

Gender Differences in Domain-Specific Self-Esteem: A Meta-Analysis

Brittany Gentile
University of Georgia

Shelly Grabe
University of California, Santa Cruz

Brenda Dolan-Pascoe
and Jean M. Twenge
San Diego State University

Brooke E. Wells
Center for HIV Educational Studies and Training (CHEST)
and National Development and Research Institutes, Inc.

Alissa Maitino
Alliant International University

This meta-analysis examines gender differences in 10 specific domains of self-esteem across 115 studies, including 428 effect sizes and 32,486 individuals. In a mixed-effects analysis, men scored significantly higher than women on physical appearance ($d = 0.35$), athletic ($d = 0.41$), personal self ($d = 0.28$), and self-satisfaction self-esteem ($d = 0.33$). Women scored higher than men on behavioral conduct ($d = -0.17$) and moral-ethical self-esteem ($d = -0.38$). The gender difference in physical appearance self-esteem was significant only after 1980 and was largest among adults. No significant gender differences appeared in academic, social acceptance, family, and affect self-esteem. The results demonstrate the influence of reflected appraisals on self-esteem.

Keywords: gender differences, self-esteem, adolescence, physical appearance, meta-analysis

Since the mid-1990s, reports in the popular media have suggested that girls—particularly teens—have distressingly low self-esteem (e.g., Pipher, 1994). Recent advertisements like the “Dove Self-Esteem Fund” continue to suggest that girls and women suffer from negative self-images. But are girls and women actually deficient in self-esteem compared with boys and men? Three previous meta-analyses found that the effect size for the gender difference in self-esteem was small, $d = 0.15$; in adolescence the difference was $d = 0.33$, a small to medium effect size (Kling, Hyde, Showers, & Buswell, 1999; Major, Barr, Zubek, & Babey, 1999; Twenge & Campbell, 2001).

However, these previous meta-analyses examined gender differences in global but not domain-specific self-esteem. These are two distinct concepts in the literature. Global self-esteem is “the positivity of the person’s self-evaluation” (Baumeister, 1998, p. 694) or “the level of global regard that one has for the self as a person” (Harter, 1993, p. 88). Domain-specific self-esteem, on the other hand, describes self-satisfaction in specific areas (e.g., ap-

pearance, academics, social). Self-esteem may vary considerably from one domain to another. Thus, domain-specific self-esteem may show larger gender differences than global self-esteem (e.g., Sondhaus, Kurtz, & Strube, 2001; Tiggemann & Rothblum, 1997). The present research undertakes a comprehensive meta-analysis of gender differences in 10 domain-specific areas of self-esteem.

We draw primarily on two theoretical approaches to self-esteem. First, the reflected appraisals model maintains that people base their self-esteem on others’ opinions and perceptions. This idea has a long history; Cooley (1902) argued that self-esteem arises from the appraisals of others. Mead (1934) took this concept a step further, maintaining that our self-esteem is also influenced by the “generalized other”—thus the entire sociocultural environment (which in modern times would include the media). Thus, the reflected appraisals model would predict gender differences in areas where societal standards are different for men and women.

The reflected appraisals model has been conceptualized more recently as a “sociometer” (Leary, Haupt, Strausser, & Chokel, 1998; Leary, Tambor, Terdal, & Downs, 1995) or the “need to belong” (Baumeister & Leary, 1995). These authors theorize that self-esteem is primarily rooted in our relationships with others—what others think of us, whether they accept us, and so on. Interactions with others are good predictors of fluctuations in self-esteem, and even at the level of nations, countries with high levels of interaction between friends have higher mean self-esteem (Denissen, Penke, & Schmitt, 2008). For gender differences in self-esteem, the reflected appraisals model predicts that other people’s perceptions of us—or of our gender as a whole—is a key component of self-esteem. If someone’s interactions with others around a particular domain are positive, then they would have high

Brittany Gentile, Department of Psychology, University of Georgia; Shelly Grabe, Department of Psychology, University of California, Santa Cruz; Brenda Dolan-Pascoe and Jean M. Twenge, Department of Psychology, San Diego State University; Brooke E. Wells, Center for HIV Educational Studies and Training (CHEST) and National Development and Research Institutes, Inc.; Alissa Maitino, California School of Professional Psychology, Alliant International University.

Correspondence concerning this article should be addressed to Jean M. Twenge, Department of Psychology, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-4611. E-mail: jtwenge@mail.sdsu.edu

self-esteem in that domain. If they are negative and judgmental, however, self-esteem would suffer. For example, some have theorized that negative interactions often occur around women's physical appearance, which is often scrutinized and discussed (Fredrickson & Roberts, 1997); thus, women might have lower self-esteem in the domain of physical appearance.

The competencies model (James, 1890) argues that people draw self-esteem from accomplishments in certain areas. The competencies model predicts gender differences in areas where actual performance differs. Thus, it is a model of self-efficacy, predicting that people will feel high self-esteem when they have performed well and will perform well when they have high self-esteem (e.g., Bandura, 1989). Although self-efficacy and self-esteem are distinct concepts, domain-specific self-esteem has some overlap with self-efficacy, because it addresses confidence in a certain area of competence. Some of these areas, like academics and athletics, are performance domains that may show a reciprocal relationship between performance and self-esteem, with each influencing the other. Thus, the competencies model (and self-efficacy theory) predicts that when females perform better in an area, the gender difference in that area will favor females. This most likely occurs through two mechanisms. First, the average individual female is likely to perform better in an area, and thus females' self-esteem will be higher than males' in that area on average. For example, the average girl earns better grades than the average boy, so the gender difference in academic self-esteem should favor females. Second, females may know that their gender group performs better in the area, increasing their self-esteem in that area apart from their individual performance. This mechanism can break down, however, if people apply shifting standards by only comparing themselves within gender groups (e.g., Biernat & Manis, 1994). Overall, however, the competencies model expects that gender differences in self-esteem in specific domains will reflect the gender differences in performance within those domains.

In some cases, the reflected appraisals and competencies models can compliment each other and work together. For example, people may appraise a woman's athletic ability differently because they realize that on average, men's strength and speed are greater (applying shifting standards: Biernat & Manis, 1994). Similarly, people might appraise a good athlete as such because he truly is competent in that area. Thus, in some cases the two models will make similar predictions. In others, however, they will predict different results because the appraisal and the competence of males versus females in certain areas are discrepant—that is, the view of others and actual competence do not agree.

This meta-analysis includes studies that administered one of the four most widely used scales of multifaceted self-esteem: the Harter Self-Perception Profile (Harter), the Self-Description Questionnaire (SDQ), the Tennessee Self-Concept Scale (TSCS), or the Piers-Harris Self-Concept Scale (Piers-Harris). Among them, these scales tap 10 domains of self-esteem: physical appearance, athletic, academic, social acceptance, family, behavioral conduct, affect, personal self, self-satisfaction, and moral-ethical. We review previous research on gender differences in each of these domains and outline what each theory (reflected appraisals or competencies) would predict for gender differences in that area.

Physical Appearance

There is no objective difference in attractiveness between the genders, thus the competencies model would predict no gender differences. Competencies might even predict a female advantage, as—at least in Western societies—both men and women prefer to look at female bodies rather than male bodies, and women focus more on their appearance. Appearance is more central to girls' self-esteem than boys', with body image a stronger predictor of global self-esteem in females compared with males (Allgood-Merten, Lewinsohn, & Hops, 1990; Polce-Lynch, Myers, Kilmartin, Forssmann-Falck, & Kliewer, 1998).

