**			

Latino	0.14	.13	[−0.12, 0.40]	.288	2			
Middle Eastern	−0.12	.10	[−0.31, 0.07]	.225	6			
Mixed	0.00	.02	[−0.03, 0.04]	.847	45			
White	−0.03	.03	[−0.08, 0.02]	.263	50			
Not reported	0.04	.02	[0.00, 0.08]	.071	96			
Annual household income						8.34	.039	203
Primarily low income	**−0.09**	**.05**	**[−0.19, 0.00]**	**.049**	**6**			
Primarily middle income	−0.01	.02	[−0.05, 0.04]	.800	58			
Primarily high income	0.04	.03	[−0.01, 0.10]	.125	10			
Not reported	**0.03**	**.02**	**[0.00, 0.06]**	**.039**	**129**			

a *Note* : The moderator results were based on the nonclinical samples. Only significant moderator results are included in the table; nonsignificant moderator results are not included. Text in boldface indicates moderator subgroups that were statistically significant ($p \leq .05$). CI = confidence interval; g = aggregated unbiased standardized mean difference (Hedges's g); SE_g = standard error for g ; p = level of statistical significance for the associated Hedges's g or heterogeneity Q test; k = number of independent samples in moderator subgroups; Q = the Q value for the heterogeneity Q test for between-subgroup differences with $K - 1$ degrees of freedom; K = the total number of independent samples included in analyses for the associated moderator variable. All effect estimates were derived from a mixed-effects model. Significant positive Hedges's g values indicate that marital satisfaction was, on average, higher for husbands than wives within the particular moderator subgroup; significant negative Hedges's g values indicate that marital satisfaction was, on average, higher for wives than husbands within the particular moderator subgroup. Levels of significance for Hedges's g values are determined by the degree to which the aggregated values are likely positive or negative (i.e., aggregated Hedges's g values with the lower and upper CI limits on the same side of zero are significant, whereas aggregated Hedges's g values with the lower and upper limits on different sides of zero are not significant). Significant Q values indicate significant differences between at least two moderator subgroups. Heterogeneity Q test significance is determined by the degree to which the confidence intervals for two or more subgroups overlap; a significant heterogeneity Q test of $p = .000$ indicates that the confidence intervals for two subgroups do not overlap.

We did not find significant differences between subgroups for the following moderators: report publication status, report type, report decade, data collection decade, husband education level, wife education level, and education level similarity. The nondyadic-data samples ($g = 0.10$) were 20% more likely to yield results indicating that wives were slightly less satisfied in their marriages than husbands compared to the dyadic-data samples, which were more likely to yield results indicating no gender differences on marital satisfaction between wives and husbands.

We identified very small gender differences ($g = 0.05$) in mixed samples (first marriages and remarriages) and no gender differences in first-marriage samples, remarriage samples, and samples for which marriage type was not reported. Despite the limited number of remarriage samples that

rendered the remarriage effect size more descriptive than conclusive, these results provide some confounded evidence that very small marital satisfaction gender differences may exist in remarriages. An analysis of marriage duration revealed very small gender differences among participants married an average of 4 years or less ($g = -0.04$) and no gender differences among participants married an average of 5 years or more, suggesting that husbands are slightly less satisfied than wives in the first 4 years of marriage. There were very small gender differences in samples comprised of primarily low-income participants ($g = -0.09$) and no gender differences in samples comprised of primarily high-income or primarily middle-income participants, suggesting that husbands are slightly less satisfied than wives in low-income households.

There were significant differences between samples from the United States and samples from other countries, with very small gender differences in international samples ($g = 0.07$) and no gender differences in domestic samples. Insufficient subgroup sample numbers prevented us from making direct comparisons between participants who belonged to domestic ethnic/racial minority subgroups because only two domestic samples had participants who exclusively belonged to one ethnic/racial minority subgroup. Analysis of domestic samples with differing levels of racial diversity indicated very small gender differences in samples with virtually no diversity ($g = -0.04$) and no gender differences in samples with some or significant diversity, suggesting that husbands are slightly less satisfied than wives in predominantly White domestic samples. The results of combined domestic and international samples indicated no gender differences in White samples and a small gender difference in Asian samples ($g = 0.26$), suggesting that wives from China, Japan, and Taiwan were, on average, 1.6 times less likely to be satisfied with their marriage than husbands; because of the limited number of samples of participants who belonged to ethnic/racial minority groups, these particular effect sizes are more descriptive than conclusive.

