Psychological Bulletin
1996, Vol. 119, No. 3, 422-447

0033-2909/96/$3.00

Gender Differences in Aggression as a Function of Provocation: A Meta-Analysis

B. Ann Bettencourt
University of Missouri—Columbia

Norman Miller
University of Southern California

In this article, we meta-analytically examine experimental studies to assess the moderating effect of provocation on gender differences in aggression. Convergent evidence shows that, whereas unprovoked men are more aggressive than women, provocation markedly attenuates this gender difference. Gender differences in appraisals of provocation intensity and fear of danger from retaliation (but not negative affect) partially mediate the attenuating effect of provocation. However, they do not entirely account for its manipulated effect. Type of provocation and other contextual variables also affect the magnitude of gender differences in aggression. The results support a social role analysis of gender differences in aggression and counter A. H. Eagly and V. Steffen's (1986) meta-analytic inability to confirm an attenuating effect of provocation on gender differences in aggression.

Many experimental studies of adult aggression show that men are more aggressive than women (for reviews, see Eagly & Steffen, 1986; Frodi, Macaulay, & Thome, 1977; Hyde, 1984; and White, 1983). Earlier literature reviews (e.g., Maccoby & Jacklin, 1974; Terman & Tyler, 1954) tend to emphasize biological contributions to this difference more strongly than do later ones (e.g., Eagly & Steffen, 1986; Frodi et al., 1977; Hyde, 1984), and in more recent reviews, some scholars seriously question whether biology plays an important role in human aggression (e.g., Adams, 1992; Benton, 1992). Whatever role biological factors play, contemporary theorists argue that gender roles and cultural norms contribute to gender differences in aggression (e.g., Bandura, 1973; B. A. Baron & Richardson, 1994; Berkowitz, 1989; Dollard, Doob, Miller, Mowrer, & Sears, 1939; Eagly & Steffen, 1986; Lightdale & Prentice, 1994; Maccoby & Jacklin, 1974; Zillman, 1979). In their meta-analysis on gender differences in aggression, Eagly and Steffen used a social role framework for explaining gender differences in aggression. They suggest that aggression can be viewed as a behavior dictated by gender roles. That is, gender differences in aggression reflect differences in normative expectations that society holds for men and women. Eagly and Steffen noted, "the male gender role includes norms encouraging many forms of aggression . . . [but] the traditional female gender role places little emphasis on aggressiveness" (p. 310). Accordingly, they argued that men tend to be more aggressive than women, unless situational features make gender role considerations less salient.

Research also emphasizes the importance of provocation in eliciting aggression (e.g., Berkowitz, 1989; Carlson & Miller, 1988; Dollard et al., 1939). In social interaction, the tit-for-tat rule or reciprocity norm exerts powerful effects (Gouldner, 1960) and often results in an escalation of conflict and hostility (Axelrod, 1984; Dodge & Coie, 1987). In consonance with the tit-for-tat rule, provocation provides justification for aggression. This unfolding dynamic of justified aggression may be of particular importance for women because it frees them from the constraints ordinarily evoked by gender role norms. However, researchers who report gender differences in aggression tend to use experimental paradigms in which participants are not provoked (Lando, Johnson-Payne, Gilbert, & Deutsch, 1977). Yet, when the experimental paradigm includes a provocation before their opportunity to aggress, gender differences often are not found (e.g., Ahmed, 1982; Dor-Shav & Dolgin, 1981; Frodi, 1978; Golin & Romanowski, 1977; Turner, Layton, & Simons, 1975). Moreover, in several studies that vary both gender of participant and level of provocation, women exhibit less aggression than men under relatively neutral conditions but behave slightly more aggressively than men when provoked (e.g., Anderson, 1993; Fischer, Kelm, & Rose, 1969; Schuck, Schuck, Hallam, Mancini, & Wells, 1971; Taylor & Epstein, 1967). Indeed, Frodi et al. (1977) reported an absence of gender differences in aggression in 61% of the studies included in their qualitative review and noted that, among those studies in which participants were provoked, few reported gender differences. Taken together,

A preliminary report of this research was presented at the meeting of the Midwestern Psychological Association, May 1993. Partial support of the preparation of this article was provided by an American Psychological Asssociation Minority Fellowship and by the Center for Research in Social Behavior at the University of Missouri—Columbia. Preparation of this article was also facilitated by National Science Foundation Grants BSN-8719439 and SBR 9319752.

We thank Lisa Allard, Mathew Fink, Alan Kelly, and Andrea Mandelblatt for their assistance in providing the judgments of key theoretical variables on the basis of the excerpted materials from the method sections; Kristina Bowman for her assistance in coding the categorical variables; and Stephanie Black for her assistance in conducting the computer literature searches. We also thank Li-Fei Wang for providing calculations of the effect size estimates. Finally, we thank Craig Anderson, Kathryn Anderson, Nancy Collins, Tiffany Ito, and especially Alice Eagly for insightful comments on drafts of this article.

Correspondence concerning this article should be addressed to B. Ann Bettencourt, Department of Psychology, University of Missouri, 210 McAlester Hall, Columbia, Missouri 65211, or to Norman Miller, Department of Psychology, University of Southern California, Los Angeles, California 90089-1061. Electronic mail may be sent via Internet to psyannb@mizzou1.missouri.edu (Bettencourt).

these findings suggest that the gender differences typically found in the absence of provocation are attenuated under high provocation and, consequently, differences in aggression between the provocation and neutral conditions of experimental studies are larger for women than for men.

The relationship between provocation and gender differences in aggression was one effect examined in Eagly and Steffen's (1986) comprehensive meta-analysis of experimental research on aggression. The approach they adopted was methodologically sound. Specifically, they dichotomized studies into those using minimal provocation (impeding progress on a task) or maximal provocation (verbal insult, assignment of an impossible task, or physical attack). Although their analysis did not confirm any moderating effect of provocation, we meta-analytically reexamined this specific relationship for several reasons. First, the theoretical basis for the expected interaction between provocation and gender of aggressor remains compelling (Deaux, 1984; Frodi et al., 1977). Second, Eagly and Steffen's test of the effects of provocation may not have been sufficiently sensitive.

Two alternative approaches may improve sensitivity. A first approach is to meta-analytically compare the gender differences in aggression observed under relatively neutral conditions with those observed under provocation conditions. Instead of meta-analytically comparing them in this way, Eagly and Steffen (1986) averaged the gender differences reported for the neutral and provocation conditions and then compared the average gender difference effect size for studies categorized as inducing minimal provocation with those categorized as inducing greater than minimal provocation (p. 313). They adopted this conservative procedure because it avoids the potential problem of interdependency among the several possible effect size estimates that may be derived from a single study. To explore the potential appropriateness of a meta-analytic comparison of effect sizes derived from those studies that had both a provocation and a neutral–control condition, we explicitly examined our dataset for problems of interdependency by using Tukey's Jackknife technique (cf. Glass, McGaw, & Smith, 1981). Because this analysis showed that dependency was not a problem for our dataset, we computed effect sizes that reflected the magnitude by which male aggressiveness exceeded female aggressiveness and meta-analytically compared the average of these effect sizes obtained in the neutral conditions with that found in the provocation conditions. In addition to these gender difference effect sizes, we also computed effect sizes that reflected the magnitude by which aggression under provocation conditions exceeded that under neutral conditions and meta-analytically compared the average of these provocation difference effect sizes for female participants with that for male participants. Clearly, these two procedures reflect two sides of the same coin and should yield conceptually parallel results. Yet, each should be more sensitive than Eagly and Steffen's approach.

A second approach to increasing meta-analytic sensitivity is to take advantage of the continuous nature of the between-study variation in levels of experimentally induced provocation. Instead of a simple dichotomization of studies in terms of minimal versus maximal provocation (e.g., Eagly & Steffen, 1986), level of provocation can be treated as a continuous variable by having judges rate the intensity of provocation that participants would perceive in the various conditions of each experimental study (Carlson, Marcus-Newhall, & Miller, 1989; Carlson & Miller, 1988; Miller, Lee, & Carlson, 1991). Then the judges' ratings of perceived provocation intensity (averaged across judges) can be treated as a continuous predictor of gender difference effect sizes.