However, the reflected appraisals model makes the opposite prediction. It is often more difficult for girls and women to feel positive about their appearance because of media messages promoting extremely high standards for female appearance (e.g., Fredrickson & Roberts, 1997; Harter, 1993). This includes the pressure to be thin, which comes not only from the media but also from family and friends (Wertheim, Paxton, Shultz, & Muir, 1997). In other words, reflected appraisals create a self-esteem deficit in appearance self-esteem among females. Critical reflected appraisals may be one of the sources for the perception that girls' self-esteem decreases during adolescence: Girls' body satisfaction and perceptions of their attractiveness decrease during the teen years, while boys' increase or remain the same (Hargreaves & Tiggemann, 2002; Harter, 1990, 1993). This dissatisfaction continues into adulthood. Tiggemann and Rothblum (1997) found that the majority of women rated themselves as overweight and said they wanted to weigh approximately nine pounds less. In contrast, men rated themselves as average weight and wanted to weigh slightly more. Women also overestimate male preferences for slender female bodies; men's ideal figure is heavier than what women believed men's ideal was (Forbes, Adams-Curtis, Rade, & Jaberg, 2001). More than 90% anorexia and bulimia sufferers are female (American Psychiatric Association, 2000). Overall, the high societal standards for female appearance mean that many women will not attain the promoted ideal. Thus, it is likely that females will have lower appearance self-esteem than males.

Athletics

Although girls are now more likely to participate in sports than they once were, athletic activity is still more emphasized for boys. Even among children who play sports, boys have higher athletic self-esteem (e.g., Eccles, Wigfield, Harold, & Blumenfeld, 1993; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002). Boys also believe they are more competent in athletics (Bosacki, 2003; Klomsten, Skaalvik, & Espnes, 2004; Wu & Smith, 1997). An assessment of the aspects of physical self-concept found that boys scored higher than girls in eight of nine subdomains, including appearance, body fat, coordination, and endurance (Klomsten et al., 2004). The gender gap in athletic self-esteem widens during adolescence. Although physical self-concept decreases from elementary to secondary school for both genders, the drop is more pronounced for girls (Klomsten et al., 2004). Both theoretical models suggest that males will score higher than females on measures of athletic self-esteem; the competencies model favors males for their measurably higher performance in most athletic

domains, and reflected appraisals value these skills more in males than in females.

At first, it may seem paradoxical that males are held to higher standards in athletics and females in physical appearance, yet we predict higher self-esteem for males in athletics and lower self-esteem for females in physical appearance. By this assumption, self-esteem should be low when expectations are high. However, athletics and physical appearance, though both involving the body, have a crucial difference: Athletics is about doing, and physical appearance is about being looked at (Fredrickson & Roberts, 1997). One is active (and thus may lead to self-esteem and good mental health) and the other is passive, and thus may lead to lowered self-esteem through rumination and eventual depression (Nolen-Hoeksema, Larson, & Grayson, 1999). Especially for adolescents, athletic standards are also more attainable than the more unrealistic standards set for appearance (Fredrickson & Roberts, 1997).

Academics

Females perform better academically and receive better grades than their male peers (Pomerantz, Altermatt, & Saxon, 2002; Stetsenko, Little, Gordeeva, Grasshof, & Oetlingen, 2000). However, this does not always result in girls and women having higher academic self-esteem. When males outperform females in academics, female self-esteem suffers, but when females perform at a higher level, their self-esteem does not increase (Eccles et al., 1993; Hyde, Fennema, & Lamon, 1990; Jacobs et al., 2002; Stetsenko et al., 2000; Weiss, Kemmler, Deisenhammer, Fleischhacker, & Delazer, 2003). Even among academically gifted adolescents, females are more self-critical of their abilities (Luscombe & Riley, 2001). This may occur because teachers give less positive and more negative feedback to girls than to boys (Dweck & Leggett, 1988; Eccles & Blumenfeld, 1985). Thus, the competencies model predicts a female advantage in academic self-esteem, but the reflected appraisals model predicts a male advantage or no gender difference.

Social Acceptance

Friendships, peer relationships, and social approval are important for self-esteem (Leary & Downs, 1995). People with higher self-esteem have closer personal relationships (Lundgren & Rudawsky, 1998). However, high self-esteem may have more bearing on females' relationships. Josephs, Markus, and Tafarodi (1992) found that men, but not women, with high self-esteem differentiated themselves from others. As children, girls play in smaller groups and maintain more emotionally intimate and less competitive relationships than boys (Lever, 1978). In adolescence, females use more expressive pathways to increase intimacy with friends, whereas males use expressive and instrumental pathways equally (Radmacher & Azmitia, 2006). During adulthood, women report that their friendships are more focused on sharing information and communicating than men's friendships are (Sheets & Lugar, 2005).

However, girls' social relationships can also be problematic. Werner and Crick (2004) showed that girls reacted to social rejection by retaliating; boys did not. This creates an element of fragility in females' friendships. Benenson and Christakos (2003)

found that across elementary, middle school, and high school samples, girls had more previous friendships that had ended, and current friendships that were shorter than those of boys. When asked to imagine their closest friendships ending, girls believed they would be more distressed. In adulthood, women are more likely than men to respond to rejection cues by becoming more self-critical (Baldwin, Granzberg, Pippus, & Pritchard, 2003). Given that women focus on maintaining relationships, the competencies model would predict that they would have higher self-esteem in this area. However, the reflected appraisals model suggests that this advantage may be eliminated by the critical nature of some of these relationships. Thus, gender differences in social acceptance self-esteem may be influenced by opposing forces, with relationships causing both higher and lower self-esteem in girls and women. These influences might well cancel each other out, leading to null or small gender differences in social acceptance self-esteem.

Family

The family can act as a source of support and help affirm a child's beliefs about his or her self-worth. Familial relationships have a significant impact on female self-esteem and levels of depression, a result not seen in males (Colarossi & Eccles, 2000). However, this influence can have both positive and negative ramifications. Girls with strong familial relationships have higher levels of self-esteem and lower levels of depression; however, parents gave girls more negative feedback than they give boys (Lundgren & Rudawsky, 1998) even when their performances were the same (Lewis, Allesandri, & Sullivan, 1992). When considered together, this creates a detrimental combination for girls' self-esteem if they are simultaneously highly influenced and highly criticized. Reflected appraisals would predict a male advantage on this scale; because there are no established gender differences in closeness to family, the competencies model predicts no difference. Given this mixed picture, we predict null or very small gender differences in family self-esteem.

Behavioral Conduct

Behavioral conduct self-esteem measures an individual's perception of how socially acceptable his or her behavior is. People with high behavioral conduct self-esteem view their behavior as appropriate (Haynes, 1990). Behavioral conduct is particularly important in school. Haynes (1990) found that behavioral conduct self-esteem was a significant predictor of classroom behavior, group participation, and attitudes toward authority in middle school children. Girls generally perceive themselves as better behaved than boys do (e.g., Bosacki, 2003; Cole et al., 2001; Wu & Smith, 1997). Boys' problems tend to manifest in the form of externalizing disorders and girls' in internalizing disorders (Wicks-Nelson & Israel, 2003). Perhaps as a result, boys are more likely to be punished or reprimanded because of their behavior. Both competencies and reflected appraisals predict a female advantage, as girls and women have less problematic behavioral conduct and are recognized for this good behavior by others. Thus, the gender difference in behavior self-esteem is likely to favor girls and women.

