The Gender Similarities Hypothesis

Janet Shibley Hyde
University of Wisconsin—Madison

The differences model, which argues that males and females are vastly different psychologically, dominates the popular media. Here, the author advances a very different view, the gender similarities hypothesis, which holds that males and females are similar on most, but not all, psychological variables. Results from a review of 46 meta-analyses support the gender similarities hypothesis. Gender differences can vary substantially in magnitude at different ages and depend on the context in which measurement occurs. Overinflated claims of gender differences carry substantial costs in areas such as the workplace and relationships.

Keywords: gender differences, gender similarities, meta-analysis, aggression

The mass media and the general public are captivated by findings of gender differences. John Gray's (1992) *Men Are From Mars, Women Are From Venus*, which argued for enormous psychological differences between women and men, has sold over 30 million copies and been translated into 40 languages (Gray, 2005). Deborah Tannen's (1991) *You Just Don't Understand: Women and Men in Conversation* argued for the *different cultures hypothesis*: that men's and women's patterns of speaking are so fundamentally different that men and women essentially belong to different linguistic communities or cultures. That book was on the *New York Times* bestseller list for nearly four years and has been translated into 24 languages (AnnOnline, 2005). Both of these works, and dozens of others like them, have argued for the *differences hypothesis*: that males and females are, psychologically, vastly different. Here, I advance a very different view—the *gender similarities hypothesis* (for related statements, see Epstein, 1988; Hyde, 1985; Hyde & Plant, 1995; Kimball, 1995).

The Hypothesis

The gender similarities hypothesis holds that males and females are similar on most, but not all, psychological variables. That is, men and women, as well as boys and girls, are more alike than they are different. In terms of effect sizes, the gender similarities hypothesis states that most psychological gender differences are in the close-to-zero ($d \leq 0.10$) or small ($0.11 < d < 0.35$) range, a few are in the moderate range ($0.36 < d < 0.65$), and very few are large ($d = 0.66$–1.00) or very large ($d > 1.00$).

Although the fascination with psychological gender differences has been present from the dawn of formalized psychology around 1879 (Shields, 1975), a few early researchers highlighted gender similarities. Thorndike (1914), for example, believed that psychological gender differences were too small, compared with within-gender variation, to be important. Leta Stetter Hollingworth (1918) reviewed available research on gender differences in mental traits and found little evidence of gender differences. Another important reviewer of gender research in the early 1900s, Helen Thompson Woolley (1914), lamented the gap between the data and scientists' views on the question:

> The general discussions of the psychology of sex, whether by psychologists or by sociologists show such a wide diversity of points of view that one feels that the truest thing to be said at present is that scientific evidence plays very little part in producing convictions. (p. 372)

The Role of Meta-Analysis in Assessing Psychological Gender Differences

Reviews of research on psychological gender differences began with Woolley's (1914) and Hollingworth's (1918) and extended through Maccoby and Jacklin's (1974) watershed book *The Psychology of Sex Differences*, in which they reviewed more than 2,000 studies of gender differences in a wide variety of domains, including abilities, personality, social behavior, and memory. Maccoby and Jacklin dismissed as unfounded many popular beliefs in psychological gender differences, including beliefs that girls are more "social" than boys; that girls are more suggestible; that girls have lower self-esteem; that girls are better at rote learning and simple tasks, whereas boys are better at higher level cognitive processing; and that girls lack achievement motivation. Maccoby and Jacklin concluded that gender differences were well established in only four areas: verbal ability, visual-spatial ability, mathematical ability, and aggression. Overall, then, they found much evidence for gender similarities. Secondary reports of their findings in textbooks and other sources, however, focused almost exclusively on their conclusions about gender differences (e.g., Gleitman, 1981; Lefrançois, 1990).

Preparation of this article was supported in part by National Science Foundation Grant REC 0207109. I thank Nicole Else-Quest, Sara Lindberg, Shelly Grabe, and Jenni Petersen for reviewing and commenting on a draft of this article.

Correspondence concerning this article should be addressed to Janet Shibley Hyde, Department of Psychology, University of Wisconsin—Madison, 1202 West Johnson Street, Madison, WI 53706. E-mail: jshyde@wisc.edu

Vol. 60, No. 6, 581–592 DOI: 10.1037/0003-066X.60.6.581

Janet Shibley Hyde

Shortly after this important work appeared, the statistical method of meta-analysis was developed (e.g., Glass, McGaw, & Smith, 1981; Hedges & Olkin, 1985; Rosenthal, 1991). This method revolutionized the study of psychological gender differences. Meta-analyses quickly appeared on issues such as gender differences in influenceability (Eagly & Carli, 1981), abilities (Hyde, 1981; Hyde & Linn, 1988; Linn & Petersen, 1985), and aggression (Eagly & Steffen, 1986; Hyde, 1984, 1986).

Meta-analysis is a statistical method for aggregating research findings across many studies of the same question (Hedges & Becker, 1986). It is ideal for synthesizing research on gender differences, an area in which often dozens or even hundreds of studies of a particular question have been conducted.

Crucial to meta-analysis is the concept of effect size, which measures the magnitude of an effect—in this case, the magnitude of gender difference. In gender meta-analyses, the measure of effect size typically is d (Cohen, 1988):

$$d = \frac{M_M - M_F}{s_w},$$

where M_M is the mean score for males, M_F is the mean score for females, and s_w is the average within-sex standard deviation. That is, d measures how far apart the male and female means are in standardized units. In gender meta-analysis, the effect sizes computed from all individual studies are averaged to obtain an overall effect size reflecting the magnitude of gender differences across all studies. In the present article, I follow the convention that negative values of d mean that females scored higher on a dimension, and positive values of d indicate that males scored higher.

Gender meta-analyses generally proceed in four steps: (a) The researcher locates all studies on the topic being reviewed, typically using databases such as PsycINFO and carefully chosen search terms. (b) Statistics are extracted from each report, and an effect size is computed for each study. (c) A weighted average of the effect sizes is computed (weighting by sample size) to obtain an overall assessment of the direction and magnitude of the gender difference when all studies are combined. (d) Homogeneity analyses are conducted to determine whether the group of effect sizes is relatively homogeneous. If it is not, then the studies can be partitioned into theoretically meaningful groups to determine whether the effect size is larger for some types of studies and smaller for other types. The researcher could ask, for example, whether gender differences are larger for measures of physical aggression compared with measures of verbal aggression.