In summary, given the absence of meta-analytic evidence that confirms the strong theoretical expectation of an attenuation of gender differences in aggression as a function of increased provocation, we applied analytical procedures designed to increase the sensitivity of our meta-analytic assessment of this interaction. In addition, our major analyses are based on a sample size that exceeds that of Eagly and Steffen (1986) by about 33%, bringing greater power to our analyses.

Gender Differences in Perceptions of Provocation Intensity, Negative Affect, and Fear of Danger

Meta-analysis can contribute to the theoretical understanding of relationships between variables (represented by effect sizes), by an examination of the effects of mediating as well as moderating variables on these relationships (Cooper & Lemke, 1991; Eagly & Wood, 1994; Hall & Rosenthal, 1991; Hall, Rosenthal, Tickle-Degnen, & Mosteller, 1994; Miller & Pollock, 1994, in press; Mullen, Salas, & Miller, 1991). The two approaches that we adopted, as described in the preceding section, are used to examine the moderating effect of provocation on gender differences in aggression. If provocation is an important moderator of gender differences, then we expect the average male versus female difference in aggression to be smaller under conditions of provocation than under neutral conditions. Similarly, we anticipate that the differences between aggression under provocation and neutral conditions will be larger for female than male participants. Finally, in consonance with these expectations, regression analyses of continuous assessments of provocation intensity, as measured by judges ratings, should yield further confirming evidence that provocation moderates gender differences in aggression.

A second major purpose of our meta-analysis, however, was to examine theoretical variables thought to mediate gender differences in aggression. Berkowitz (1989) argued that aggression is a direct result of the negative affect elicited by antecedent aversive events (e.g., provocations) and that, consequently, people are most aggressive when events elicit strong negative affect. In other words, subjectively experienced negative affect and provocation intensity should mediate aggressive responding. If so, then gender differences in aggression should be mediated by gender differences in subjective appraisals of negative affect and provocation intensity. In accordance with this expectation, research (e.g., Harris, 1993; Lohr, Hamberger, & Bonge, 1988) suggests that men do perceive some behaviors as more provoking than do women, and vice versa.

Thus, extrapolating from the theorizing of Berkowitz (1989), we reasoned that the difference between male and female judges' ratings of the level of negative affect that would be elicited by the respective conditions of each study should provide an index of the differences in subjective negative affect experienced by the male and female participants in the experimental studies. These gender differences, in turn, should mediate gen-

der differences in aggression (Berkowitz, 1989). Likewise, the difference between female and male judges' appraisals of the subjective degree of provocation experienced by the participants of each study should provide an index of gender differences in levels of subjectively experienced provocation intensity. We expected that these gender differences in appraisals of provocation intensity also would mediate gender differences in aggression.

Generally, those who have reason to fear that their aggressive acts will bring about retaliation are more likely to control their aggression (Berkowitz, 1988; Dollard et al., 1939; Rogers, 1980). Although Frodi et al. (1977) minimized its importance, Eagly and Steffen (1986) argued that gender differences in anticipated danger of retaliation from the person who may experience one's aggressive action importantly affects the magnitude of gender differences in aggression. In Eagly and Steffen's meta-analysis, although provocation did not directly attenuate gender differences in aggression, female judges who read relevant information from the method section of each study did anticipate more danger than male judges, and larger gender differences in perceived danger were associated with larger gender differences in aggression. Thus, based on these findings, as well as the predictions of the social role model, we also assessed the mediating role of differential perceptions of danger from retaliation. Specifically, we also expected differences between female and male judges' appraisals of danger from retaliation to predict gender differences in aggression.

Contextual Variables That Moderate Gender Differences in Aggression

In addition to providing a framework for understanding the effects of provocation, the social role model suggests that other contextual variables function as moderators of gender differences in aggression (Eagly, 1987). Consistent with this view, Deaux and Major (1987) highlighted the flexibility of gender-related behaviors and emphasize the importance of considering immediate situations when studying the effects of gender roles on gender differences in aggression. In the paragraphs that follow, we discuss variables that have been thought to moderate the magnitude of gender differences in aggression and the degree to which these differences are attenuated by provocation.

The type of aggressive response made available to participants appears to determine the magnitude of gender differences (Eagly & Steffen, 1986; Frodi et al., 1977; White, 1983). Following the reasoning of social role explanations, levels of aggression in men and women may differ because particular types of aggressive behaviors are considered more appropriate for one gender than for the other (Deaux & Major, 1987; Harris, 1991). Physical aggression may be deemed more appropriate for men, whereas verbal aggression may be considered comparatively more appropriate for women. Although several specific studies in which was measured either physical (e.g., Buss, 1963; Hynan, Harper, Wood, & Kallas, 1980) or verbal aggression (Berkowitz, 1960; Berkowitz & Holmes, 1959; Stagner & Gongdon, 1975) did not find gender differences, meta-analytic integration of the literature did indicate that men were both more physically and verbally aggressive than women but that the magnitude of this difference was larger when aggression was physical (Eagly & Steffen, 1986).

In addition to these differences in the use of a distinct type of aggressive behavior, men and women may react to particular types of provocations differently. For example, women may be provoked more by verbal insult than by physical instigation (Frodi et al., 1977; White, 1983). Consistent with this view, Harris (1993) found that men perceived physical attack as more provoking than insensitive or condescending behaviors, whereas for women the relative magnitude of perceived provocation produced by each was reversed. In addition, Van Goozen, Frijda, Kindt, and van de Poll (1994) have shown that women are more likely to report that they would be angry as a result of impolite treatment, abuses, and frustrations than as a consequence of their own inability or incompetence. Such considerations argue for examining the moderating effects of specific types of instigation on gender differences in aggression. In doing so, we distinguished condescending, insensitive, or impolite comments (i.e., insults) from feedback to participants about their level of competence, as provided by a high-status individual (e.g., experimenter or professor), and we compared these types of instigations with each other, as well as with physical attack and frustration.

Another situational factor likely to moderate gender differences in aggression is the specific configurations of gender of actor and gender of target. In some situations, chivalry norms may temper men's aggressiveness toward women, making them less aggressive toward female than male targets (Bjorkqvist & Niemela, 1992; Eagly & Steffen, 1986; Geen, 1990). At the same time, women's fear of more severe retaliation from men may inhibit aggressive responses toward male targets, leading women to display relatively more aggression toward female than male targets. Taken together, these two propositions suggest an interaction between gender of actor and gender of target, such that both men and women behave more aggressively toward targets of the same gender. Although this interaction has been obtained in some studies (Harris, 1973, 1992b; Yousseff, 1968), it is qualified by a gender of target main effect in which male targets receive more aggression from both female and male actors (Harris, 1992b; Taylor & Epstein, 1967). Adding to this complexity, other studies have not found support for this interaction (Hynan, 1982; Larsen, Coleman, Forbes, & Johnson, 1972). The meta-analytic results of Eagly and Steffen showed that male targets do elicit slightly more aggression than female targets. Lack of power, however, precluded any firm conclusions about the effects of both gender of actor on the difference in aggression toward male versus female targets and gender of target on the difference in aggression displayed by male and female actors. Because we believed that our own procedures would increase sensitivity, we meta-analytically examined again the effect of target on gender differences in aggression, separately for neutral and provocation conditions. In addition, separately for male and female actors, we examined the effect of gender of target on the difference in aggression that is found between provocation and neutral conditions.

Finally, gender of experimenter might affect the outcomes of aggression studies. The experimenter's role as an authority may imply evaluation or criticism of participants when they violate norms regarding aggression (Rogers, 1980). Borden (1975) showed that men were more aggressive when an observer was a man than when a woman. To explain this difference, he argued

that whereas male observers approve of aggression, female observers do not. On the basis of these results, we expected men to be more aggressive in the presence of a male experimenter than in the presence of a female experimenter. Women, too, may perceive male observers as more approving of aggression. If so, women may behave more aggressively when the experimenter is a man than when a woman. Although some studies have included both female and male experimenters (e.g., DaGloria & DeRidder, 1981; Lando et al., 1977), we know of no single study that has examined this interaction. Meta-analysis provides a means of examining it correlationally.