Affect

Feeling happy, satisfied, and free from anxiety are all elements of emotional well-being. In contrast, negative emotions are related to depression and low self-esteem. In childhood, girls and boys are equally likely to be depressed, but in early adolescence female depression rates begin to rise and by late adolescence females are twice as likely as males to be depressed (Nolen-Hoeksema, 2001; Twenge & Nolen-Hoeksema, 2002). This incidence rate remains high throughout adulthood. This large gender difference has been attributed to many sources, including chronic strain, rumination, and unequal status at home and work (Nolen-Hoeksema, 1999). Because women experience more sexist incidents than men, they were subsequently more likely to report feeling angry, anxious, and depressed (Swim, Hyers, Cohen, & Ferguson, 2001). Furthermore, being a member of a devalued social group increases one's risk for emotional distress (Katz, Joiner, & Kwon, 2002). Women are also more likely to be anxious. Two meta-analyses found that women score higher in anxiety than men do, with an effect size around $d = 0.24$ (Feingold, 1994; Twenge, 2000). However, the gender difference in positive affect also favors women. When only positive emotions are measured, women report slightly greater happiness and life satisfaction than men (Wood, Rhodes, & Whelan, 1989). Women are also more gregarious, trusting, and tender-minded than men (Feingold, 1994). Thus, the gender difference in affect self-esteem might be null, because women's higher negative affect may be cancelled out by their higher positive affect. Both competencies and reflected appraisals recognize the gender difference in emotionality in both positive and negative affect and are thus in agreement with the prediction for no or only small gender differences.

Personal Self

Personal self-concept is a measure of one's evaluation of personality apart from the physical body or relationship to others (Haynes, 1990). Research on personal self-esteem is limited and conflicting. One study found that females score significantly higher than males (Gadzella & Williamson, 1984); however, personal self-esteem overlaps with global self-esteem, which usually favors males. This suggests that the gender difference will slightly favor males. Reflected appraisals also suggest a male advantage, because men are usually seen as more autonomous and assertive; the competencies model predicts no differences, because gender differences in agentic traits have disappeared (Twenge, 1997, 2001).

Self-Satisfaction

Self-satisfaction is also a measure of happiness with oneself as a person (Stringer, Reynolds, & Simpson, 2003). Women have been shown to have higher life satisfaction than men (Wood et al., 1989), but little research has investigated self-satisfaction. One study found that women score significantly higher on this subscale (Gadzella & Williamson, 1984). On the other hand, self-satisfaction, like personal self-esteem, overlaps considerably with global self-esteem. This suggests that the gender difference will slightly favor males. Similar to personal self, reflected appraisals would favor males, whereas competencies would predict no differences.

Moral–Ethical Self-Concept

Moral–ethical self-concept is a measure of one's perceptions of moral–ethical attributes and satisfaction with one's religion or lack of it (Haynes, 1990). Among Christian populations, women are more religious than men on every measure of religiosity (Walter & Davie, 1998). However, there is little gender difference in religiosity among Hindu, Jewish, or Muslim adherents (Loewenthal, MacLeod, & Cinnirella, 2002).

Women are also more likely to focus on caring for others when faced with moral dilemmas (Gilligan & Attanucci, 1988; Wark & Krebs, 1996). Females also exhibit more moral maturity than males (Wark & Krebs, 1996); however, no gender differences have been found for moral judgment stage scores (Lifton, 1985). Both competencies and reflected appraisals recognize that females value this area more. Thus, we expect that the gender difference in moral–ethical self-esteem will favor girls and women.

Overview

This meta-analysis aims to determine the size and significance of gender differences in specific domains of self-esteem. The small gender differences found in global self-esteem (Kling et al., 1999; Major et al., 1999; Twenge & Campbell, 2001) may be masking larger differences among specific domains. It is also not clear which specific facets favor men and which favor women, and how large any differences are. There may also be significant moderators of the effects. The current meta-analysis examines gender differences in self-esteem based on appearance, athletics, academics, social acceptance, family, behavioral conduct, affect, personal self, self-satisfaction, and moral-ethical self.

Moderator Variables

The meta-analytic design also allows us to examine moderator variables of gender differences in specific self-esteem. For example, does the gender difference in appearance self-esteem grow larger during adolescence? We can also examine changes over time to see whether gender differences in specific self-concept have changed from the 1970s to the present. In addition, because we are using several measures of specific self-concept, we can also examine how consistent gender differences are across the four measures.

Method

Measures of Self-Esteem

We focused on studies that used the four most widely used scales of multifaceted self-esteem: the Harter, the SDQ, the TSCS, and the Piers-Harris, each of which demonstrates high reliability and validity (Blascovich & Tomaka, 1991). Of course, reliability and validity differs by population. These scales appear to be equally valid for males and females, although they differ in age group targeted. The Harter, SDQ, and Piers-Harris scales are intended for children and are more valid for that age group, whereas the TSCS were intended for adolescents and adults (Blascovich & Tomaka, 1991). These scales are the most popular measures of specific self-esteem and thus have the most samples available for analysis. Of note, each of these scales assesses

physical appearance self-esteem, which we hypothesize to be one of the most important facets for gender differences.

Three of these scales (the SDQ, the TSCS, and the Piers-Harris) use a similar format, with respondents asked to note if they agree or disagree with simple first-person statements such as “I am good-looking,” “I am well-behaved in school,” “I am a member of a happy family,” “I am a good athlete,” or “I am satisfied with my moral behavior.” The Harter, intended for young children, instead uses pictures of children performing various activities—one who is good at the activity (“This boy isn’t very good at numbers”) and one who isn’t (“This boy is pretty good at numbers”). Children are then asked if they are good at that activity (“Are you: Not too good at numbers OR sort of good, pretty good, OR really good at numbers”).

Locating Studies

We searched the Web of Science by Thompson Scientific (previously known as the Science and Social Sciences Citation Index), a comprehensive database that includes both major and minor journals from all fields in science and social science. We searched for articles that cited the original sources for the four scales (Fitts, 1965; Harter, 1985; Marsh & O’Neill, 1984; Marsh, Relich, & Smith, 1983; Marsh, Smith, & Barnes, 1983; Piers, 1969, 1984; Piers & Harris, 1969; Roid & Fitts, 1988).

We also conducted a search for dissertations using the database ProQuest Dissertations and Theses. We searched for the keywords “Harter Self-Perception Profile,” “Self-Description Questionnaire,” “Tennessee Self-Concept Scale,” and “Piers-Harris.” An initial search yielded a prohibitively high number of dissertations. We examined a subsample of the available dissertations, locating those published in 1970, 1975, 1980, 1985, 1990, 1995, 2000, and 2005.

We included in the meta-analysis studies that met the following criteria: (1) the study did not preselect participants on any relevant variable (e.g., extreme depression scores, at-risk status, alcoholics, learning-disabled students, etc.); (2) the study involved at least 10 participants; and (3) the authors reported the domain-specific self-esteem scores broken down by gender. Finally, studies had to report data in such a way that an effect size could be calculated (see following section on the calculation of effect sizes).