Discussion

The results of this meta-analysis indicate that, on average, women report less marital satisfaction than men. The magnitude of the overall difference was very small, with the effect size being 0.04, indicating that wives are only 7% less likely to be satisfied with their marital relationship than husbands. However, results of the moderator analyses indicated that the overall significant gender difference is largely attributable to the 21 clinical samples that were included in the meta-analysis. The effect size for couples in marital therapy was 0.22, with wives 51% less likely to be satisfied than husbands. When the 201 community-based samples were analyzed separately from the clinical samples, the effect size was reduced to 0.02, which was not significantly different from 0. These findings indicate that there is no significant gender difference in marital satisfaction among couples in community-based samples.

The lack of significant gender differences in community-based samples is reinforced by the additional finding that there were no significant gender differences when the level of marital satisfaction of husbands and wives in the same relationship (i.e., dyadic data) was compared. These findings are consistent with analyses of data from the National Study of Families and Households that indicated that nearly 80% of couples had the same marital satisfaction scores and, among the remaining couples who reported different scores, the proportion of husbands and wives who reported lower scores was virtually equal (Gager & Sanchez, **2003**). In contrast, we found a small significant gender difference

in marital satisfaction among samples comprised of individual husbands and wives (i.e., non-dyadic data), with wives reporting lower mean levels of marital satisfaction. It is possible that the marital satisfaction difference between nonrelated men and women could be attributable to sample selection bias, with dissatisfied wives possibly more likely than dissatisfied husbands to respond to marital surveys.

Thus, at least from the perspective of marital satisfaction, our results fail to support Bernard's (**1972**) widely held dictum that “his” marriage is better than “her” marriage and that wives consistently report lower marital satisfaction than husbands. Marital dissolution may provide a possible explanation for our findings of no gender differences in marital satisfaction. The finding from our meta-analysis that wives in marital therapy were substantially less satisfied with their marriages than their husbands suggests that the pathway to marital distress most commonly lies with the wife becoming dissatisfied in the relationship; hence, wives are more likely to seek marital therapy (Doss et al., **2003**) and to initiate divorce (Montenegro, **2004** ; Rokach et al., **2004**). Thus, with women more likely to exit marriage, it is possible that the lack of gender differences in relationship satisfaction in intact marriages may reflect the fact that unhappy wives have already dissolved their marriage. There was a substantial increase in the divorce rate in the 1970s (Schoen & Canudas-Romo, **2006**) due to societal increases in individualism and policy changes related to divorce (Bumpass, **1990**). As a result, the difference between Bernard's observation of pervasive unhappiness among wives in 1972 and the findings of this meta-analysis could be due to the rise in the divorce rate, which has increasingly given dissatisfied wives the opportunity to terminate their marriage.

Another possible explanation for the lack of significant differences between husbands and wives in nonclinical samples is that both wives and husbands experience different difficulties that similarly affect marital satisfaction. For example, husbands report experiencing high levels of pressure to meet both family and work demands and are more likely to report wanting to spend additional time with their spouse and children compared to wives (Roxburgh, **2006**). Husbands also tend to struggle with understanding their wives' needs, managing their wives' emotional demands, experiencing rejection within the context of sexual intimacy, and feeling unable to satisfy their wives' desired communication levels and styles (Thompson & Walker, **1989**). These difficulties, along with others, may decrease marital satisfaction for husbands, essentially counterbalancing the marital difficulties wives tend to experience to the point in which husbands and wives are similarly satisfied or dissatisfied with their marriages.

Perhaps our findings of marital satisfaction gender similarity were due to changes that have occurred within the role of the stay-at-home housewife over the past four decades, to which Bernard (**1972**) attributed much of the marital dissatisfaction women experience. Compared to the early 1970s, when Bernard wrote her book, the percentage of working married women and working married mothers has increased dramatically, such that over two thirds of both groups are currently in the labor force (Juhn & Potter, **2006**). Also, although women still shoulder the majority of responsibility child care and household work, the gender gap in the division of these responsibilities has narrowed considerably (Sayer, **2005**). Despite indications of decreased prevalence of patriarchal marriages and increased egalitarian marriages in the United States since the 1980s (Amato et al., **2007**), we found no indication of marital satisfaction gender differences between samples based on report decade or on data-collection decade. Given that we analyzed data collected since the 1970s, our results represent

marital satisfaction gender differences from the 1970s to the present. Thus, we were unable to investigate gender differences occurring prior to the 1970s, the period on which Bernard's comments about marital satisfaction gender differences were based.