Overview

In summary, we meta-analytically reevaluated the moderating effect of provocation on gender differences in aggression. We defined this as provoking those conditions in which participants were intentionally frustrated or attacked.[1] As Geen (1990) noted, provocation can include "frustration, . . . physical attacks, verbal insults, and blows to the victim's self esteem" (p. 31). For each study, an effect size estimate that compared the levels of aggression displayed by men with that of women was calculated separately for a neutral condition, a provocation condition, or both. Our term for this comparison is *gender difference* effect size. In addition, for those studies that had both a neutral and a provocation condition, an effect size that compared aggressive behavior under each condition was calculated separately for women and men. Our term for this second type of comparison is *provocation difference* effect size. Finally, we also examined the moderating effect of provocation by using a continuous measure based on judges' ratings of the experimental inductions. These three sets of analyses should yield convergent, mutually confirming outcomes.

A second set of analyses examined key theoretical variables that might mediate gender differences in aggression. As discussed in our preceding overview, these include male versus female differences in perceptions of negative affect, provocation intensity, and danger from retaliation, as indexed by the differences between male and female judges' ratings of these variables. (A fourth potential mediator, gender differences in perceptions of role appropriateness of aggression, was assessed but not directly discussed in our introduction because our judges' ratings of it were too unreliable for it to be useful in analyses.) Judges' ratings of these variables were made separately for the provocation and neutral conditions of each study. If these variables function as mediators of the effect of experimentally manipulated provocation, they should predict gender differences in aggression.

Finally, for reasons previously discussed, we examined the following categorical contextual variables: type of aggressive response, type of instigation, gender of participant, gender of experimenter, and the gender configurations produced by variation in gender of participant and both that of target of aggression and of experimenter. In addition, we examined the potential effects of a number of other contextual moderators.

Method

Sample of Studies

To collect relevant studies, each volume of the following journals, published before the year 1995, was carefully examined: *Journal of Personality and Social Psychology, Journal of Experimental Social Psychology, Personality and Social Psychology Bulletin, Social Psychology Quarterly, Journal of Applied Social Psychology, Journal of Social Psychology, British Journal of Social Psychology, European Journal of Social Psychology*, and *Aggressive Behavior*.[2] The reference sections of several relevant meta-analyses (Carlson, 1988; Carlson et al., 1989; Carlson & Miller, 1988; Eagly & Steffen, 1986) and qualitative reviews (Frodi et al., 1977; White, 1983) were examined for further citations. Finally, a *PsycLIT* search of psychological abstracts and an *ERIC* search was conducted for the years of 1974 to 1994 and 1966 to 1994, respectively, using the key words *aggression, hostile behavior, gender, mood*, and *human sex differences*. All potentially qualifying articles were copied, and their reference sections checked for relevant additional citations. This process was repeated until no new citations were found.

Inclusions Criteria for Studies

Studies were included in our analyses if it was possible to calculate an effect size estimate of the difference between the aggression of adult female and male participants. Most of the studies had both a neutral and provocation condition, but some studies had only a neutral condition and others had only a provocation condition. Because one goal was to compare neutral conditions with those that were provoking, studies that reported results collapsed across a neutral and provoking condition were excluded from the analyses.[3] In addition, we excluded specific conditions that imposed procedures intentionally designed to diminish the effect of prior provocation, as well as studies that used self-report measures, projective test measures, children under 14 years old, or administered alcohol or other drugs. A few studies included cues of violence (e.g., guns or knives), sexual material (e.g., films or pictures), or humorous stimuli (e.g., cartoons or jokes)—all of which have been shown to have important effects in experimental studies of aggression (Carlson, Marcus-Newhall, & Miller, 1990; Donnerstein & Hallam, 1978; Mueller & Donnerstein, 1977). Because of their potential noncomparability with the remainder of the sample of studies, specific conditions within them that included these manipulations were excluded from the analysis. When studies had experimental conditions absent of these cues and reported statistics that allowed the effects of the no-cue conditions to be separated from the cue conditions, the neutral and provocation conditions were included in the meta-analyses.

---

[1] Experiments that used a teacher–learner paradigm but did not specifically provoke the participant with prior frustration, competition, insult, or physical attack were categorized as neutral.

[2] Predecessors of these journals were also scanned, including the *Journal of Abnormal and Social Psychology*.

[3] The following study reported only an F value that did not separate male from female participants: Knott and Drost (1970). The following study reported a male versus female effect that collapsed across the provocation versus neutral–control conditions: Buss (1966a). The following studies reported statistics that collapsed across the aggressive cue and no-cue conditions: Harris, Liguori, and Joniak (1973); and Harris and Samerotte (1975). The following studies reported statistics that collapsed across both the aggressive cue versus no-cue conditions, as well as the provocation versus neutral conditions: Chaikin, Derlega, Yoder, and Phillips (1974); Harris (1974); Siegman and Dintzer (1977). The following study was not included because participants were offered a $10 reward for the best trainer, which made aggression instrumental: Bond and Dutton (1975). Each of the following studies were not included because there was no behavioral or verbal measure of aggression (e.g., a self-report or Thematic Apperception Test measure of aggression was used): Deturck (1987); Doyle and Biaggio (1981); Gustafson, Hemlin, and Soderberg (1987); Infante (1989); Kanekar, Bulsara, Duarte, and Kolsawalla (1981); Reinisch and Sanders (1986).

In total, 64 studies yielded 107 gender difference effect sizes.[4] Eight studies had only a neutral condition, 17 studies had only a provocation condition, and 28 studies had both a neutral and provocation condition. Several studies yielded more than one effect size per neutral condition, provocation condition, or both; these studies varied conditions of interest, such as female versus male target. Thus, 5 studies yielded two neutral conditions, 5 yielded two provocation conditions, and 2 yielded two of each condition. For the provocation difference effect sizes, 23 studies yielded both a female and male effect size, 2 studies yielded two of each, and 12 studies yielded only a female (6 studies) or a male (6 studies) effect size.[5]

Variables Coded From Each Research Report

The following information was coded for each study: (a) type of design (laboratory or field experiment); (b) type of instigation (physical [shock or noise], verbal [insulted by another, evaluated poorly by a confederate, or yelled at], negative intelligence feedback [high-status experimenter provides negative feedback about the participant's intelligence][6], and frustration [difficult puzzle, participant loses a competitive game, confederate cuts into a line in front of the participant, or confederate remains stopped at an intersection in front of participant's car]); (c) type of aggression (physical [administering electric shock or noise blasts to another], verbal [negative evaluation or disparaging comments directed at a target], horn honking [blaring one's car horn]); (d) gender of experimenter (woman, man, or both); (e) gender of target (woman, man, or both); (f) type of paradigm (competition, evaluation, frustration, and teacher–learner); (g) degree of surveillance (none, experimenter only, experimenter and confederate, or many others); (h) opportunity to aggress (participant forced to aggress or participant free to aggress); and (i) confederate feedback (no feedback or feedback). To maximize the number of studies included in the overall analysis, when there was insufficient information for coding a study on a given variable, it was assigned a missing value. For each study, B. Ann Bettencourt and a research assistant coded each of the variables. Their percentage agreement on the nine variables ranged from 80% (gender of experimenter) to 100% (type of design and type of aggression). Coders discussed and resolved discrepancies.

Variables Rated From Each Research Report

In addition to the categorical variables coded directly from the research reports, two female and two male judges rated each of four variables, including perceptions of provocation intensity, negative affect, danger from retaliation, and gender-role appropriateness. By comparing the mean ratings of male versus female judges, we could index gender differences in perceptions of the levels of each of these four variables and examine whether they plausibly function as mediators of gender differences in aggression. For the assessment of provocation intensity, judges were asked to rate the extent to which the participant would perceive that they were attacked or provoked by another person. To assess negative affect, judges were asked to rate the extent to which the events would make the participant feel unpleasant or negative emotions. For the judgment of danger from retaliation, we used a question similar to that used by Eagly and Steffen (1986): How much danger would you face if you enacted the available aggressive act? Gender role appropriateness was rated as how likely is it that the average (man/woman) would enact this behavior (Is it appropriate to act aggressively in the situation, using the behavior described).