Coding of Studies

Studies were coded for the following variables: (a) age or age group of respondents, (b) year of data collection, coded as two years prior to the date of publication unless another year was noted in the article (following Oliver & Hyde, 1993), (c) ethnicity/racial breakdown of the sample, and (d) measure of self-esteem used. Not enough studies reported data on socioeconomic status to code this variable. Most studies reported region, but there were not enough studies conducted outside the United States to analyze this variable. Studies reporting statistics separately for different age groups, self-esteem measures, or ethnic groups were coded as separate data points.

Final Sample of Studies

The search and review procedures led to a final sample of 115 articles and dissertations. These studies included 32,486 participants and yielded 428 effect sizes (most studies included multiple effect sizes, as the self-esteem measures all include multiple domains of self-esteem.) A table with a list of the studies included in the meta-analysis and their reference list can be obtained from either the first or second author.

Calculation of Effect Sizes

Formulas for the effect size, d , and homogeneity tests were taken from Hedges and Becker (1986). When means and SD s were available, the effect size was computed as the mean self-esteem score for males minus the mean self-esteem score for females, divided by the pooled SD . When means and SD s were not available, the effect size was calculated from reported t or F tests. When t or F was reported, d was calculated by using the formula provided by Hedges and Becker (1986).

Because effect sizes tend to be upwardly biased when based on small sample sizes, effect sizes were corrected for bias in the estimation of population effect sizes using the formula provided by Hedges (1981). All effect size analyses were weighted analyses (i.e., each effect size was weighted by an inverted variance; Hedges & Versa, 1998; Lipsey & Wilson, 2001).

To conduct the meta-analyses, we used *mixed-effects* models, which assume that effect size variance can be explained by both systematic and random components (Lipsey & Wilson, 2001). In mixed-effects models, certain identifiable study characteristics may act as moderator variables that are associated with systematic differences among effect sizes at the same time that a random component of residual variance remains after the systematic portion is accounted for. The mixed-effects model is preferable in this case because a fixed-effects model assumes that the only source of variation is from systematic variation, and the random effects model assumes none of the variation is from systematic sources. Mixed-effects models assume the effects of between-study variables are systematic, but that there is a remaining unmeasured random effect in the effect size distribution in addition to sampling error. As is done in random effects models, a random effects variance component (derived from the residual homogeneity value after the moderators are taken into account) is estimated and added to the standard error associated with each effect size and inverted variance weights are calculated.

Results

Mean Effect Sizes

Mean effect sizes were calculated for each of the self-esteem domains. The results are reported in Table 1. The number of samples (k), the weighted d , the 95% confidence interval for d , and the total homogeneity statistic (Q_T) for each self-esteem domain are reported. To address outlier data points, we eliminated effect sizes more than 2 SD s above or below the mean (Lipsey & Wilson, 2001). Out of 428 effect sizes, this procedure identified 20 total outliers that were eliminated from further analyses (4 for physical

Table 1
Summary of Main Effects for Gender Differences in Domain-Specific Self-Esteem

Self-esteem domain	<i>k</i>	<i>d</i>	95% CI	Q_T
Physical appearance	76	.35	.31, .40	288.07***
Athletic	68	.41	.36, .46	322.33***
Academic	75	.04	.00, .08	155.60***
Social acceptance	81	.04	-.02, .10	109.84*
Family	21	-.02	-.07, .04	20.41
Behavioral conduct	56	-.17	-.28, -.06	88.85**
Affect	17	.11	.04, .19	19.18
Personal self	9	.28	.11, .45	23.82**
Self-satisfaction	10	.33	.18, .49	58.78***
Moral-ethical	15	-.38	-.48, -.29	59.47***

Note. *k* = number of studies; *d* = effect size; CI = confidence interval; Q_T = total homogeneity. A positive *d* indicates that males scored higher; a negative *d* indicates that females scored higher.

* $p < .05$. ** $p < .01$. *** $p < .001$.

appearance, 3 for athletic, 3 for academic, 2 for social acceptance, 1 for family, 4 for behavioral conduct, 1 for affect, 1 for personal self, and 1 for self-satisfaction).

Significant mean effect sizes ranged from -0.17 to 0.41 (see Table 1). As expected, males scored higher than females on physical appearance, athletic, personal, and self-satisfaction self-esteem, but females scored significantly higher than males on behavioral conduct and moral-ethical self-esteem. Academic, affect, social acceptance, and family self-esteem did not show significant gender differences, because their 95% confidence intervals included zero (or nearly included zero, in the case of affect). Table 2 compares the predictions from the reflected appraisals and competencies models with the results of the analyses. Overall, the reflected appraisals model receives more support.

Only six out of the 10 domains included enough datapoints from dissertations to compare unpublished versus published data. Of these, four showed no significant differences by publication status. Unpublished dissertations produced a higher effect size ($d = 0.72$, $k = 4$) than published journal articles ($d = 0.33$, $k = 72$), $\chi^2(2) = 16.02$, $p < .001$, for physical appearance, and dissertations showed a higher effect size ($d = 0.70$, $k = 3$) than published articles ($d = 0.18$, $k = 7$), $\chi^2(2) = 8.91$, $p < .05$ for self-satisfaction. Thus, there were few differences by publication sta-

tus, and the higher effect sizes for dissertations argue against the possibility of a file-drawer problem.

Moderator Analyses

Physical appearance. Homogeneity analyses using procedures specified by Hedges and Becker (1986) and Lipsey and Wilson (2001) indicated that the set of 76 effect sizes was significantly heterogeneous, $Q_T = 288.07$, $\chi^2(75) = 118.60$, $p < .001$. The significant between-groups homogeneity statistics for age, $\chi^2(4) = 13.28$, $p < .01$, data collection year $\chi^2(3) = 16.27$, $p < .001$, and measure, $\chi^2(3) = 7.82$, $p < .05$ suggest that there is a significant difference in the magnitude of the effect sizes as a function of these moderator variables. As can be seen from Table 3, gender differences in physical appearance self-esteem were largest during adulthood. However, that effect size was based on only two samples. Of the age groups with more samples, the largest gender difference in appearance self-esteem occurred during early adolescence, consistent with the idea that the standards of others for a sexualized adult female appearance is at the root of the difference (Fredrickson & Roberts, 1997).

Time period also moderated the effect; there was not a significant gender difference during the 1970s, but beginning in the 1980s women scored significantly lower than men in physical appearance self-esteem. There are also slight differences based on measure, with the smallest differences on the Harter and the largest on the TSCS.

Athletic. Homogeneity analyses showed that the set of 68 effect sizes was significantly heterogeneous. As shown in Table 4, measure is the only significant moderator variable, with effect sizes for the Harter and the TSCS moderate and comparable but the effect size for the SDQ near zero.

Academic. The set of 75 effect sizes was significantly heterogeneous, $Q_T = 155.60$, $\chi^2(74) = 117.35$, $p < .001$. As shown in Table 5, age is a significant moderator. While there are no significant differences between males and females on academic self-esteem at most ages, a small difference emerges during adulthood favoring females.

Social acceptance. Homogeneity analyses indicated that the set of 81 effect sizes was significantly heterogeneous, $Q_T = 109.84$, $\chi^2(80) = 101.88$, $p < .05$. As shown in Table 6, year was a significant moderator, with the effect closer to zero during the 1980s and very small during the other eras.