The Evidence

To evaluate the gender similarities hypothesis, I collected the major meta-analyses that have been conducted on psychological gender differences. They are listed in Table 1, grouped roughly into six categories: those that assessed cognitive variables, such as abilities; those that assessed verbal or nonverbal communication; those that assessed social or personality variables, such as aggression or leadership; those that assessed measures of psychological well-being, such as self-esteem; those that assessed motor behaviors, such as throwing distance; and those that assessed miscellaneous constructs, such as moral reasoning. I began with meta-analyses reviewed previously by Hyde and Plant (1995), Hyde and Frost (1993), and Ashmore (1990). I updated these lists with more recent meta-analyses and, where possible, replaced older meta-analyses with more up-to-date meta-analyses that used larger samples and better statistical methods.

Hedges and Nowell (1995; see also Feingold, 1988) have argued that the canonical method of meta-analysis—which often aggregates data from many small convenience samples—should be augmented or replaced by data from large probability samples, at least when that is possible (e.g., in areas such as ability testing). Test-norming data as well as data from major national surveys such as the National Longitudinal Study of Youth provide important information. Findings from samples such as these are included in the summary shown in Table 1, where the number of reports is marked with an asterisk.

Inspection of the effect sizes shown in the rightmost column of Table 1 reveals strong evidence for the gender similarities hypothesis. These effect sizes are summarized in Table 2. Of the 128 effect sizes shown in Table 1, 4 were unclassifiable because the meta-analysis provided such a wide range for the estimate. The remaining 124 effect sizes were classified into the categories noted earlier: close-to-zero ($d \leq 0.10$), small ($0.11 < d < 0.35$), moderate ($0.36 < d < 0.65$), large ($d = 0.66$–1.00), or very large (>1.00). The striking result is that 30% of the effect sizes are in the close-to-zero range, and an additional 48% are in the small range. That is, 78% of gender differences are

Table 1
Major Meta-Analyses of Research on Psychological Gender Differences

Study and variable	Age	No. of reports	*d*
	Cognitive variables		
Hyde, Fennema, & Lamon (1990)			
Mathematics computation	All	45	−0.14
Mathematics concepts	All	41	−0.03
Mathematics problem solving	All	48	+0.08
Hedges & Nowell (1995)			
Reading comprehension	Adolescents	5*	−0.09
Vocabulary	Adolescents	4*	+0.06
Mathematics	Adolescents	6*	+0.16
Perceptual speed	Adolescents	4*	−0.28
Science	Adolescents	4*	+0.32
Spatial ability	Adolescents	2*	+0.19
Hyde, Fennema, Ryan, et al. (1990)			
Mathematics self-confidence	All	56	+0.16
Mathematics anxiety	All	53	−0.15
Feingold (1988)			
DAT spelling	Adolescents	5*	−0.45
DAT language	Adolescents	5*	−0.40
DAT verbal reasoning	Adolescents	5*	−0.02
DAT abstract reasoning	Adolescents	5*	−0.04
DAT numerical ability	Adolescents	5*	−0.10
DAT perceptual speed	Adolescents	5*	−0.34
DAT mechanical reasoning	Adolescents	5*	+0.76
DAT space relations	Adolescents	5*	+0.15
Hyde & Linn (1988)			
Vocabulary	All	40	−0.02
Reading comprehension	All	18	−0.03
Speech production	All	12	−0.33
Linn & Petersen (1985)			
Spatial perception	All	62	+0.44
Mental rotation	All	29	+0.73
Spatial visualization	All	81	+0.13
Voyer et al. (1995)			
Spatial perception	All	92	+0.44
Mental rotation	All	78	+0.56
Spatial visualization	All	116	+0.19
Lynn & Irwing (2004)			
Progressive matrices	6–14 years	15	+0.02
Progressive matrices	15–19 years	23	+0.16
Progressive matrices	Adults	10	+0.30
Whitley et al. (1986)			
Attribution of success to ability	All	29	+0.13
Attribution of success to effort	All	29	−0.04
Attribution of success to task	All	29	−0.01
Attribution of success to luck	All	29	−0.07
Attribution of failure to ability	All	29	+0.16
Attribution of failure to effort	All	29	+0.15
Attribution of failure to task	All	29	−0.08
Attribution of failure luck	All	29	−0.15
	Communication		
Anderson & Leaper (1998)			
Interruptions in conversation	Adults	53	+0.15
Intrusive interruptions	Adults	17	+0.33
Leaper & Smith (2004)			
Talkativeness	Children	73	−0.11
Affiliative speech	Children	46	−0.26
Assertive speech	Children	75	+0.11

(table continues)

Table 1 (*continued*)

Study and variable	Age	No. of reports	*d*
	Communication (*continued*)		
Dindia & Allen (1992)			
Self-disclosure (all studies)	—	205	−0.18
Self-disclosure to stranger	—	99	−0.07
Self-disclosure to friend	—	50	−0.28
LaFrance et al. (2003)			
Smiling	Adolescents and adults	418	−0.40
Smiling: Aware of being observed	Adolescents and adults	295	−0.46
Smiling: Not aware of being observed	Adolescents and adults	31	−0.19
McClure (2000)			
Facial expression processing	Infants	29	−0.18 to −0.92
Facial expression processing	Children and adolescents	89	−0.13 to −0.18
	Social and personality variables		
Hyde (1984, 1986)			
Aggression (all types)	All	69	+0.50
Physical aggression	All	26	+0.60
Verbal aggression	All	6	+0.43
Eagly & Steffen (1986)			
Aggression	Adults	50	+0.29
Physical aggression	Adults	30	+0.40
Psychological aggression	Adults	20	+0.18
Knight et al. (2002)			
Physical aggression	All	41	+0.59
Verbal aggression	All	22	+0.28
Aggression in low emotional arousal context	All	40	+0.30
Aggression in emotional arousal context	All	83	+0.56
Bettencourt & Miller (1996)			
Aggression under provocation	Adults	57	+0.17
Aggression under neutral conditions	Adults	50	+0.33
Archer (2004)			
Aggression in real-world settings	All	75	+0.30 to +0.63
Physical aggression	All	111	+0.33 to +0.84
Verbal aggression	All	68	+0.09 to +0.55
Indirect aggression	All	40	−0.74 to +0.05
Stuhlmacher & Walters (1999)			
Negotiation outcomes	Adults	53	+0.09
Walters et al. (1998)			
Negotiator competitiveness	Adults	79	+0.07
Eagly & Crowley (1986)			
Helping behavior	Adults	99	+0.13
Helping: Surveillance context	Adults	16	+0.74
Helping: No surveillance	Adults	41	−0.02
Oliver & Hyde (1993)			
Sexuality: Masturbation	All	26	+0.96
Sexuality: Attitudes about casual sex	All	10	+0.81
Sexual satisfaction	All	15	−0.06
Attitudes about extramarital sex	All	17	+0.29
Murnen & Stockton (1997)			
Arousal to sexual stimuli	Adults	62	+0.31
Eagly & Johnson (1990)			
Leadership: Interpersonal style	Adults	153	−0.04 to −0.07
Leadership: Task style	Adults	154	0.00 to −0.09
Leadership: Democratic vs. autocratic	Adults	28	+0.22 to +0.34
Eagly et al. (1992)			
Leadership: Evaluation	Adults	114	+0.05
Eagly et al. (1995)			
Leadership effectiveness	Adults	76	−0.02