When rating each variable separately for each study, the judges referred to a form that provided a definition of the construct. They rated all variables on 9-point scales which contained descriptors that labeled the odd-numbered points. For example, the rating scale for provocation was as follows: 1 = *no provocation*, 3 = *slight provocation*, 5 = *moderate provocation*, 7 = *notable provocation*, 9 = *extreme provocation*. For some studies, only a neutral condition or provocation condition was available for judgment; but for other studies, both neutral and provocation conditions were available. Although we expected the neutral conditions to yield ratings of minimal levels of provocation intensity and negative affect, we asked judges to rate these conditions because prior ratings of some of the neutral conditions in the aggression literature suggest that they can include features perceived as aversive (Carlson et al., 1990; Carlson & Miller, 1988).

Each judge received photocopies of the Method sections of all studies.[7] To maximize the validity of their judgments, judges rated all studies on a given variable before proceeding to the next judged variable, rated the studies in different randomly assigned orders for each variable, and rated each of the four variables in a distinct, randomly determined order. Exact replications were rated only once. These rating procedures and materials were adapted from those we used in previous meta-analyses (cf. Carlson et al., 1990; Carlson & Miller, 1988) and have been shown to produce valid ratings (Miller & Carlson, 1990; Miller et al., 1991).

The intraclass correlations (Shrout & Fleiss, 1979; also see Lahey, Downey, & Saal, 1983) for the female, male, and all judges are presented in Table 1, along with the effective reliabilities (*R*s; Rosenthal, 1984) and the Pearson correlations (*r*s). For provocation intensity and negative affect, the reliabilities between opposite, as well as same gender, raters were fairly high. For danger of retaliation, reliabilities were high for

---

[4] Eagly and Steffen's (1986) meta-analysis included 58 reports, which yielded 77 effect sizes.

[5] Twelve of the 66 provocation difference comparisons were taken from studies that could not be included in the calculation of the gender difference effect sizes. Six of these provocation difference effect sizes were derived from studies that included only female participants. To obtain effect sizes for male participants that were comparable with those of the female-only studies, the method sections of studies that included only male participants were carefully selected to match each study that only used female participants. For five of the study pairs, the male-only study had been cited in the female-only study, which lends to the validity of the comparability between the methods and measures used. For one female–male study pair, the two separate studies were reported in one single publication, thus we felt confident that the procedures of those studies were likely to be relatively consistent with each other. An analysis of the provocation difference effect sizes that excluded these 12 effect sizes revealed very similar results to one which did include them. The mean weighted effect size for the female participants, $d+ = 0.73$, was only directionally higher than that for male participants, $d+ = 0.66$, but this difference was not significant, $Q_b(1) = 0.54$, $p < .50$.

[6] A number of studies provoked participants by having an authority figure threaten their intellectual abilities. In all of these conditions, the participants were criticized by the experimenter for their limited intelligence or lack of academic ability. Other researchers, seeking to study the effects of self-esteem, have used either similar or identical inductions to manipulate participants' level of self-esteem (e.g., Craparo, Hines, & Kayson, 1981). In light of Van Goozen et al.'s (1994) results showing that women are not provoked by this type of incompetence feedback and because this type of instigation seemed qualitatively as well as quantitatively different from the negative evaluation or insult instigations (which typically were induced by a confederate who posed as a participant), for a study in which an experimenter criticized the intelligence or academic ability of participants, the instigation was coded as negative intelligence feedback.

[7] Relevant portions of studies (including explanations of the procedure, materials, and instruments) necessary to judge the variables were highlighted in yellow. Any sections that were not critical for understanding the methods or that referred to the hypotheses or results were cut out of the portions that the judges read.

same gender judges but, as expected, discernibly lower but positive for opposite gender judges. Finally, the reliabilities for gender role appropriateness were markedly low and were not used in any further analyses.

Effect Size Calculation

We used *d* as an effect size index, which is the difference between the means of two groups divided by a pooled standard deviation (*SD*)and corrected for small sample bias (Hedges & Olkin, 1985). Whenever possible, the pooled *SD* was based on the *SD*s of each group, but when not reported, an estimate of the pooled *SD* was obtained from analysis of variance results. Using DSTAT software (Johnson, 1989), we estimated each *d* from the condition means, an *F* for main effects, an *F* from an interaction along with the corresponding means and sample sizes, a χ^2 statistic, or a correlation. For studies that did not supply statistics but reported that there were no significant differences between conditions, we adopted the conservative procedure of estimating the effect size as 0. Each effect size was calculated separately by B. Ann Bettencourt and an assistant. When discrepancies arose, they were resolved by discussion.

To retain the distinction between the neutral and provocation conditions of a single study, a gender difference effect size was separately calculated for each condition. A positive gender difference effect size indicates greater aggression by male participants. In addition, separately for female and male participants, the second type of effect size, the provocation difference effect size, was calculated by comparing aggression under provocation with that under neutral conditions. A positive provocation difference effect size indicates greater aggression under provocation.

Thus, any single experiment could potentially yield four effect sizes: (a) male versus female participants, neutral; (b) male versus female participants, provocation; (c) provocation versus neutral, female participants; and (d) provocation versus neutral, male participants. In many cases, however, only one or two of these effect sizes could be calculated because not all studies included both neutral and provocation conditions. Although the gender difference and provocation difference effect sizes are at best only partially independent, the two alternative analytical approaches to examining the effect of provocation on gender differences in aggression should yield mutually confirming convergent outcomes.

Results

In the following sections, the effect size estimates that compared the aggressive behavior of men with women, as previously noted, are referred to as gender difference effect sizes. They are reported separately for the neutral and provocation condition of each study in Table 2. In addition, Table 2 contains the coded values for type of instigation, type of aggression, gender of target, and gender of experimenter. Separately for female and male participants, Table 3 presents the effect size estimates that compare aggression under the provocation conditions with that under the neutral conditions of each study. As previously indicated, we refer to these as provocation difference effect sizes. For convenience, the coded values of the categorical variables presented in Table 2 are repeated in Table 3. Whenever the results of the analyses of the provocation difference effect sizes (reported in Table 3) help clarify the results of the analyses of the gender difference effect sizes (reported in Table 2), we discuss them as well.

We used two methods to analyze the effect sizes. In our primary analyses, each effect size estimate was weighted by the reciprocal of its variance. However, we also analyzed the effect sizes using parametric statistics (i.e., *F*), which do not weight the effect sizes. In the following sections, we focus primarily on the results of the weighted analyses, but we also present the results of the unweighted analyses and discuss the degree of correspondence between the outcomes of two types of analysis. When differences in sample sizes are associated with paradigm differences, the results obtained with weighted analyses can be misleading. Indeed, as we subsequently show, such concerns are important for our own dataset.

(*text continues on page 433*)

Table 1
Correlations and Reliabilities for Female and Male Judges

Source of judgment	Provocation intensity	Negative affect	Danger perception	Gender role appropriateness
Female judges				
ICC	.90	.85	.88	.47
R	.89	.86	.92	.18
r	.81	.78	.80	.11
Male judges				
ICC	.89	.89	.84	.47
R	.89	.89	.86	.40
r	.80	.83	.78	.26
Female and male judges				
ICC	.72	.70	.36	.26
R	.86	.85	.26	.26
r	.75	.74	.16	.16

Note. The intraclass correlations (ICCs) between judges of the same gender were calculated ([BMS − EMS]/BMS), which assumes no interaction between judges and observations, and the ICC between all four judges was calculated ([BMS − EMS]/BMS + [k − 1] × EMS), which allows interaction between judges (i.e., female and male) and observations (Shrout & Fleiss, 1979). Effective reliabilities (*R*s) are Spearman–Brown coefficients. Effective reliabilities and raw correlations between female and male judges are based on the average rating within each gender. BMS = mean square between observations; EMS = residual mean square; *k* = number of judges; *r* = Pearson correlation.