Family. Homogeneity analyses indicated that the set of 21 effect sizes was not heterogeneous, $Q_T = 20.41$, $\chi^2(20) =$

Table 2
Comparison of the Model Predictions and Actual Outcomes

Self-esteem domain	Reflected appraisals model predictions	Competencies model predictions	Actual findings
Physical appearance	Males higher	Females higher/no difference	Males higher
Athletic	Males higher	Males higher	Males higher
Academic	Males higher/no difference	Females higher	No difference
Social acceptance	No difference	Females higher	No difference
Family	Males higher	No difference	No difference
Behavioral conduct	Females higher	Females higher	Females higher
Affect	No difference	No difference	No difference
Personal	Males higher	No difference	Males higher
Self-satisfaction	Males higher	No difference	Males higher
Moral-ethical	Females higher	Females higher	Females higher

Table 3
Moderating Variables in Physical Appearance Self-Esteem

Variable and class	Between-groups H	k	d	95% CI for d	Within-group H
Age group	17.13**				
Elementary school (ages 5–10)		24	0.30	0.22, 0.38	102.92***
Junior high (ages 11–13)		25	0.41	0.33, 0.49	57.47***
High school (ages 14–17)		14	0.30	0.19, 0.42	59.71***
College (ages 18–22)		8	0.25	0.10, 0.40	17.53*
Adult (ages 23–58)		5	0.73	0.51, 0.96	33.31***
Year	32.14***				
1970–1979		8	0.09	–0.09, 0.28	14.39*
1980–1989		22	0.33	0.25, 0.41	50.24***
1990–1999		34	0.38	0.31, 0.46	70.74***
2000–2006		11	0.32	0.20, 0.45	120.54***
Self-Esteem measure	9.15*				
Piers-Harris		10	0.39	0.24, 0.54	41.69***
Self-Description Questionnaire		13	0.33	0.23, 0.43	61.14***
Harter Self-Perception Profile		15	0.21	0.09, 0.32	70.09***
Tennessee Self-Concept Scale		38	0.41	0.34, .047	106.00

Note. k = number of studies; d = difference in terms of SDs ; CI = confidence interval; H = homogeneity. A positive d indicates that males scored higher; a negative d indicates that females scored higher.

* $p < .05$. ** $p < .01$. *** $p < .001$.

31.41, $p < .05$. Thus, no moderator variable analyses were conducted.

Behavioral conduct. The set of 56 effect sizes was significantly heterogeneous, $Q_T = 88.85$, $\chi^2(55) = 82.29$, $p < .01$. As shown in Table 7, age is a significant moderator, with the female advantage in behavioral conduct self-esteem increasing as youth get older. Year is also a moderator, with the female advantage increasing until the 1990s and then reversing in the 2000s.

Affect. Homogeneity analyses indicated that the set of 17 effect sizes was not heterogeneous, $Q_T = 19.18$, $\chi^2(16) = 26.30$, $p < .05$. Thus, no moderator variable analyses were conducted.

Personal self. Homogeneity analyses indicated that the set of 9 effect sizes was significantly heterogeneous, $Q_T = 23.82$, $\chi^2(8) = 20.09$, $p < .01$. However, none of the moderator variable analyses reached significance.

Self-satisfaction. The set of 10 effect sizes was significantly heterogeneous, $Q_T = 58.78$, $\chi^2(9) = 27.88$, $p < .001$. Not all groups had 2 or more studies, which restricted moderator analyses. It appears that gender differences in self-satisfaction self-esteem are largest during high school versus college and are larger in the most recent decade in comparison to the 1970s (see Table 8).

Moral-ethical. Homogeneity analyses indicated that the set of 15 effect sizes was significantly heterogeneous, $Q_T = 59.47$,

$\chi^2(14) = 36.12$, $p < .001$. The female advantage in moral-ethical self-esteem gets increasingly larger over time and is larger in studies that employ the SDQ rather than the Piers-Harris (see Table 9).

Overall. If we make the assumption that global self-esteem equals the sum of all domains of specific self-esteem, we can estimate the gender difference for global self-esteem here as $d = 0.10$. This is smaller than the effect size for measures of global self-esteem ($d = 0.15$: Kling et al., 1999), perhaps because appearance plays a disproportionate role in determining global self-esteem (Harter, 1993), especially among younger people. Thus, this estimate must be considered only a rough approximation, as it does not weight the domains of self-esteem.

Discussion

This study used meta-analytic techniques to investigate 428 effect sizes of gender differences in specific domains of self-esteem. Males scored significantly higher than females on physical appearance ($d = 0.35$), athletic ($d = 0.41$), personal self ($d = 0.28$), and self-satisfaction ($d = 0.33$) self-esteem. Females scored higher than males on behavioral conduct ($d = -0.17$) and moral-ethical self-esteem ($d = -0.38$). No significant gender dif-

Table 4
Moderating Variables in Athletic Self-Esteem

Variable and class	Between-groups H	k	d	95% CI for d	Within-group H
Self-Esteem measure	37.60***				
Self-Description Questionnaire		12	0.06	–0.05, 0.17	48.01***
Harter Self-Perception Profile		12	0.58	0.45, 0.71	70.49***
Tennessee Self-Concept Scale		44	0.48	0.41, 0.54	166.23***

Note. k = number of studies; d = difference in terms of SDs ; CI = confidence interval; H = homogeneity. A positive d indicates that males scored higher; a negative d indicates that females scored higher.

*** $p < .001$.

Table 5
Moderating Variables in Academic Self-Esteem

Variable and class	Between-groups <i>H</i>	<i>k</i>	<i>d</i>	95% CI for <i>d</i>	Within-group <i>H</i>
Age group	10.25*				
Elementary school (ages 5–10)		30	0.10	0.04, 0.16	74.92***
Junior high (ages 11–13)		30	0.01	–0.05, 0.07	50.26*
High school (ages 14–17)		8	0.00	–0.15, 0.15	1.91
College (ages 18–22)		5	–0.07	–0.24, 0.10	17.01**
Adult (ages 23–58)		2	–0.21	–0.45, 0.04	1.25

Note. *k* = number of studies; *d* = difference in terms of *SDs*; CI = confidence interval; *H* = homogeneity. A positive *d* indicates that males scored higher; a negative *d* indicates that females scored higher.

* $p < .05$. ** $p < .01$. *** $p < .001$.

ferences appeared in academic, social acceptance, family, and affect self-esteem. Most of the significant gender differences are moderate in size (between .20 and .40 *SDs*). Nevertheless, the differences in appearance, athletic, self-satisfaction, and moral–ethical self-esteem are larger than the established gender differences in verbal abilities, empathy, and adult aggression (Ashmore, 1990) and the White-Asian difference in self-esteem (Twenge & Crocker, 2002). Thus, the results challenge the idea that self-esteem differences between men and women are small. It is more accurate to say that the difference varies depending on the specific domain. Within some domains, such as physical appearance, athletic, and moral-ethical self-esteem, the gender differences are more than double those found with general self-esteem measures.

Overall, the results favor the reflected appraisals model. Both models made correct predictions in 4 domains (athletic, behavioral conduct, affect, moral-ethical); competencies were correct and reflected appraisals wrong in 1 (family); and reflected appraisals were right and competencies wrong in 5 (physical appearance, academics, social acceptance, personal, and self-satisfaction). If people evaluated their abilities based on objective competence, males would score higher on athletic self-esteem, females would score higher on academic, social acceptance, behavioral conduct, and moral-ethical self-esteem, and there would be no gender differences in the other domains. However, one of the largest gender differences was in physical appearance, a domain where competencies predicted no differences or even a female advantage but reflected appraisals correctly predicted a considerable male advantage. Competencies and reflected appraisals apparently canceled each other out in the domains of academics and social acceptance.