Table 1 (*continued*)

Study and variable	Age	No. of reports	*d*
Social and personality variables (*continued*)			
Eagly et al. (2003)			
Leadership: Transformational	Adults	44	−0.10
Leadership: Transactional	Adults	51	−0.13 to +0.27
Leadership: Laissez-faire	Adults	16	+0.16
Feingold (1994)			
Neuroticism: Anxiety	Adolescents and adults	13*	−0.32
Neuroticism: Impulsiveness	Adolescents and adults	6*	−0.01
Extraversion: Gregariousness	Adolescents and adults	10*	−0.07
Extraversion: Assertiveness	Adolescents and adults	10*	+0.51
Extraversion: Activity	Adolescents and adults	5	+0.08
Openness	Adolescents and adults	4*	+0.19
Agreeableness: Trust	Adolescents and adults	4*	−0.35
Agreeableness: Tendermindedness	Adolescents and adults	10*	−0.91
Conscientiousness	Adolescents and adults	4	−0.18
Psychological well-being			
Kling et al. (1999, Analysis I)			
Self-esteem	All	216	+0.21
Kling et al. (1999, Analysis II)			
Self-esteem	Adolescents	15*	+0.04 to +0.16
Major et al. (1999)			
Self-esteem	All	226	+0.14
Feingold & Mazzella (1998)			
Body esteem	All	—	+0.58
Twenge & Nolen-Hoeksema (2002)			
Depression symptoms	8–16 years	310	+0.02
Wood et al. (1989)			
Life satisfaction	Adults	17	−0.03
Happiness	Adults	22	−0.07
Pinquart & Sörensen (2001)			
Life satisfaction	Elderly	176	+0.08
Self-esteem	Elderly	59	+0.08
Happiness	Elderly	56	−0.06
Tamres et al. (2002)			
Coping: Problem-focused	All	22	−0.13
Coping: Rumination	All	10	−0.19
Motor behaviors			
Thomas & French (1985)			
Balance	3–20 years	67	+0.09
Grip strength	3–20 years	37	+0.66
Throw velocity	3–20 years	12	+2.18
Throw distance	3–20 years	47	+1.98
Vertical jump	3–20 years	20	+0.18
Sprinting	3–20 years	66	+0.63
Flexibility	5–10 years	13	−0.29
Eaton & Enns (1986)			
Activity level	All	127	+0.49
Miscellaneous			
Thoma (1986)			
Moral reasoning: Stage	Adolescents and adults	56	−0.21
Jaffee & Hyde (2000)			
Moral reasoning: Justice orientation	All	95	+0.19
Moral reasoning: Care orientation	All	160	−0.28
Silverman (2003)			
Delay of gratification	All	38	−0.12
Whitley et al. (1999)			
Cheating behavior	All	36	+0.17
Cheating attitudes	All	14	+0.35

(table continues)

Table 1 (*continued*)

Study and variable	Age	No. of reports	*d*
Whitley (1997)			
Computer use: Current	All	18	+0.33
Computer self-efficacy	All	29	+0.41
Konrad et al. (2000)			
Job attribute preference: Earnings	Adults	207	+0.12
Job attribute preference: Security	Adults	182	−0.02
Job attribute preference: Challenge	Adults	63	+0.05
Job attribute preference: Physical work environment	Adults	96	−0.13
Job attribute preference: Power	Adults	68	+0.04

Note. Positive values of *d* represent higher scores for men and/or boys; negative values of *d* represent higher scores for women and/or girls. Asterisks indicate that data were from major, large national samples. Dashes indicate that data were not available (i.e., the study in question did not provide this information clearly). No. = number; DAT = Differential Aptitude Test.

small or close to zero. This result is similar to that of Hyde and Plant (1995), who found that 60% of effect sizes for gender differences were in the small or close-to-zero range.

The small magnitude of these effects is even more striking given that most of the meta-analyses addressed the classic gender differences questions—that is, areas in which gender differences were reputed to be reliable, such as mathematics performance, verbal ability, and aggressive behavior. For example, despite Tannen's (1991) assertions, gender differences in most aspects of communication are small. Gilligan (1982) has argued that males and females speak in a different moral "voice," yet meta-analyses show that gender differences in moral reasoning and moral orientation are small (Jaffee & Hyde, 2000).

The Exceptions

As noted earlier, the gender similarities hypothesis does not assert that males and females are similar in absolutely every domain. The exceptions—areas in which gender differences are moderate or large in magnitude—should be recognized.

The largest gender differences in Table 1 are in the domain of motor performance, particularly for measures such as throwing velocity ($d = 2.18$) and throwing distance ($d = 1.98$) (Thomas & French, 1985). These differences are particularly large after puberty, when the gender gap in muscle mass and bone size widens.

A second area in which large gender differences are found is some—but not all—measures of sexuality (Oliver & Hyde, 1993). Gender differences are strikingly large for incidences of masturbation and for attitudes about sex in a casual, uncommitted relationship. In contrast, the gender difference in reported sexual satisfaction is close to zero.

Across several meta-analyses, aggression has repeatedly shown gender differences that are moderate in magnitude (Archer, 2004; Eagly & Steffen, 1986; Hyde, 1984, 1986). The gender difference in physical aggression is particularly reliable and is larger than the gender difference in verbal aggression. Much publicity has been given to gender differences in relational aggression, with girls scoring higher (e.g., Crick & Grotpeter, 1995). According to the Archer (2004) meta-analysis, indirect or relational aggression showed an effect size for gender differences of −0.45 when measured by direct observation, but it was only −0.19 for peer ratings, −0.02 for self-reports, and −0.13 for teacher reports. Therefore, the evidence is ambiguous regarding the magnitude of the gender difference in relational aggression.