Table 2
Gender Difference Effect Sizes and Categorical Codes

Study	GD *d*	95% CI	*N*	Instigation	Aggression	Experimenter gender	Target gender
			Neutral conditions				
Schuck et al. (1971)	2.31	0.71–3.91	10	0	1	7	3
Yinon et al. (1975)[a]	2.06	1.23–2.88	35	0	1	3	3
Richardson et al. (1994)	2.03	1.26–2.80	39	0	2	7	7
Yinon et al. (1975)[b]	1.92	1.09–2.88	34	0	1	3	3
DaGloria & DeRidder (1979)	1.87	0.38–3.36	10	0	1	7	3
Edwards (1968)[c]	1.27	0.59–1.94	40	0	1	1	3
Scharff & Schlottmann (1973)	1.13	0.18–2.07	20	0	1	2	2
Raden (1974)	1.01	0.52–1.50	69	0	1	2	2
Titley & Viney (1969)	0.80	−0.49–2.08	10	0	1	7	1
Epstein (1965)	0.78	0.13–1.42	40	0	1	7	2
Atkinson & Polivy (1976)	0.71	−0.20–1.61	20	0	2	1	1
Rogers (1980)	0.66	0.25–1.07	96	0	1	3	3
Lando et al. (1977)	0.63	−0.01–1.27	40	0	1	5	5
Fischer et al. (1969)	0.61	−0.81–2.03	8	0	2	7	3
Mueller & Donnerstein (1977)	0.60	−0.56–1.76	12	0	1	3	3
Caprara et al. (1984)	0.52	−0.11–1.15	40	0	1	7	3
Carver (1975)	0.52	−0.12–1.14	40	0	1	2	2
Hammock & Richardson (1992)	0.49	0.20–0.78	194	0	1	7	7
Rothaus & Worchel (1960)	0.48	0.08–0.89	96	0	2	2	2
Yousseff (1968)	0.42	0.05–0.78	120	0	1	7	5
Anderson (1993)	0.40	−0.37–1.16	27	0	1	1	3
Zillman et al. (1981)	0.39	0.07–0.87	72	0	2	3	3
Edwards (1968)[d]	0.38	−0.25–1.01	40	0	1	1	3
Gaebelein (1977)	0.37	−0.51–1.25	20	0	1	2	1
Leventhal & Shemberg (1969)	0.36	−0.09–0.80	80	0	1	7	7
Gentry (1972)	0.31	−0.74–1.37	14	0	2	3	3
Goodwin (1973)	0.23	−0.58–1.03	24	0	1	2	2
Leventhal et al. (1968)	0.18	−0.44–0.80	40	0	1	7	2
Larsen et al. (1972)	0.16	−0.64–0.96	27	0	1	2	1
Pytkowicz et al. (1967)	0.15	−1.10–1.39	10	0	2	1	1
Frodi (1978)	0.15	−0.98–1.29	12	0	1	4	3
Worchel (1966)	0.08	−0.42–0.59	60	0	2	2	2
Taylor & Epstein (1967)	0.08	−1.06–1.21	12	0	1	2	2
Buss (1963)	0.07	−0.55–0.69	40	0	1	7	5
Buss (1966b)	0.07	−0.81–0.94	20	0	1	7	1
Buss (1966b)	0.04	−0.84–0.92	20	0	1	7	2
Taylor & Epstein (1967)	0.02	−1.12–1.15	12	0	1	2	1
Baron & Ball (1974)	0.00	−1.24–1.24	10	0	1	2	2
Berkowitz (1960)	0.00	−0.44–0.44	80	0	2	7	3
Berkowitz & Holmes (1959)	0.00	−0.68–0.68	34	0	2	7	3
Golin & Romanawski (1977)	0.00	−0.62–0.62	40	0	2	2	5
Stagner & Gongdon (1975)	0.00	−0.88–0.88	20	0	2	7	7

Table 2 (*continued*)

Study	GD *d*	95% CI	*N*	Instigation	Aggression	Experimenter gender	Target gender
Koerner (1977)	−0.02	−0.52–0.49	60	0	5	3	3
Lando et al. (1977)	−0.10	−0.72–0.52	40	0	5	5	5
Fisher & Harris (1976)	−0.15	−0.64–0.93	25	0	2	1	1
Caprara et al. (1983)	−0.16	−0.52–0.20	240	0	1	7	3
Larsen et al. (1972)	−0.16	−0.61–0.30	78	0	1	2	2
Shemberg et al. (1968)	−0.31	−0.90–0.27	45	0	1	7	2
Caprara (1982)	−0.58	−0.98–−0.18	100	0	1	7	3
Titley & Viney (1969)	−1.21	−2.56–0.14	10	0	1	7	2
			Provocation conditions				
Lando & Donnerstein (1978)	1.77	0.31–3.23	10	1	1	7	3
Epstein (1965)	0.98	0.33–1.64	40	3	1	7	2
Unger et al. (1974)	0.95	0.66–1.24	204	4	4	6	1
Pytkowicz et al. (1967)	0.87	−0.43–2.17	10	3	2	1	1
Caprara (1982)	0.78	0.37–1.18	100	3	1	7	3
Gaebelein (1977)	0.64	−0.26–1.54	20	5	1	2	1
Worchel (1966)	0.53	0.16–0.89	120	3	2	2	2
Lando et al. (1977)	0.52	−0.11–1.15	40	1	1	5	5
Doob & Gross (1968)	0.50	0.03–0.97	74	4	4	6	2
Caprara et al. (1984)	0.46	−0.17–1.09	40	3	1	7	3
Buvinic (1975)	0.45	0.04–0.85	48	2	2	3	3
Caprara et al. (1983)	0.42	0.07–1.79	240	3	1	7	3
Harris (1976)	0.39	−0.24–1.02	40	4	3	6	2
Rothaus & Worchel (1960)	0.38	−0.02–0.79	96	3	2	2	2
Hynan (1982)	0.38	0.04–0.71	140	4	1	7	5
Hedrick (1976)	0.34	−0.41–1.08	28	1	1	1	2
Fisher & Harris (1976)	0.30	−1.09–0.49	25	2	2	1	1
Hoppe (1977)	0.30	−0.27–0.86	48	1	1	7	2
Hammock & Richardson (1992)	0.28	−0.01–0.57	194	1	1	7	7
Hynan & Esselman (1981)	0.28	−0.52–1.09	25	4	1	7	2
Unger et al. (1974)	0.28	0.00–0.55	204	4	4	6	2
Deaux (1971)	0.24	−0.33–0.81	61	4	4	6	1
Deaux (1971)	0.22	−0.34–0.78	62	4	4	6	2
Frodi (1978)	0.21	−0.92–1.34	12	5	1	4	3
Buss (1963)	0.19	−0.43–0.81	40	5	1	7	5
Hynan et al. (1980)	0.19	−0.14–0.88	100	4	1	7	1
Chase & Mills (1973)	0.10	−0.52–0.72	40	4	4	6	2
Lightdale & Prentice (1994)	0.07	−0.38–0.53	74	5	5	1	1
Hedrick (1976)	0.06	−0.68–0.80	28	1	1	1	1
Harris & Huang (1974)	0.03	−0.47–0.51	64	2	3	4	4
Ahmed (1982)	0.00	−0.22–0.22	320	4	3	6	2
Baron & Ball (1974)	0.00	−1.24–1.24	10	1	1	2	2
Berkowitz (1960)	0.00	−0.44–0.44	80	2	2	7	3

(*table continues*)

Table 2 (*continued*)