The male advantage in physical appearance self-esteem was significant at all ages but was most pronounced during adulthood. The gender gap did not increase consistently from childhood to

adulthood, but rather increased from childhood to junior high, then decreased throughout high school and college before rising again in adulthood. These results are consistent with research showing that female body satisfaction decreases during adolescence while males' stabilizes or increases (Hargreaves & Tiggemann, 2002; Harter, 1990, 1993). It is also consistent with studies showing that female body dissatisfaction persists during adulthood (Forbes et al., 2001; Tiggemann & Rothblum, 1997). Prior to adulthood, the largest gender gap occurs during junior high school. This may reflect the fact that females begin puberty earlier than males and thus may show more concern over their development compared with their male peers, whose development is more delayed. This is consistent with the traditional theory that lower self-esteem in females is related to the physical changes of puberty (Rosenberg, 1986).

The gender difference in appearance self-esteem was not significantly different from zero during the 1970s. After 1980, the difference rose to about a third of a *SD* and stayed there. This may have been caused by the increasing media focus on appearance during the 1980s and afterward. As Fredrickson and Roberts' (1997) objectification theory would predict, this may have led to lowered appearance self-esteem for women through reflected appraisals.

There were no significant gender differences on academic self-esteem. Given that females show higher academic performance, this is consistent with research showing that females may discount their academic abilities even when they excel (Eccles et al., 1993; Hyde et al., 1990; Jacobs et al., 2002; Pomerantz et al., 2002; Stetsenko et al., 2000; Weiss et al., 2003). In this case, actual competencies are washed out by reflected (or perhaps self-) appraisals.

Table 6
Moderating Variables in Social Acceptance Self-Esteem

Variable and class	Between-groups <i>H</i>	<i>k</i>	<i>d</i>	95% CI for <i>d</i>	Within-group <i>H</i>
Year	16.51***				
1970–1979		8	0.12	–0.11, 0.34	7.85
1980–1989		19	–0.01	–0.13, 0.11	20.28
1990–1999		36	0.07	–0.03, 0.16	34.16
2000–2006		17	0.09	–0.04, 0.22	31.04

Note. *k* = number of studies; *d* = difference in terms of *SDs*; CI = confidence interval; *H* = homogeneity. A positive *d* indicates that males scored higher; a negative *d* indicates that females scored higher.

*** $p < .001$.

Table 7
Moderating Variables in Behavioral Conduct Self-Esteem

Variable and class	Between-groups <i>H</i>	<i>k</i>	<i>d</i>	95% CI for <i>d</i>	Within-group <i>H</i>
Age group	10.09*				
Elementary school (ages 5–10)		19	–0.10	–0.29, 0.09	47.13***
Junior high (ages 11–13)		26	–0.23	–0.39, –0.08	25.27
High school (ages 14–17)		7	–0.25	–0.55, 0.06	5.62
College (ages 18–22)		3	–0.27	–0.73, 0.20	0.74
Year	14.18**				
1970–1979		7	–0.09	–0.42, 0.24	3.42
1980–1989		9	–0.12	–0.37, 0.13	3.34
1990–1999		30	–0.30	–0.45, –0.15	32.21
2000–2006		9	0.22	–0.05, 0.49	35.70

Note. *k* = number of studies; *d* = difference in terms of *SDs*; CI = confidence interval; *H* = homogeneity. A positive *d* indicates that males scored higher; a negative *d* indicates that females scored higher.

* *p* < .05. ** *p* < .01. *** *p* < .001.

Females scored significantly higher on behavioral conduct self-esteem. This reflects research showing that females behave better (e.g., Bosacki, 2003; Cole et al., 2001; Wu & Smith, 1997), whereas males tend to act out more (Wicks-Nelson & Israel, 2003). This difference grows as children grow older, suggesting that experience in school reinforces these beliefs. Moral–ethical self-esteem was significantly higher in females, supporting previous research finding greater religiosity among women (at least in Christian populations). The gender difference is largest during high school; thus females may mature faster in their morality than males (Wark & Krebs, 1996). The difference has grown over the decades; there was no significant gender difference in the 1970s, but the effect size now exceeds half a *SD*.

Males scored significantly higher than females on the personal self and self-satisfaction subscales. Personal self and self-satisfaction are similar to global self-esteem (they all measure happiness with oneself as a person). The differences here are higher than those in the previous meta-analyses on global self-esteem, but are very close to the meta-analytic effect sizes for adolescents, who comprised the majority of the samples on these domains in this meta-analysis.

Implications

An influential review (Baumeister, Campbell, Kruger, & Vohs, 2003) found that global self-esteem was linked to happiness, but had few benefits for academic achievement, work performance, or healthy behaviors. In contrast, domain-specific measures of self-esteem are consistently correlated with performance within that domain, apparently in a reciprocal process in which each causes the other (Marsh & Craven, 2006). These results are explained by the specificity matching hypothesis, which contends that the most meaningful linkages are between attitudes and behavior within the same domain. Nevertheless, many interventions for girls and women conflate domain-specific self-esteem with general self-esteem (e.g., the Dove Self-Esteem Fund, which opines that “Too many girls develop low self-esteem from hang-ups about looks and, consequently, fail to reach their full potential in later life.”) Research on global vs. domain-specific self-esteem suggests instead that most girls with “low self-esteem from hangups about looks” will do just fine in reaching their potential (e.g., academically), but may continue to have low appearance self-esteem and might be at risk for eating disorders.

Table 8
Moderating Variables in Self-Satisfaction Self-Esteem

Variable and class	Between-groups <i>H</i>	<i>k</i>	<i>d</i>	95% CI for <i>d</i>	Within-group <i>H</i>
Age group	50.32***				
High school (ages 14–17)		4	0.36	0.13 0.59	3.63
College (ages 18–22)		4	0.05	–0.21 0.30	4.83
Year	14.20***				
1970–1979		7	0.11	–0.08 0.31	6.15
2000–2006		2	0.81	0.46 1.16	38.43

Note. *k* = number of studies; *d* = difference in terms of *SDs*; CI = confidence interval; *H* = homogeneity. A positive *d* indicates that males scored higher; a negative *d* indicates that females scored higher.

*** *p* < .001.

Table 9
Moderating Variables in Moral–Ethical Self-Esteem

Variable and class	Between-groups <i>H</i>	<i>k</i>	<i>d</i>	95% CI for <i>d</i>	Within-group <i>H</i>
Year	44.81***				
1970–1979		5	0.01	–0.19, 0.21	6.01
1980–1989		5	–0.40	–0.57, –0.24	8.40
1990–1999		3	–0.68	–0.91, –0.46	0.25
Self-esteem measure	12.47**				
Piers-Harris		6	–0.32	–0.48, –0.17	34.46***
Self-Description Questionnaire		8	–0.50	–0.63, –0.37	12.54

Note. *k* = number of studies; *d* = difference in terms of *SDs*; CI = confidence intervals; *H* = homogeneity. A positive *d* indicates that males scored higher; a negative *d* indicates that females scored higher.