Table 2
Effect Sizes (n = 124) for Psychological Gender Differences, Based on Meta-Analyses, Categorized by Range of Magnitude

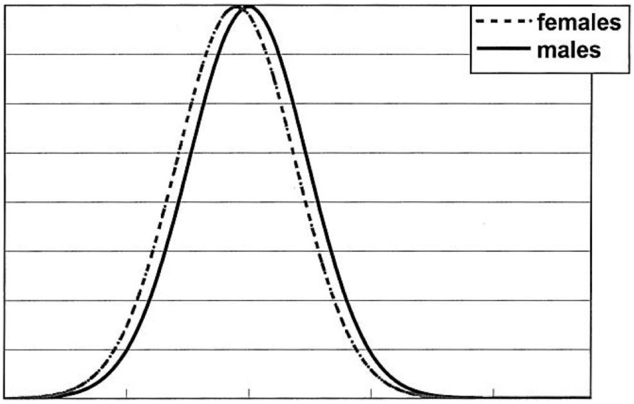
	Effect size range				
Effect sizes	0–0.10	0.11–0.35	0.36–0.65	0.66–1.00	>1.00
Number	37	59	19	7	2
% of total	30	48	15	6	2

The Interpretation of Effect Sizes

The interpretation of effect sizes is contested. On one side of the argument, the classic source is the statistician Cohen (1969, 1988), who recommended that 0.20 be considered a small effect, 0.50 be considered medium, and 0.80 be considered large. It is important to note that he set these guidelines before the advent of meta-analysis, and they have been the standards used in statistical power analysis for decades.

In support of these guidelines are indicators of overlap between two distributions. For example, Kling, Hyde, Showers, and Buswell (1999) graphed two distributions differing on average by an effect size of 0.21, the effect size they found for gender differences in self-esteem. This graph is shown in Figure 1. Clearly, this small effect size

Figure 1
Graphic Representation of a 0.21 Effect Size



Note. Two normal distributions that are 0.21 standard deviations apart (i.e., $d = 0.21$). This is the approximate magnitude of the gender difference in self-esteem, averaged over all samples, found by Kling et al. (1999). From "Gender Differences in Self-Esteem: A Meta-Analysis," by K. C. Kling, J. S. Hyde, C. J. Showers, and B. N. Buswell, 1999, *Psychological Bulletin, 125,* p. 484. Copyright 1999 by the American Psychological Association.

reflects distributions that overlap greatly—that is, that show more similarity than difference. Cohen (1988) developed a *U* statistic that quantifies the percentage of nonoverlap of distributions. For $d = 0.20$, $U = 15\%$; that is, 85% of the areas of the distributions overlap. According to another Cohen measure of overlap, for $d = 0.20$, 54% of individuals in Group A exceed the 50th percentile for Group B.

For another way to consider the interpretation of effect sizes, *d* can also be expressed as an equivalent value of the Pearson correlation, *r* (Cohen, 1988). For the small effect size of 0.20, $r = .10$, certainly a small correlation. A *d* of 0.50 is equivalent to an *r* of .24, and for $d = 0.80$, $r = .37$.

Rosenthal (1991; Rosenthal & Rubin, 1982) has argued the other side of the case—namely, that seemingly small effect sizes can be important and make for impressive applied effects. As an example, he took a two-group experimental design in which one group is treated for cancer and the other group receives a placebo. He used the method of binomial effect size display (BESD) to illustrate the consequences. Using this method, for example, an *r* of .32 between treatment and outcome, accounting for only 10% of the variance, translates into a survival rate of 34% in the placebo group and 66% in the treated group. Certainly, the effect is impressive.

How does this apply to the study of gender differences? First, in terms of costs of errors in scientific decision making, psychological gender differences are quite a different matter from curing cancer. So, interpretation of the magnitude of effects must be heavily conditioned by the costs of making Type I and Type II errors for the particular question under consideration. I look forward to statisticians developing indicators that take these factors into account. Second, Rosenthal used the *r* metric, and when this is translated into *d*, the effects look much less impressive. For example, a *d* of 0.20 is equivalent to an *r* of 0.10, and Rosenthal's BESD indicates that that effect is equivalent to cancer survival increasing from 45% to 55%—once again, a small effect. A close-to-zero effect size of 0.10 is equivalent to an *r* of .05, which translates to cancer survival rates increasing only from 47.5% to 52.5% in the treatment group compared with the control group. In short, I believe that Cohen's guidelines provide a reasonable standard for the interpretation of gender differences effect sizes.

One caveat should be noted, however. The foregoing discussion is implicitly based on the assumption that the variabilities in the male and female distributions are equal. Yet the greater male variability hypothesis was originally proposed more than a century ago, and it survives today (Feingold, 1992; Hedges & Friedman, 1993). In the 1800s, this hypothesis was proposed to explain why there were more male than female geniuses and, at the same time, more males among the mentally retarded. Statistically, the combination of a small average difference favoring males and a larger standard deviation for males, for some trait such as mathematics performance, could lead to a lopsided gender ratio favoring males in the upper tail of the distribution reflecting exceptional talent. The statistic used to investigate this question is the variance ratio (VR), the ratio of the male variance to the female variance. Empirical investigations of the VR have found values of 1.00–1.08 for vocabulary (Hedges & Nowell, 1995), 1.05–1.25 for mathematics performance (Hedges & Nowell), and 0.87–1.04 for self-esteem (Kling et al., 1999). Therefore, it appears that whether males or females are more variable depends on the domain under consideration. Moreover, most VR estimates are close to 1.00, indicating similar variances for males and females. Nonetheless, this issue of possible gender differences in variability merits continued investigation.

Developmental Trends

Not all meta-analyses have examined developmental trends and, given the preponderance of psychological research on college students, developmental analysis is not always possible. However, meta-analysis can be powerful for identifying age trends in the magnitude of gender differences. Here, I consider a few key examples of meta-analyses that have taken this developmental approach (see Table 3).

At the time of the meta-analysis by Hyde, Fennema, and Lamon (1990), it was believed that gender differences in mathematics performance were small or nonexistent in childhood and that the male advantage appeared beginning around the time of puberty (Maccoby & Jacklin, 1974). It was also believed that males were better at high-level mathematical problems that required complex processing, whereas females were better at low-level mathematics that required only simple computation. Hyde and colleagues addressed both hypotheses in their meta-analysis. They found a small gender difference favoring girls in computation in elementary school and middle school and no gender difference in computation in the high school years.