Study	GD *d*	95% CI	*N*	Instigation	Aggression	Experimenter gender	Target gender
			Provocation conditions (cont'd)				
Berkowitz & Holmes (1959)	0.00	−0.68–0.68	34	2	2	7	3
Dor-Shav & Dolgin (1981)	0.00	−0.98–0.98	16	2	1	7	5
Golin & Romanowski (1977)	0.00	−0.62–0.62	40	2	2	2	5
Harris & Klingbeil (1976)	0.00	−0.41–0.41	96	2	3	1	1
Juhnke et al. (1977)	0.00	−0.59–0.59	44	4	2	2	2
Stagner & Gongdon (1975)	0.00	−0.62–0.62	40	4	2	7	7
Richardson et al. (1994)	−0.02	−0.65–0.60	39	2	2	7	7
Taylor & Epstein (1967)	−0.02	−1.15–1.11	12	1	1	2	2
Taylor & Epstein (1967)	−0.03	−1.16–1.10	12	1	1	2	1
Scharff & Schlottmann (1973)	−0.09	−0.97–0.78	20	2	1	2	2
Mueller & Donnerstein (1977)	−0.10	−1.23–1.03	12	1	1	3	3
Schuck et al. (1971)	−0.11	−1.35–1.13	10	2	1	7	3
Rohsenow & Bachorowski (1984)	−0.12	−0.69–0.46	47	2	2	5	3
Hoppe (1979)	−0.25	−0.81–0.32	48	1	1	7	1
Koerner (1977)	−0.29	−0.80–0.22	60	2	5	3	3
Zillman et al. (1981)	−0.44	0.91–0.23	72	2	2	3	3
Harris (1973)	−0.45	−1.34–0.44	20	4	4	6	2
Turner et al. (1975)	−0.56	−1.18–0.05	49	4	4	6	2
Anderson (1993)	−0.67	−1.54–0.21	28	1	1	1	3
Lando et al. (1977)	−0.69	−1.33–−0.06	40	5	5	5	5
Atkinson & Polivy (1976)	−0.84	−1.76–0.07	20	2	2	1	1
Fischer et al. (1969)	−1.19	−2.69–0.32	8	2	2	7	3
Dagloria & DeRidder (1979)	−1.20	−2.55–0.15	10	1	1	7	3
Gentry (1972)	−2.52	−3.92–−1.12	14	2	2	3	3

Note. GD *d* = gender difference effect size (effect sizes that are positive indicate more aggressive from male than female participants); CI = confidence interval; *N* = total number of female and male participants; Instigation: 0 = none, 1 = physical, 2 = peer's insult, 3 = negative intelligence feedback, 4 = frustration, 5 = other; Aggression: 1 = physical, 2 = evaluation, 3 = spoken, 4 = horn honk, 5 = other; Experimenter gender: 1 = female, 2 = male, 3 = same, 4 = crossed, 5 = both female and male, 6 = no experiementer, 7 = not available; Target gender: 1 = female, 2 = male, 3 = same, 4 = crossed, 5 = both female and male, 6 = no target, 7 = not available.

[a] Participants made a decision about level of aggression with a group of people. [b] Participants were alone when administering aggression. [c] Participants were told that a confederate would not have an opportunity to retaliate. [d] Participants were told that a confederate would have an opportunity to retaliate.

Table 3
Studies, Provocation Difference Effect Sizes, and Categorical Codes

Study	GD *d*	95% CI	*N*	Instigation	Aggression	Experimenter gender	Target gender
			Female participants				
DaGloria & DeRidder (1979)	4.67	2.28–7.07	10	1	1	7	3
Wilson & Rogers (1975)	3.24	2.25–4.24	16	2	1	1	3
Zillman et al. (1981)	3.00	2.23–3.68	72	2	2	3	3
Atkinson & Polivy (1976)	2.61	1.42–3.80	20	2	2	1	1
Mueller & Donnerstein (1977)	2.58	1.05–4.11	12	1	1	3	3
Anderson (1993)	2.42	1.54–3.31	35	1	1	1	3
Fischer et al. (1969)	2.37	0.56–4.18	8	2	2	7	3
Richardson et al. (1994)	2.14	1.34–2.93	38	2	2	7	7
Donnerwerth & Foa (1974)	2.03	0.95–3.11	20	2	5	2	3
DaGloria & DeRidder (1981)	1.98	0.70–3.26	14	1	1	1	3
Konecni & Ebbesen (1976)	1.81	0.65–2.98	16	1	1	3	3
Gentry (1972)	1.81	0.56–3.05	14	2	2	3	3
Schuck et al. (1971)	1.47	0.07–2.87	10	2	1	7	3
Konecni (1975)	1.39	0.91–1.56	42	2	1	7	1
Caprara et al. (1984)	1.29	0.85–1.82	40	3	1	7	3
Lando et al. (1977)	1.27	0.59–1.95	40	1	1	5	4
Lando et al. (1977)	1.16	0.49–1.83	40	5	5	5	5
Landy & Mettee (1969)	1.04	0.19–1.89	24	3	2	2	2
Frodi (1978)	0.84	−0.34–2.01	12	5	1	4	3
Worchel (1966)	0.77	0.32–1.22	90	3	2	2	2
Gaebelein (1977)	0.71	−0.19–1.62	20	5	1	2	1
Rothaus & Worchel (1960)	0.67	0.26–1.08	96	3	2	2	2
Baron (1979)	0.64	−0.58–1.85	16	2	1	2	3
Caprara et al. (1983)	0.44	−0.01–0.88	240	3	1	7	3
Pytkowicz et al. (1967)	0.24	−1.00–1.49	10	3	2	1	1
Buss (1963)	0.16	−0.46–0.78	40	5	1	7	5
Taylor & Epstein (1967)[a]	0.08	−1.05–1.22	12	1	1	2	2
Taylor & Epstein (1967)[b]	0.05	−1.09–1.18	12	1	1	2	1
Scharff & Schlottmann (1973)	0.02	−0.86–0.89	20	2	1	2	2
Caprara (1982)	0.00	−0.39–0.39	100	3	1	7	3
Koerner (1977)	−0.04	−0.55–0.47	60	2	5	3	3
Fisher & Harris (1976)	−0.68	−0.13–1.49	25	3	2	1	1
Epstein (1965)	−1.71	−2.43–−0.98	40	3	1	7	2

(*table continues*)

Table 3 (*continued*)

Study	GD *d*	95% CI	*N*	Instigation	Aggression	Experimenter gender	Target gender
			Male participants				
Zillman et al. (1981)	2.17	1.60–2.76	72	2	2	3	3
Berkowitz & Knurek (1969)	2.14	1.25–3.05	30	3	2	3	2
Mueller & Donnerstein (1977)	1.88	0.52–3.24	12	1	1	3	3
Geen et al. (1975)	1.80	0.59–2.90	30	1	1	3	3
DaGloria & DeRidder (1979)	1.60	0.18–3.03	10	1	1	7	3
Caprara (1982)	1.36	0.93–1.81	100	3	1	7	3
Caprara et al. (1984)	1.23	1.23–2.26	40	3	1	7	3
Baron & Bell (1977)	1.21	0.18–2.25	17	2	1	3	3
Worchel (1966)	1.21	0.74–1.68	90	3	2	2	2
Lando et al. (1977)	1.17	0.50–1.84	40	1	1	5	5
Caprara et al. (1983)	1.02	0.55–1.49	240	3	1	7	3
Gaebelein (1977)	0.93	0.01–1.85	20	5	1	2	1
Frodi (1978)	0.92	−0.27–2.11	12	5	1	4	3
Anderson (1993)	0.88	−0.08–1.84	20	1	1	1	3
Pytkowicz et al. (1967)	0.88	−0.41–2.18	10	3	2	1	1
Geen (1968)	0.79	−0.13–1.68	30	2	1	2	3
Konecni (1975)	0.75	0.23–1.27	30	2	1	7	1
Atkinson & Polivy (1976)	0.74	−0.16–1.65	20	2	2	1	1
DaGloria & DeRidder (1981)	0.66	−0.24–1.56	20	1	1	1	3
Lando et al. (1977)	0.58	−0.05–1.21	40	5	5	5	5
Rothaus & Worchel (1960)	0.57	0.16–0.98	96	3	2	2	2
Baron (1979)	0.41	−0.29–1.11	16	2	1	1	3
Buss (1963)	0.28	−0.34–0.90	40	5	1	7	5
Richardson et al. (1994)	0.08	−0.54–0.70	40	2	2	7	7
Taylor & Epstein (1967)[a]	0.03	−1.10–1.16	12	1	1	2	2
Taylor & Epstein (1967)[b]	0.02	−1.11–1.15	12	1	1	2	1
Gentry (1972)	0.00	−1.05–1.05	14	2	2	3	3
Fischer et al. (1969)	−0.23	−1.62–1.16	8	2	2	7	3
Fisher & Harris (1976)	−0.23	−0.55–1.02	25	3	2	1	1
Koerner (1977)	−0.31	−0.82–0.20	60	2	5	3	3
Epstein (1965)	−0.81	−1.45–−0.16	40	3	1	7	2
Schuck et al. (1971)	−0.96	−2.26–0.35	10	2	1	7	3
Scharff & Schlottmann (1973)	−1.19	−2.15–−0.24	20	2	1	2	2

Note. GD *d* = provocation difference effect size (effect sizes that are positive indicate more aggressive under provocation conditions); CI = confidence interval; *N* = total number of female or male participants; Instigation: 1 = physical, 2 = peer's insult–evaluation, 3 = negative intelligence feedback, 4 = frustration, 5 = other; Aggression: 1 = physical, 2 = evaluation, 3 = spoken, 4 = horn honk, 5 = other; Experimenter gender: 1 = female, 2 = male, 3 = same, 4 = crossed, 5 = both female and male, 6 = no experimenter, 7 = not available; Target gender: 1 = female, 2 = male, 3 = same, 4 = crossed, 5 = both female and male, 6 = no target, 7 = not available.