Other researchers have also failed to find substantial gender differences in marital relationships. After presenting her findings that marital status and marital quality had similar effects on women's and men's psychological well-being, Williams (**2003**) stated,

> The idea that marriage is bad for women and good for men … has rarely been questioned. Research presented here indicates that … continued acceptance of gender differences in the effects of marital status and marital quality on psychological well-being is unwarranted. (p. 483)

In addition, research indicates that there are few gender differences in the effect of interpersonal processes, such as communication, problem solving, and decision making, on marital outcomes (Kurdek, **2005**). A review of 115 longitudinal studies on marriage concluded that previous theoretical assumptions of gender differences "in the effects on marriage may have been exaggerated" (Karney & Bradbury, **1995** , p. 20). These findings lend empirical support to the *gender similarities hypothesis* , which argues that men and women are similar on most psychological variables (Hyde, **2005**). Hyde's (**2005**) meta-meta-analysis of 46 meta-analyses on psychological variables found that gender differences were nonsignificant or minimal for most psychological variables. She concluded that men and women are much more similar than different and called for the cessation of "overinflated claims of gender differences" (p. 590).

With regard to moderators of marital satisfaction gender differences, we found that approximately half of the moderators were not associated with significant difference. Almost all of the significant moderators had very small effect sizes, with confidence intervals close to zero, indicating unappreciable gender differences in marital satisfaction. The only moderator exception to the general conclusion that there were no significant marital satisfaction gender differences in nonclinical dyadic data samples was a small gender difference in marital satisfaction among couples in China, Japan, and Taiwan; this result should be viewed tentatively, given the limited number of Asian samples.

Limitations

Missing effect sizes are considered the most pervasive limitation of meta-analytic research because they increase sampling error and result bias and decrease the accuracy and generalizability of results (R. A. Peterson & Brown, **2005**). Although we made every effort to maximize inclusivity by identifying relevant samples through (a) implementing multiple strategies for identifying relevant reports, (b) including unpublished reports, (c) imputing effect sizes, and (d) contacting report authors, we undoubtedly were unable to identify, obtain, and utilize some effect sizes. Nevertheless, the result of several data censoring tests suggested the absence of significant publication bias. In addition, there were no differences between samples in regard to whether the report was published or unpublished, further suggesting the absence of significant publication bias.

Another limitation to meta-analysis is that aggregated findings are not impervious to the

methodological issues associated with included sample data (Lipsey & Wilson, **2001**). An additional common limitation to meta-analysis is the loss of detail through aggregation. Because the reports we analyzed rarely included the necessary effect sizes by moderator subgroup for marriage duration, education, and annual household income, we applied the average score to the entire sample, thus restricting our ability to more accurately detect between-subgroup differences for these participant characteristics. Similarly, a major impediment to analyzing marital satisfaction gender differences by ethnic/racial group was that the reports did not provide summary statistics by race. Despite our efforts to identify possible marital satisfaction gender differences within and between races by coding for the degree to which members of ethnic/racial minority groups were represented in samples, this method had inherent limitations, rendering inconclusive our finding that there were no marital satisfaction gender differences by race.

Guidelines for Future Research

Researchers should use dyadic data when studying marital satisfaction because of the gender bias associated with nondyadic data. Furthermore, it is important that researchers use dyadic data analysis to appropriately manage data-nonindependence issues linked with dyadic data. The highly correlated average marital satisfaction scores between husband and wife pairs (r_{hw} = .51) clearly demonstrate the necessity of using appropriate statistical models, such as multilevel modeling, to account for dyadic interaction effects (Kenny et al., **2006**).

Additional research exploring marital satisfaction gender differences for members of ethnic/racial minority groups is needed. It would also be helpful to more closely examine how major life course transitions, such as the transition to parenthood and retirement, affect gender differences in marital satisfaction. Major life course transitions are associated with significant changes in marital interactions and expectations (Miller, **2000** ; Miller & Yorgason, **2009**). Likewise, health problems often affect marital satisfaction, especially for the spouse who is still well (Yorgason, Booth, & Johnson, **2008**). Hence, it would be important to examine possible gender differences in marital satisfaction during and after these transitions.

Note

We extend special thanks to Matthew Martin and Sherie Christensen for their assistance in identifying and locating potential studies; David Curtis, Meagen B. Mace, Lexie Pfeiffer, and William Walker for dedicated efforts in identifying, locating, evaluating, and coding studies; Kimberly Ortiz for her assistance in preparing the data for analysis; and Alan J. Hawkins and Joseph Olsen at Brigham Young University and Michael Borenstein and Steven Tarlow at Biostat Inc. for their data analysis recommendations.

Supporting Information

References

Citing Literature