Table 3
Selected Meta-Analyses Showing Developmental Trends in the Magnitude of Gender Differences

Study and variable	Age (years)	No. of reports	*d*
Hyde, Fennema, & Lamon (1990)			
Mathematics: Complex problem solving	5–10	11	0.00
	11–14	21	−0.02
	15–18	10	+0.29
	19–25	15	+0.32
Kling et al. (1999)			
Self-esteem	7–10	22	+0.16
	11–14	53	+0.23
	15–18	44	+0.33
	19–22	72	+0.18
	23–59	16	+0.10
	>60	6	−0.03
Major et al. (1999)			
Self-esteem	5–10	24	+0.01
	11–13	34	+0.12
	14–18	65	+0.16
	19 or older	97	+0.13
Twenge & Nolen-Hoeksema (2002)			
Depressive symptoms	8–12	86	−0.04
	13–16	49	+0.16
Thomas & French (1985)			
Throwing distance	3–8	—	+1.50 to +2.00
	16–18	—	+3.50

Note. Positive values of *d* represent higher scores for men and/or boys; negative values of *d* represent higher scores for women and/or girls. Dashes indicate that data were not available (i.e., the study in question did not provide this information clearly). No. = number.

There was no gender difference in complex problem solving in elementary school or middle school, but a small gender difference favoring males emerged in the high school years ($d = 0.29$). Age differences in the magnitude of the gender effect were significant for both computation and problem solving.

Kling et al. (1999) used a developmental approach in their meta-analysis of studies of gender differences in self-esteem, on the basis of the assertion of prominent authors such as Mary Pipher (1994) that girls' self-esteem takes a nosedive at the beginning of adolescence. They found that the magnitude of the gender difference did grow larger from childhood to adolescence: In childhood (ages 7–10), $d = 0.16$; for early adolescence (ages 11–14), $d = 0.23$; and for the high school years (ages 15–18), $d = 0.33$. However, the gender difference did not suddenly become large in early adolescence, and even in high school, the difference was still not large. Moreover, the gender difference was smaller in older samples; for example, for ages 23–59, $d = 0.10$.

Whitley's (1997) analysis of age trends in computer self-efficacy are revealing. In grammar school samples, $d = 0.09$, whereas in high school samples, $d = 0.66$. This dramatic trend leads to questions about what forces are at work transforming girls from feeling as effective with computers as boys do to showing a large difference in self-efficacy by high school.

These examples illustrate the extent to which the magnitude of gender differences can fluctuate with age. Gender differences grow larger or smaller at different times in the life span, and meta-analysis is a powerful tool for detecting these trends. Moreover, the fluctuating magnitude of gender differences at different ages argues against the differences model and notions that gender differences are large and stable.

The Importance of Context

Gender researchers have emphasized the importance of context in creating, erasing, or even reversing psychological gender differences (Bussey & Bandura, 1999; Deaux & Major, 1987; Eagly & Wood, 1999). Context may exert influence at numerous levels, including the written instructions given for an exam, dyadic interactions between participants or between a participant and an experimenter, or the sociocultural level.

In an important experiment, Lightdale and Prentice (1994) demonstrated the importance of gender roles and social context in creating or erasing the purportedly robust gender difference in aggression. Lightdale and Prentice used the technique of deindividuation to produce a situation that removed the influence of gender roles. *Deindividuation* refers to a state in which the person has lost his or her individual identity; that is, the person has become anonymous. Under such conditions, people should feel no obligation to conform to social norms such as gender roles. Half of the participants, who were college students, were assigned to an individuated condition by having them sit close to the experimenter, identify themselves by name, wear large name tags, and answer personal questions. Participants in the deindividuation condition sat far from the experimenter, wore no name tags, and were simply told to wait. All participants were also told that the experiment required information from only half of the participants, whose behavior would be monitored, and that the other half would remain anonymous. Participants then played an interactive video game in which they first defended and then attacked by dropping bombs. The number of bombs dropped was the measure of aggressive behavior.

The results indicated that in the individuated condition, men dropped significantly more bombs ($M = 31.1$) than women did ($M = 26.8$). In the deindividuated condition, however, there were no significant gender differences and, in fact, women dropped somewhat more bombs ($M = 41.1$) than men ($M = 36.8$). In short, the significant gender difference in aggression disappeared when gender norms were removed.

Steele's (1997; Steele & Aronson, 1995) work on stereotype threat has produced similar evidence in the cognitive domain. Although the original experiments concerned African Americans and the stereotype that they are intellectually inferior, the theory was quickly applied to gender and stereotypes that girls and women are bad at math (Brown & Josephs, 1999; Quinn & Spencer, 2001; Spencer, Steele, & Quinn, 1999; Walsh, Hickey, & Duffy, 1999). In one experiment, male and female college students with equivalent math backgrounds were tested (Spencer et al., 1999). In one condition, participants were told that the math test had shown gender difference in the past, and in the other condition, they were told that the test had been shown to be gender fair—that men and women had performed equally on it. In the condition in which participants had been told that the math test was gender fair, there were no gender differences in performance on the test. In the condition in which participants expected gender differences, women underperformed compared with men. This simple manipulation of context was capable of creating or erasing gender differences in math performance.

Meta-analysts have addressed the importance of context for gender differences. In one of the earliest demonstrations of context effects, Eagly and Crowley (1986) meta-analyzed studies of gender differences in helping behavior, basing the analysis in social-role theory. They argued that certain kinds of helping are part of the male role: helping that is heroic or chivalrous. Other kinds of helping are part of the female role: helping that is nurturant and caring, such as caring for children. Heroic helping involves danger to the self, and both heroic and chivalrous helping are facilitated when onlookers are present. Women's nurturant helping more often occurs in private, with no onlookers. Averaged over all studies, men helped more ($d = 0.34$). However, when studies were separated into those in which onlookers were present and participants were aware of it, $d = 0.74$. When no onlookers were present, $d = -0.02$. Moreover, the magnitude of the gender difference was highly correlated with the degree of danger in the helping situation; gender differences were largest favoring males in situations with the most danger. In short, the gender difference in helping behavior can be large, favoring males, or close to zero, depending on the social context in which the behavior is measured. Moreover, the pattern of gender differences is consistent with social-role theory.