[a] Target gender was male. [b] Target gender was female.

For the weighted analyses, between class statistics (Q_b) and within-class homogeneity statistics (Q_w) were used to analyze the effect size estimates (Hedges & Olkin, 1985). The between-class statistic is analogous to an F statistic and indicates the magnitude of variance between the average effect sizes. The within-class homogeneity statistics in the weighted analyses indicate significant heterogeneity among the effect sizes in a class. A confidence interval, using a fixed effects approach (Hedges & Olkin, 1985), was also computed for each mean weighted effect size in each class. In addition, when there were more than two classes for a between-class comparison, pair-wise contrasts between mean weighted effect sizes (χ^2) were conducted.

Summary of Effect Size Estimates

Table 4 reports the summary effect size estimates and their respective confidence intervals. For gender differences, a positive mean effect size indicates that men were more aggressive than women. For provocation differences (pooled across female and male participants), a positive mean effect size indicates that the participants who were provoked were more aggressive than those who were not. These analyses are not of primary concern, but as seen they show the expected main effects for gender and provocation. The mean effect size comparing male with female participants differs significantly from 0, showing that, overall, men were more aggressive than women.[8,9,10] The mean effect size for the effect of provocation, presented in the lower panel of Table 4, confirms that people were more aggressive when they were provoked.

Weighting the effect size of each study by its sample size is recommended (e.g., Hedges & Olkin, 1985). The weighted mean assigns greater weight to studies with larger sample sizes on the assumption that their effects are more reliable. As we have previously indicated, although arguably appropriate on statistical grounds, this procedure ignores the fact that sample size is sometimes confounded with other study characteristics that affect the outcome of analyses. When it is, the researcher must balance the statistical appropriateness of weighting against the conceptual distortion (i.e., misleading outcomes) that it may impose. To assess this potential problem in the present set of studies, differences in sample size for studies using each of four distinct paradigms were examined with a nonparametric Kruskal–Wallis test (Siegel, 1956). As shown in the upper panel of Table 5, the mean sample size used in field studies was much larger than that used in evaluation, teacher–learner, and competition paradigms, $\chi^2(3) = 8.68, p < .05$.

Because provocation is the primary variable of interest, a nonparametric test was also conducted for the variable, type of instigation. As shown in the lower panel of Table 5, the sample sizes of studies vary considerably across the four operationali-

Table 4
Summary of Effect Sizes

Difference of effect size	Effect size	95% CI	Median effect size
Gender (k = 107)			
M weighted with modified *N*s[a]	.24	0.18–0.29	.19
M unweighted	.23	0.11–0.36	
M weighted with unmodified *N*s	.24	0.19–0.30	
Provocation (k = 66)			
M weighted with modified *N*s[a]	.76	0.66–0.85	.86
M unweighted	.91	0.65–1.17	
M weighted with unmodified *N*s	.75	0.66–0.83	

Note. Gender difference effect sizes that are positive indicate that men are more aggressive. Provocation effect sizes that are positive indicate higher aggression under provoking conditions. k = number of effect sizes in the category; CI = confidence interval.
[a] For this mean value, the weights of studies that had *N*s greater than 90 were modified by specifying that $N = 90$.

[8] In many cases (but not all), a gender difference effect size for a neutral condition and a gender difference effect size for a provocation condition were derived from a single study. As previously noted, using more than one effect size per study may yield statistically dependent data, not because they are based on the responses from the same participants (within subjects) but because they are based on the responses from participants in the same experiment. Tukey's Jackknife is one method used to assess whether statistical dependencies exist within a dataset (Glass, et al., 1981). To do so, the meta-analyst first calculates the average effect size for the entire set of observations. That set of observations in this sample included 107 effect sizes (from 64 studies). Next, a set of average effect sizes (partial means) are calculated by removing, one at a time, each study's effect size or effect sizes. Thus, for this study, there were 64 partial means. Third, pseudovalues are calculated using the following formula: (No. of Studies × Mean Effect Size for All Observations) − ([No. of Studies − 1] × [Each Partial Mean]). The number of pseudovalues calculated is equivalent to the number of studies in the dataset. There were 64 pseudovalues in our analysis. The pseudovalues are then analyzed to obtain a Tukey's Jackknife mean effect size and confidence interval (CI). Finally, the Jackknife mean and CI is compared with that from the entire set of observations. If the two mean values and CIs are similar, dependence in the sample is not a problem. The overall mean gender difference effect size, d = 0.23, and CI, 95% CI = 0.11–0.36, derived from the Tukey's Jackknife procedure closely paralleled the results based on the full set of 107 separate observations, d = 0.21, 95% CI = 0.09–0.33. Similarly, the mean provocation difference effect size from the Tukey's Jackknife analysis, d = 0.83, 95% CI = 0.45–1.21, was very similar to that for the 66 separate female and male provocation difference effect sizes, $d+$ = 0.91, 95% CI = 0.66–1.21. These Jackknife analyses strongly suggest that statistical dependence is not a problem in the present dataset.

[9] Fourteen effect size estimates of 0 were derived from nine studies. For five of these studies, an effect size of 0 was estimated for both the provocation and the neutral conditions. As would be expected by the prediction that provocation reduces gender differences in aggression, there were several more effect sizes estimated as 0 from studies that only had a provocation condition than from those that only had a neutral condition. To determine if the four gender difference estimates derived from the provocation conditions biased the comparison between the provocation and neutral conditions, the four estimates of 0 were excluded from the analysis of gender difference effect sizes. The weighted mean effect size for the provocation conditions without these four cases, $d+$ = 0.20, was notably similar to that which included them, $d+$ = 0.17, and the difference between the neutral and provocation conditions remained significant, $Q_b(1) = 5.71, p < .01$.

[10] The unweighted mean and weighted mean analyses included 14 effect sizes estimated as 0. For these studies, the researchers reported no differences between female and male participants but did not provide statistics. Analyses with these effect sizes removed from the dataset produced similar outcomes, $d+$ = 0.26, CI = 0.21–0.34.

zations of provocation, with frustration manipulations having larger samples, $\chi^2(3) = 15.59, p < .01$.

Clearly, for this sample, weighted analyses would assign more impact to field studies, which typically implement less control than laboratory studies, and would overrepresent the effect of frustration. To deal with this problem, we adopted a compromise procedure. An examination of the distribution of the study sample sizes revealed that those equivalent to 90 or more were outliers. To minimize potential problems that might result from weighting mean effect sizes by their sample sizes, the mean effect sizes were weighted according to their sample size when the sample size did not exceed 90. However, when the sample size exceeded 90, we weighted the mean effect size by a reduced "total sample size" of 90. For the gender difference effect sizes, this procedure was invoked for eight studies; for the provocation difference effect sizes, it was invoked for two studies.