** *p* < .01. *** *p* < .001.

As appearance is an area in which “performance” is difficult to measure (and looking good may have come at the expense of an eating disorder) the consequences of low appearance self-esteem are more difficult to quantify than the consequences of, for example, low academic self-esteem. Outcomes linked to low appearance self-esteem may be fueled by mediators such as objectification (Fredrickson & Roberts, 1997) and other risk factors for eating disorders. Thus any interventions designed to help girls and women avoid the negative consequences of low appearance self-esteem should be focused specifically on appearance and not on general self-esteem. An even more expeditious route might be to bypass appearance self-esteem and directly target the root causes of negative outcomes such as eating disorders. Using the appraisals model as a basis, such interventions might help girls and women develop a more objective view of how others see them rather than relying on media images as universal standards. This may be more effective than self-esteem boosting efforts that fail to account for this important social context. “You’re beautiful just for being you” falls on deaf ears if girls believe that others will see them as beautiful only if they meet an impossible standard.

Measurement Effects and Limitations

The effect sizes were fairly consistent across the self-esteem measures used for most of the subscales, with few exceptions. A large gender difference favoring males was found on the athletic self-esteem subscale of the Harter and TSCS, compared with a difference of near zero on the SDQ. However, the SDQ showed a larger female advantage on moral–ethical self-esteem than the Piers-Harris. One limitation is that age and measure were confounded, because adult samples completed the TSCS and child samples the other measures. In general, the biggest limitation of this analysis is the small number of samples from adult populations; the majority of the data here come from children and adolescents. Future research should explore how gender differences in the domains of self-esteem change during the course of adult life.

Conclusions

The current meta-analysis found that gender differences vary widely across the subdomains of self-esteem, some showing no difference at all, and others with gender differences in the mod-

erate range. The differences obtained in this study were moderated by variables such as age, year of data collection, and measure. Like the meta-analyses of global self-esteem, these analyses do not find extremely large gender differences in self-esteem. However, many of the gender differences in specific domains of self-esteem are considerably larger than the *d* = 0.15 difference in global self-esteem. In some areas, like that of athletic ability, these gender differences reflect actual gender differences in competence and performance. In other areas, such as physical appearance, women’s lower self-esteem derives not from actual deficits but from the more critical reflected appraisals of others—including the larger “other” of idealized media images.

References

- Allgood-Merten, B., Lewinsohn, P. M., & Hops, H. (1990). Sex differences and adolescent depression. *Journal of Abnormal Psychology, 99*, 55–63.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.
- 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.
- Baldwin, M. W., Granzberg, A., Pippus, L., & Pritchard, E. T. (2003). Cued activation of relational schemas: Self-evaluation and gender effects. *Canadian Journal of Behavioural Science, 35*, 153–163.
- Bandura, A. (1989). Regulation of cognitive processes through perceived self-efficacy. *Developmental Psychology, 25*, 729–735.
- Baumeister, R. F. (1998). The self. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 680–740). New York: McGraw-Hill.
- Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest, 4*, 1–44.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*, 497–529.
- Benenson, J. F., & Christakos, A. (2003). The greater fragility of females’ versus males’ closest same-sex friendships. *Child Development, 74*, 1123–1129.
- Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology, 66*, 5–20.
- Blascovich, J., & Tomaka, J. (1991). Measures of self-esteem. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 115–160). San Diego, CA: Academic Press.

- Bosacki, S. L. (2003). Psychological pragmatics in preadolescents: Sociomoral understanding, self-worth, and school behavior. *Journal of Youth and Adolescents*, 32, 141–155.
- Colarossi, L. G., & Eccles, J. S. (2000). A prospective study of adolescents' peer support: Gender differences and the influence of parental relationships. *Journal of Youth and Adolescence*, 29, 661–678.
- Cole, D. A., Maxwell, S. E., Martin, J. M., Peeke, L. G., Seroczynski, A. D., Tram, J. M., et al. (2001). The development of multiple domains of child and adolescent self-concept: A cohort sequential longitudinal design. *Child Development*, 72, 1723–1746.
- Cooley, C. H. (1902). *Human nature and the social order*. New York: Scribner's.
- Denissen, J. J. A., Penke, L., & Schmitt, D. P. (2008). Self-esteem reactions to social interactions: Evidence for sociometer mechanisms across days, people, and nations. *Journal of Personality and Social Psychology*, 95, 181–196.
- Dweck, C. S., & Leggett, E. L. (1988). A social cognitive approach to motivation and personality. *Psychological Review*, 95, 256–273.
- Eccles, J., & Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, 64, 830–847.
- Eccles, J. S., & Blumenfeld, P. (1985). Classroom experiences and student gender: Are there differences and do they matter? In L. C. Wilkinson & C. B. Marrett (Eds.), *Gender influences in the classroom* (pp. 79–113). New York: Academic Press.
- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin*, 116(3), 429–456.
- Fitts, W. H. (1965). *Manual for the Tennessee Self Concept Scale*. Nashville, TN: Counselor Recordings and Tests.
- Forbes, G. B., Adams-Curtis, L. E., Rade, B., & Jaberg, P. (2001). Body dissatisfaction in women and men: The role of gender-typing and self-esteem. *Sex Roles*, 44, 461–484.
- Fredrickson, B. L., & Roberts, T. A. (1997). Objectification theory: Toward understanding women's lived experiences and mental health risks. *Psychology of Women Quarterly*, 21, 173–206.
- Gadzella, B. M., & Williamson, J. D. (1984). Differences between men and women on selected Tennessee Self-Concept Scales. *Psychological Reports*, 55, 939–942.
- Gilligan, C., & Attanucci, J. (1988). Two moral orientations: Gender differences and similarities. *Merrill-Palmer Quarterly*, 34(3), 223–237.
- Hargreaves, D., & Tiggemann, M. (2002). The role of appearance schematicity in the development of adolescent body dissatisfaction. *Cognitive Therapy and Research*, 26, 691–700.
- Harter, S. (1985). *Manual for the self-perception profile for children*. Denver, CO: University of Denver.
- Harter, S. (1990). Causes, correlates and the functional role of global self-worth: A life Span perspective. In R. J. Sternberg & J. Kolligian, Jr. (Eds.), *Competence considered* (pp. 67–98). New Haven, CT: Yale University Press.
- Haynes, N. M. (1990). Influence of self-concept on school adjustment among middle School students. *Journal of Social Psychology*, 130, 199–208.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107–128.
- Hedges, L. V., & Becker, B. J. (1986). Statistical methods in the meta-analysis of research in 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., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, 3, 486–504.
- Hyde, J. S., Fennema, E., & Lamon, J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107, 139–155.
- Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, 73, 509–527.
- James, W. (1890). *The principles of psychology* (Vol. 1). New York: Holt.
- Josephs, R. A., Markus, H. R., & Tafarodi, R. W. (1992). Gender and self-esteem. *Journal of Personality and Social Psychology*, 63, 391–402.
- Katz, J., Joiner Jr., T. E., & Kwon, P. (2002). Membership in a devalued social group and emotional well-being: Developing a model of personal self-esteem, collective self-esteem, and group socialization. *Sex Roles*, 47, 419–431.
- 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.
- Klomsten, A. T., Skaalvik, E. M., & Espnes, G. A. (2004). Physical self-concept and sports: Do gender differences still exist? *Sex Roles*, 50, 119–127.
- Leary, M. R., & Downs, D. L. (1995). Interpersonal functions of the self-esteem motive: The self-esteem system as a sociometer. In M. H. Kernis (Ed.), *Efficacy, agency, and self-esteem* (pp. 123–144). New York: Plenum Press.
- Leary, M. R., Haupt, A. L., Strausser, K. S., & Chokel, J. T. (1998). Calibrating the sociometer: The relationship between interpersonal appraisals and the state self-esteem. *Journal of Personality and Social Psychology*, 74, 1290–1299.
- Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. (1995). Self-esteem as an interpersonal monitor: The sociometer hypothesis. *Journal of Personality Psychology*, 68, 518–530.
- Lever, J. (1978). Sex differences in the complexity of children's play and games. *American Sociological Review*, 43, 471–483.
- Lewis, M., Allesandri, S. M., & Sullivan, M. W. (1992). Differences in shame and pride as a function of children's gender and task difficulty. *Child Development*, 63, 630–638.
- Lifton, P. D. (1985). Individual differences in moral development: The relation of sex, gender, and personality to morality. *Personality*, 53, 306–334.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Loewenthal, K. M., MacLeod, A. K., & Cinnirella, M. (2002). Are women more religious than men? Gender differences in religious activity among different religious groups in the UK. *Personality and Individual Differences*, 32, 133–139.
- Lundgren, D. C., & Rudawsky, D. J. (1998). Female and male college students' responses to negative feedback from parents and peers. *Sex Roles*, 39, 409–429.
- Luscombe, A., & Riley, T. L. (2001). An examination of self-concept in academically gifted adolescents: Do gender differences occur? *Roeper Review*, 24, 20–22.
- Major, B., Barr, L., Zubek, J., & Babey, S. H. (1999). Gender and self-esteem: A meta-analysis. In W. B. Swann Jr., 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.
- Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, 1, 133–163.
- Marsh, H. W., & O'Neill, R. (1984). Self-description questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement*, 24, 153–174.
- Marsh, H. W., Smith, I. D., & Barnes, J. (1983). Multitrait-multimethod analyses of the self-description questionnaire: Student-teacher agree-