Anderson and Leaper (1998) obtained similar context effects in their meta-analysis of gender differences in conversational interruption. At the time of their meta-analysis, it was widely believed that men interrupted women considerably more than the reverse. Averaged over all studies, however, Anderson and Leaper found a d of 0.15, a small effect. The effect size for intrusive interruptions (excluding back-channel interruptions) was larger: 0.33. It is important to note that the magnitude of the gender difference varied greatly depending on the social context in which interruptions were studied. When dyads were observed, $d = 0.06$, but with larger groups of three or more, $d = 0.26$. When participants were strangers, $d = 0.17$, but when they were friends, $d = -0.14$. Here, again, it is clear that gender differences can be created, erased, or reversed, depending on the context.

In their meta-analysis, LaFrance, Hecht, and Paluck (2003) found a moderate gender difference in smiling ($d = -0.41$), with girls and women smiling more. Again, the magnitude of the gender difference was highly dependent on the context. If participants had a clear awareness that they were being observed, the gender difference was larger ($d = -0.46$) than it was if they were not aware of being observed ($d = -0.19$). The magnitude of the gender difference also depended on culture and age.

Dindia and Allen (1992) and Bettencourt and Miller (1996) also found marked context effects in their gender meta-analyses. The conclusion is clear: The magnitude and even the direction of gender differences depends on the context. These findings provide strong evidence against the differences model and its notions that psychological gender differences are large and stable.

Costs of Inflated Claims of Gender Differences

The question of the magnitude of psychological gender differences is more than just an academic concern. There are serious costs of overinflated claims of gender differences (for an extended discussion of this point, see Barnett & Rivers, 2004; see also White & Kowalski, 1994). These costs occur in many areas, including work, parenting, and relationships.

Gilligan's (1982) argument that women speak in a different moral "voice" than men is a well-known example of the differences model. Women, according to Gilligan, speak in a moral voice of caring, whereas men speak in a voice of justice. Despite the fact that meta-analyses disconfirm her arguments for large gender differences (Jaffee & Hyde, 2000; Thoma, 1986; Walker, 1984), Gilligan's ideas have permeated American culture. One consequence of this overinflated claim of gender differences is that it reifies the stereotype of women as caring and nurturant and men as lacking in nurturance. One cost to men is that they may believe that they cannot be nurturant, even in their role as father. For women, the cost in the workplace can be enormous. Women who violate the stereotype of being nurturant and nice can be penalized in hiring and evaluations. Rudman and Glick (1999), for example, found that female job applicants who displayed agentic qualities received considerably lower hireability ratings than agentic male applicants ($d = 0.92$) for a managerial job that had been "feminized" to require not only technical skills and the ability to work under pressure but also the ability to be helpful and sensitive to the needs of others. The researchers concluded that women must present themselves as competent and agentic to be hired, but they may then be viewed as interpersonally deficient and uncaring and receive biased work evaluations because of their violation of the female nurturance stereotype.

A second example of the costs of unwarranted validation of the stereotype of women as caring nurturers comes from Eagly, Makhijani, and Klonsky's (1992) meta-analysis of studies of gender and the evaluation of leaders. Overall, women leaders were evaluated as positively as men leaders ($d = 0.05$). However, women leaders portrayed as uncaring autocrats were at a more substantial disadvantage than were men leaders portrayed similarly ($d = 0.30$). Women who violated the caring stereotype paid for it in their evaluations. The persistence of the stereotype of women as nurturers leads to serious costs for women who violate this stereotype in the workplace.

The costs of overinflated claims of gender differences hit children as well. According to stereotypes, boys are better at math than girls are (Hyde, Fennema, Ryan, Frost, & Hopp, 1990). This stereotype is proclaimed in mass media headlines (Barnett & Rivers, 2004). Meta-analyses, however, indicate a pattern of gender similarities for math performance. Hedges and Nowell (1995) found a d of 0.16 for large national samples of adolescents, and Hyde, Fennema, and Lamon (1990) found a d of -0.05 for samples of the general population (see also Leahey & Guo, 2000). One cost to children is that mathematically talented girls may be overlooked by parents and teachers because these adults do not expect to find mathematical talent among girls. Parents have lower expectations for their daughters' math success than for their sons' (Lummis & Stevenson, 1990), despite the fact that girls earn better grades in math than boys do (Kimball, 1989). Research has shown repeatedly that parents' expectations for their children's mathematics success relate strongly to outcomes such as the child's mathematics self-confidence and performance, with support for a model in which parents' expectations influence children (e.g., Frome & Eccles, 1998). In short, girls may find their confidence in their ability to succeed in challenging math courses or in a mathematically oriented career undermined by parents' and teachers' beliefs that girls are weak in math ability.

In the realm of intimate heterosexual relationships, women and men are told that they are as different as if they came from different planets and that they communicate in dramatically different ways (Gray, 1992; Tannen, 1991). When relationship conflicts occur, good communication is essential to resolving the conflict (Gottman, 1994). If, however, women and men believe what they have been told—that it is almost impossible for them to communicate with each other—they may simply give up on trying to resolve the conflict through better communication. Therapists will need to dispel erroneous beliefs in massive, unbridgeable gender differences.

Inflated claims about psychological gender differences can hurt boys as well. A large gender gap in self-esteem beginning in adolescence has been touted in popular sources (American Association of University Women, 1991; Orenstein, 1994; Pipher, 1994). Girls' self-esteem is purported to take a nosedive at the beginning of adolescence, with the implication that boys' self-esteem does not. Yet meta-analytic estimates of the magnitude of the gender difference have all been small or close to zero: $d = 0.21$ (Kling et al., 1999, Analysis I), $d = 0.04–0.16$ (Kling et al., 1999, Analysis II), and $d = 0.14$ (Major, Barr, Zubek, & Babey, 1999). In short, self-esteem is roughly as much a problem for adolescent boys as it is for adolescent girls. The popular media's focus on girls as the ones with self-esteem problems may carry a huge cost in leading parents, teachers, and other professionals to overlook boys' self-esteem problems, so that boys do not receive the interventions they need.

As several of these examples indicate, the gender similarities hypothesis carries strong implications for practitioners. The scientific evidence does not support the belief that men and women have inherent difficulties in communicating across gender. Neither does the evidence support the belief that adolescent girls are the only ones with self-esteem problems. Therapists who base their practice in the differences model should reconsider their approach on the basis of the best scientific evidence.

Conclusion

The gender similarities hypothesis stands in stark contrast to the differences model, which holds that men and women, and boys and girls, are vastly different psychologically. The gender similarities hypothesis states, instead, that males and females are alike on most—but not all—psychological variables. Extensive evidence from meta-analyses of research on gender differences supports the gender similarities hypothesis. A few notable exceptions are some motor behaviors (e.g., throwing distance) and some aspects of sexuality, which show large gender differences. Aggression shows a gender difference that is moderate in magnitude.

It is time to consider the costs of overinflated claims of gender differences. Arguably, they cause harm in numerous realms, including women's opportunities in the workplace, couple conflict and communication, and analyses of self-esteem problems among adolescents. Most important, these claims are not consistent with the scientific data.

REFERENCES

American Association of University Women. (1991). *Shortchanging girls, shortchanging America: Full data report.* Washington, DC: Author.
Anderson, K. J., & Leaper, C. (1998). Meta-analyses of gender effects on conversational interruption: Who, what, when, where, and how. *Sex Roles, 39,* 225–252.
AnnOnline. (2005). *Biography: Deborah Tannen.* Retrieved January 10, 2005, from http://www.annonline.com
Archer, J. (2004). Sex differences in aggression in real-world setting: A meta-analytic review. *Review of General Psychology, 8,* 291–322.
Ashmore, R. D. (1990). Sex, gender, and the individual. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 486–526). New York: Guilford Press.
Barnett, R., & Rivers, C. (2004). *Same difference: How gender myths are hurting our relationships, our children, and our jobs.* New York: Basic Books.
Bettencourt, B. A., & Miller, N. (1996). Gender differences in aggression as a function of provocation: A meta-analysis. *Psychological Bulletin, 119,* 422–447.
Brown, R. P., & Josephs, R. A. (1999). A burden of proof: Stereotype relevance and gender differences in math performance. *Journal of Personality and Social Psychology, 76,* 246–257.
Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review, 106,* 676–713.
Cohen, J. (1969). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social–psychological adjustment. *Child Development, 66,* 710–722.
Deaux, K., & Major, B. (1987). Putting gender into context: An interactive model of gender-related behavior. *Psychological Review, 94,* 369–389.
Dindia, K., & Allen, M. (1992). Sex differences in self-disclosure: A meta-analysis. *Psychological Bulletin, 112,* 106–124.
Eagly, A. H., & Carli, L. L. (1981). Sex of researchers and sex-typed communications as determinants of sex differences in influenceability: A meta-analysis of social influence studies. *Psychological Bulletin, 90,* 1–20.
Eagly, A. H., & Crowley, M. (1986). Gender and helping behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin, 100,* 283–308.
Eagly, A. H., Johannesen-Schmidt, M. C., & van Engen, M. L. (2003). Transformational, transactional, and laissez-faire leadership styles: A meta-analysis comparing women and men. *Psychological Bulletin, 129,* 569–591.
Eagly, A. H., & Johnson, B. T. (1990). Gender and leadership style: A meta-analysis. *Psychological Bulletin, 108,* 233–256.
Eagly, A. H., Karau, S. J., & Makhijani, M. G. (1995). Gender and the effectiveness of leaders: A meta-analysis. *Psychological Bulletin, 117,* 125–145.
Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin, 111,* 3–22.
Eagly, A. H., & Steffen, V. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin, 100,* 309–330.
Eagly, A. H., & Wood, W. (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist, 54,* 408–423.
Eaton, W. O., & Enns, L. R. (1986). Sex differences in human motor activity level. *Psychological Bulletin, 100,* 19–28.
Epstein, C. F. (1988). *Deceptive distinctions: Sex, gender, and the social order.* New Haven, CT: Yale University Press.
Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist, 43,* 95–103.
Feingold, A. (1992). Sex differences in variability in intellectual abilities: A new look at an old controversy. *Review of Educational Research, 62,* 61–84.
Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116,* 429–456.
Feingold, A., & Mazzella, R. (1998). Gender differences in body image are increasing. *Psychological Science, 9,* 190–195.
Frome, P. M., & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology, 74,* 435–452.
Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge, MA: Harvard University Press.
Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.
Gleitman, H. (1981). *Psychology.* New York: Norton.
Gottman, J. (1994). *Why marriages succeed or fail.* New York: Simon & Schuster.
Gray, J. (1992). *Men are from Mars, women are from Venus: A practical guide for improving communication and getting what you want in your relationships.* New York: HarperCollins.
Gray, J. (2005). *John Gray, Ph.D. is the best-selling relationship author of all time.* Retrieved January 10, 2005, from http://www.marsvenus.com
Hedges, L. V., & Becker, B. J. (1986). Statistical methods in the meta-analysis of research on gender differences. In J. S. Hyde & M. C. Linn (Eds.), *The psychology of gender: Advances through meta-analysis* (pp. 14–50). Baltimore: Johns Hopkins University Press.
Hedges, L. V., & Friedman, L. (1993). Sex differences in variability in intellectual abilities: A reanalysis of Feingold's results. *Review of Educational Research, 63,* 95–105.
Hedges, L. V., & Nowell, A. (1995, July 7). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269,* 41–45.
Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* San Diego, CA: Academic Press.
Hollingworth, L. S. (1918). Comparison of the sexes in mental traits. *Psychological Bulletin, 15,* 427–432.
Hyde, J. S. (1981). How large are cognitive gender differences? A meta-analysis using $\omega^2$and *d*. *American Psychologist, 36,* 892–901.
Hyde, J. S. (1984). How large are gender differences in aggression? A developmental meta-analysis. *Developmental Psychology, 20,* 722–736.
Hyde, J. S. (1985). *Half the human experience: The psychology of women* (3rd ed.). Lexington, MA: Heath.
Hyde, J. S. (1986). Gender differences in aggression. In J. S. Hyde & M. C. Linn (Eds.), *The psychology of gender: Advances through meta-analysis* (pp. 51–66). Baltimore: Johns Hopkins University Press.
Hyde, J. S., Fennema, E., & Lamon, S. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107,* 139–155.
Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly, 14,* 299–324.
Hyde, J. S., & Frost, L. A. (1993). Meta-analysis in the psychology of women. In F. L. Denmark & M. A. Paludi (Eds.), *Psychology of women: A handbook of issues and theories* (pp. 67–103). Westport, CT: Greenwood Press.
Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104,* 53–69.
Hyde, J. S., & Plant, E. A. (1995). Magnitude of psychological gender differences: Another side to the story. *American Psychologist, 50,* 159–161.
Jaffee, S., & Hyde, J. S. (2000). Gender differences in moral orientation: A meta-analysis. *Psychological Bulletin, 126,* 703–726.
Kimball, M. M. (1989). A new perspective on women's math achievement. *Psychological Bulletin, 105,* 198–214.
Kimball, M. M. (1995). *Feminist visions of gender similarities and differences.* Binghamton, NY: Haworth Press.
Kling, K. C., Hyde, J. S., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin, 125,* 470–500.
Knight, G. P., Guthrie, I. K., Page, M. C., & Fabes, R. A. (2002). Emotional arousal and gender differences in aggression: A meta-analysis. *Aggressive Behavior, 28,* 366–393.
Konrad, A. M., Ritchie, J. E., Lieb, P., & Corrigall, E. (2000). Sex differences and similarities in job attribute preferences: A meta-analysis. *Psychological Bulletin, 126,* 593–641.
LaFrance, M., Hecht, M. A., & Paluck, E. L. (2003). The contingent

smile: A meta-analysis of sex differences in smiling. *Psychological Bulletin, 129,* 305–334.