- ment on multidimensional ratings of student self-concept. *American Education Research Journal*, 20, 333–357.
- Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.
- Nolen-Hoeksema, S. (1999). Gender differences in depression. In C. Hammen & I. Gotlib (Eds.), *Handbook of depression*. NY: Guilford.
- Nolen-Hoeksema, S. (2001). Gender differences in depression. *Current Directions in Psychological Science*, 10, 173–176.
- Nolen-Hoeksema, S., Larson, J., & Grayson, C. (1999). Explaining the gender difference in depressive symptoms. *Journal of Personality and Social Psychology*, 77, 1061–1072.
- Oliver, M. B., & Hyde, J. S. (1993). Gender differences in sexuality: A meta-analysis. *Psychological Bulletin*, 114, 29–51.
- Piers, E. V. (1969). *Manual for the Piers-Harris Children's Self-Concept Scale*. Nashville, TN: Counselor Recordings and Tests.
- Piers, E. V. (1984). *Revised manual for the Piers-Harris Children's Self-Concept Scale*. Los Angeles: Western Psychological Services.
- Piers, E. V., & Harris, D. B. (1969). *Piers-Harris Children's Self-Concept Scale*. Los Angeles: Western Psychological Services.
- Pipher, M. (1994). *Reviving Ophelia: Saving the selves of adolescent girls*. New York: Ballantine Books.
- Polce-Lynch, M., Myers, B. J., Kilmartin, C. T., Forssmann-Falck, R., & Kliewer, W. (1998). Gender and age patterns in emotional expression, body image, and self-esteem: A qualitative analysis. *Sex Roles*, 38, 1025–1048.
- Pomerantz, E. M., Altermatt, E. R., & Saxon, J. L. (2002). Making the grade but feeling distressed: Gender differences in academic performance and internal distress. *Journal of Educational Psychology*, 94, 396–404.
- Radmacher, K., & Azmitia, M. (2006). Are there gendered pathways to intimacy in early adolescents' and emerging adults' friendships? *Journal of Adolescent Research*, 21, 415–448.
- Roid, G. H., & Fitts, W. H. (1988). *Tennessee self-concept scale (TSCS), revised manual*. Los Angeles: Western Psychological Services.
- Rosenberg, M. (1986). Self-concept from middle childhood through adolescence. In J. Suls & A. G. Greenwald (Eds.), *Psychological perspectives on the self* (pp. 107–136). Hillsdale, NJ: Erlbaum.
- Sheets, V. L., & Lugar, R. (2005). Friendship and gender in Russia and the United States. *Sex Roles*, 52(1/2), 131–140.
- Sondhaus, E. L., Kurtz, R. M., & Strube, M. J. (2001). Body attitude, gender, and self-concept: A 30-year perspective. *The Journal of Psychology*, 135, 413–429.
- Stetsenko, A., Little, T. D., Gordeeva, T., Grasshof, M., & Oettingen, G. (2000). Gender effects in children's beliefs about school performance: A cross-cultural study. *Child Development*, 71, 517–527.
- Stringer, S. J., Reynolds, G. P., & Simpson, F. M. (2003). Collaboration between Classroom teachers and a school counselor through literature circles: Building self-esteem. *Journal of Instructional Psychology*, 30, 69–76.
- Swim, J. K., Hyers, L. L., Cohen, L. L., & Ferguson, M. J. (2001). Everyday sexism: Evidence for its incidence, nature, and psychological impact from three daily diary studies. *Journal of Social Issues*, 57, 31–53.
- Tiggemann, M., & Rothblum, E. D. (1997). Gender differences in internal beliefs about weight and negative attitudes towards self and others. *Psychology of Women Quarterly*, 21, 581–593.
- Twenge, J. M. (1997). Changes in masculine and feminine traits over time: A meta-analysis. *Sex Roles*, 36, 305–325.
- Twenge, J. M. (2000). The age of anxiety? Birth cohort change in anxiety and neuroticism, 1952–1993. *Journal of Personality and Social Psychology*, 79, 1007–1021.
- Twenge, J. M., & Campbell, W. K. (2001). Age and birth cohort differences in self-esteem: A cross-temporal meta-analysis. *Personality and Social Psychology Review*, 5, 321–344.
- Twenge, J. M., & Crocker, J. (2002). Race and self-esteem: Meta-analysis comparing Whites, Blacks, Hispanics, Asians, and American Indians. *Psychological Bulletin*, 128, 371–408.
- 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.
- Walter, T., & Davie, G. (1998). The religiosity of women in the modern west. *British Journal of Sociology*, 49, 640–660.
- Wark, G. R., & Krebs, D. L. (1996). Gender and dilemma differences in real-life moral judgment. *Developmental Psychology*, 32, 220–230.
- Weiss, E. M., Kemmler, G., Deisenhammer, E. A., Fleischhacker, W. W., & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and Individual Differences*, 35, 863–875.
- Werner, N. E., & Crick, N. R. (2004). Maladaptive peer relationships and the development of relational and physical aggression during middle childhood. *Social Development*, 13, 495–514.
- Wertheim, E. H., Paxton, S. J., Shultz, H. K., & Muir, S. L. (1997). Why do adolescent girls watch their weight? An interview study examining sociocultural pressures to be thin. *Journal of Psychosomatic Research*, 42, 345–355.
- Wicks-Nelson, R., & Israel, A. C. (2003). *Behavior disorders of childhood*. Upper Saddle River, NJ: Prentice Hall.
- 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.
- Wu, Y., & Smith, D. E. (1997). Self-esteem of Taiwanese children. *Child Study Journal*, 27, 1–19.

Received December 10, 2007

Revision received July 21, 2008

Accepted July 22, 2008 ■