Leahey, E., & Guo, G. (2000). Gender differences in mathematical trajectories. *Social Forces, 80,* 713–732.

Leaper, C., & Smith, T. E. (2004). A meta-analytic review of gender variations in children's language use: Talkativeness, affiliative speech, and assertive speech. *Developmental Psychology, 40,* 993–1027.

Lefrançois, G. R. (1990). *The lifespan* (3rd ed.). Belmont, CA: Wadsworth.

Lightdale, J. R., & Prentice, D. A. (1994). Rethinking sex differences in aggression: Aggressive behavior in the absence of social roles. *Personality and Social Psychology Bulletin, 20,* 34–44.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56,* 1479–1498.

Lummis, M., & Stevenson, H. W. (1990). Gender differences in beliefs and achievement: A cross-cultural study. *Developmental Psychology, 26,* 254–263.

Lynn, R., & Irwing, P. (2004). Sex differences on the progressive matrices: A meta-analysis. *Intelligence, 32,* 481–498.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Major, B., Barr, L., Zubek, J., & Babey, S. H. (1999). Gender and self-esteem: A meta-analysis. In W. B. Swann, J. H. Langlois, & L. A. Gilbert (Eds.) *Sexism and stereotypes in modern society: The gender science of Janet Taylor Spence* (pp. 223–253). Washington, DC: American Psychological Association.

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin, 126,* 424–453.

Murnen, S. K., & Stockton, M. (1997). Gender and self-reported sexual arousal in response to sexual stimuli: A meta-analytic review. *Sex Roles, 37,* 135–154.

Oliver, M. B., & Hyde, J. S. (1993). Gender differences in sexuality: A meta-analysis. *Psychological Bulletin, 114,* 29–51.

Orenstein, P. (1994). *Schoolgirls: Young women, self-esteem, and the confidence gap.* New York: Anchor Books.

Pinquart, M., & Sörensen (2001). Gender differences in self-concept and psychological well-being in old age: A meta-analysis. *Journal of Gerontology: Psychological Sciences, 56B,* P195–P213.

Pipher, M. (1994). *Reviving Ophelia: Saving the selves of adolescent girls.* New York: Ballantine Books.

Quinn, D. M., & Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues, 57,* 55–72.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. ed.). Newbury Park, CA: Sage.

Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74,* 166–169.

Rudman, L. A., & Glick, P. (1999). Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality and Social Psychology, 77,* 1004–1010.

Shields, S. A. (1975). Functionalism, Darwinism, and the psychology of women: A study in social myth. *American Psychologist, 30,* 739–754.

Silverman, I. W. (2003). Gender differences in delay of gratification: A meta-analysis. *Sex Roles, 49,* 451–463.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4–28.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52,* 613–629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797–811.

Stuhlmacher, A. C., & Walters, A. E. (1999). Gender differences in negotiation outcome: A meta-analysis. *Personnel Psychology, 52,* 653–677.

Tamres, L. K., Janicki, D., & Helgeson, V. S. (2002). Sex differences in coping behavior: A meta-analytic review and an examination of relative coping. *Personality and Social Psychology Review, 6,* 2–30.

Tannen, D. (1991). *You just don't understand: Women and men in conversation.* New York: Ballantine Books.

Thoma, S. J. (1986). Estimating gender differences in the comprehension and preference of moral issues. *Developmental Review, 6,* 165–180.

Thomas, J. R., & French, K. E. (1985). Gender differences across age in motor performance: A meta-analysis. *Psychological Bulletin, 98,* 260–282.

Thorndike, E. L. (1914). *Educational psychology* (Vol. 3). New York: Teachers College, Columbia University.

Twenge, J. M., & Nolen-Hoeksema, S. (2002). Age, gender, race, socioeconomic status, and birth cohort differences on the Children's Depression Inventory: A meta-analysis. *Journal of Abnormal Psychology, 111,* 578–588.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117,* 250–270.

Walker, L. J. (1984). Sex differences in the development of moral reasoning: A critical review. *Child Development, 55,* 677–691.

Walsh, M., Hickey, C., & Duffy, J. (1999). Influence of item content and stereotype situation on gender differences in mathematical problem solving. *Sex Roles, 41,* 219–240.

Walters, A. E., Stuhlmacher, A. F., & Meyer, L. L. (1998). Gender and negotiator competitiveness: A meta-analysis. *Organizational Behavior and Human Decision Processes, 76,* 1–29.

White, J. W., & Kowalski, R. M. (1994). Deconstructing the myth of the nonaggressive woman: A feminist analysis. *Psychology of Women Quarterly, 18,* 487–508.

Whitley, B. E. (1997). Gender differences in computer-related attitudes and behavior: A meta-analysis. *Computers in Human Behavior, 13,* 1–22.

Whitley, B. E., McHugh, M. C., & Frieze, I. H. (1986). Assessing the theoretical models for sex differences in causal attributions of success and failure. In J. S. Hyde & M. C. Linn (Eds.), *The psychology of gender: Advances through meta-analysis* (pp. 102–135). Baltimore: Johns Hopkins University Press.

Whitley, B. E., Nelson, A. B., & Jones, C. J. (1999). Gender differences in cheating attitudes and classroom cheating behavior: A meta-analysis. *Sex Roles, 41,* 657–677.

Wood, W., Rhodes, N., & Whelan, M. (1989). Sex differences in positive well-being: A consideration of emotional style and marital status. *Psychological Bulletin, 106,* 249–264.

Woolley, H. T. (1914). The psychology of sex. *Psychological Bulletin, 11,* 353–379.