Effect of Provocation on Gender Differences in Aggression

Gender difference effect sizes. Our major goal was to assess whether provocation affects the magnitude of gender differences in aggression. As expected, and as seen in the upper panel of Table 6, men were more aggressive than women under neutral conditions. The mean effect size was of moderate magnitude, $d+ = 0.33$, and the 95% CI excluded 0. When provoked, although 0 was excluded in its 95% CI, the mean effect size reflecting the difference between men and women was of small absolute magnitude, $d+ = 0.17$. In comparison with the weighted analysis, the unweighted mean effect sizes indicated a similar magnitude of gender differences in aggression under the neutral conditions, although the absolute value of the mean effect size was larger than that obtained in weighted analysis. In contrast with the weighted analysis, however, in the unweighted analysis, gender difference under provocation was not reliably greater than 0. Most important, as predicted, gender differences were more pronounced under neutral than under provocation conditions, $Q_b(1) = 6.65, p < .01$.[11] The within-class homogeneity statistic, however, shows that there is substantial variance within each class of effect sizes, indicating that this classification does not account for the variation between studies within the neutral and provocation conditions, respectively.

Table 5
Mean Sample Size for Effect Size Comparisons as a Function of Experimental Paradigm and Type of Instigation

Type	*M*	*N*
Paradigm[a]		
Evaluation	44	25
Competition	48	24
Teacher–learner	50	39
Field	106	11
Instigation[b]		
Physical attack	38	21
Insult or negative evaluation	39	27
Negative intelligence feedback	80	16
Frustration	90	16

[a] Kruskal–Wallis, $\chi^2 = 8.77, p < .05$. [b] Kruskal–Wallis, $\chi^2 = 15.59, p < .01$.

Table 6
Gender Difference Sizes for Neutral and Provocation Conditions

Condition	*k*	*M* effect size	95% CI	Q_w
Weighted gender difference effect sizes				
Neutral	50	.33	0.23–0.42	136.98***
Provocation	57	.17	0.09–0.25	106.89***
Unweighted gender difference effect sizes				
Neutral	50	.43	0.24–0.63	
Provocation	57	.06	−0.10–0.22	

Note. Effect sizes are positive for differences in the male direction. *k* = number of effect sizes in the category; CI = confidence interval; Q_w = homogeneity within class.
*** $p < .001$.

Provocation difference effect sizes. The overall comparison of the provocation difference effect sizes was only directionally consistent with our confirmation of the predicted effect for the gender difference effect sizes. As shown in the lower panel of Table 6, both men, $d+ = 0.70$, 95% CI = 0.57–0.82, and women, $d+ = 0.82$, 95% CI = 0.69–0.95, were more aggressive when provoked. Because studies show that men are more aggressive than women in neutral conditions, however, we expected that the provocation effect for women would exceed that of men. This directional effect, although seen in the absolute magnitudes of these male and female provocation weighted effect sizes, was only marginally reliable, $Q_b(1) = 1.78, p < .10$. Once again, the unweighted comparison shows a similar pattern of differences between female (mean unweighted effect size = 1.19, 95% CI = 0.75–1.63) and male (mean unweighted effect size = 0.64, 95% CI = 0.35–0.92) participants. Moreover, for the unweighted mean effect sizes, this difference is reliable, $F(1, 105) = 8.78, p < .01$. We isolate the source of our failure to find support for our prediction for the provocation difference weighted effect sizes in the next section.

Type of instigation. Although both men and women were generally more aggressive when provoked, specific types of instigations did differentially affect the degree to which provocation attenuated gender differences in aggression. The mean unweighted and weighted gender difference effect size for each type

[11] When the studies were categorized and pooled across neutral and provocation conditions using Eagly and Steffen's (1986) dichotomization of provocation into minimal versus maximal levels, the mean weighted effect size for minimal provocation was .38 (CI = 0.29–0.42) and for maximal provocation, .21 (CI = 0.15–0.28). The comparison between these mean effect sizes was significant, $Q_b = 8.37, p < .01$, which was not consistant with Eagly and Steffen's findings. This discrepancy between the findings of the present analysis and that of Eagly and Steffen may be due to differences in the inclusionary criteria of the two meta-analyses and to the fact that the present meta-analysis included a greater number of effect sizes, thus having more power to detect differences.

of instigation is presented in Table 7, along with that for the neutral conditions (as a reference point).[12] The pattern of results for the unweighted and weighted analysis are similar to each other. As seen in Table 7, the type of instigation used as a provocation moderates gender differences in aggressive behavior, $Q_b(3) = 31.08, p < .0001$, and the within-class homogeneity statistics suggest that the effect sizes within each class of provocation manipulations are sufficiently homogeneous. Inspection of the weighted mean effect sizes reveals a relatively large gender difference among studies that use negative intelligence feedback as the instigation. Comparison between the weighted mean effect sizes for physical attack and insult are marginally different from each other, $\chi^2(1) = 3.51, p = .06$. Although the effect size for frustration exceeds that for insult, $\chi^2(1) = 11.68, p < .001$, these three types of provocation yield mean gender difference effect sizes that are smaller than that produced by negative intelligence feedback, physical attack, $\chi^2(1) = 9.04, p < .01$; insult, $\chi^2(1) = 29.94, p < .0001$; frustration, $\chi^2(1) = 7.29, p < .01$. Similarly, in a post hoc comparison, the gender difference effect size of negative intelligence feedback reliably exceeded the combined average effect size of the other types of instigations, $\chi^2(1) = 19.24, p < .0001$.

In summary, these results suggest that although provocation generally minimizes gender differences in aggression, when exposed to a bogus feedback manipulation—which suggests that they are unintelligent—female participants remain relatively unprovoked. These results are consistent with those of Van Goozen et al. (1994), who found that women tend not to feel anger as a result of suggestions of intellectual incompetence. Instead, they were more apt to express disappointment or sadness. Supporting this suggestion, an analysis comparing male and female judges' mean provocation ratings for the four types of instigation revealed a significant interaction between gender of judge and type of instigation, $F(3, 48) = 2.81, p < .05$. More specifically, the mean value for the male judges' rating of the negative intelligence feedback manipulations was 5.78 (a high score indicates greater perceived provocation), whereas the female judges' rating of negative intelligence feedback was 3.89. Moreover, this mean rating by the female judges was lower than their average ratings of physical attack, $M = 6.91$, and verbal insult, $M = 6.06$.

Examination of the effects found under specific types of provocation further explains our failure to find statistical support for our expectations regarding the difference in the weighted provocation difference effect sizes for male and female participants. When the studies that used negative intelligence feedback as a provocation were removed from the overall analysis of the provocation difference effect sizes, the predicted outcome of a larger provocation difference effect among women than among men was confirmed, female participant, $d+ = 1.22$, k = 23, 95% CI = 1.05–1.41; male participant, $d+ = 0.61$, k = 23, 95% CI = 0.46–0.79, $Q_b(1) = 22.65, p < .001$. Further examination of the provocation effect sizes corroborates this result. As previously suggested, it shows that, whereas provocation ordinarily increases aggression for both men and women, negative intelligence feedback does not seem to provoke women toward aggression. When female participants were either physically attacked, $d+ = 1.50$, or insulted by a peer, $d+ = 1.42$, the magnitude of the difference between the provocation and neutral conditions was much larger than when they were exposed to a negative intelligence feedback, $d+ = 0.33$—physical attack versus negative intelligence feedback, $\chi^2(1) = 29.21, p < .0001$; peer insult versus negative intelligence feedback, $\chi^2(1) = 49.23, p < .0001$. By contrast, for men the mean effect size for negative intelligence feedback, $d+ = 0.82$, was smaller than that for physical attack, $d+ = 0.96$. Moreover, it exceeded that which was produced by a peer's insult, $d+ = 0.45$—negative intelligence feedback versus peer insult, $\chi^2(1) = 6.15, p < .05$—as did physical attack, $\chi^2(1) = 5.66, p < .05$.[13]

Assessment of provocation intensity as a continuous variable. Although the preceding analyses provide evidence for the moderating effects of provocation on gender differences in aggression, an average of the female and male judges' pooled ratings of provocation intensity allows a third and more sensitive test of the moderating effects of provocation on gender differences in aggression by providing a continuous measure of provocation intensity. A series of least squares regression analyses, calculated with each effect size weighted by the reciprocal of its variance, were conducted to assess the relationships between gender differences in aggression and both the average ratings of provocation intensity and negative affect. Because weighted regression outcomes from standard programs such as SAS and SPSSx are based on a model somewhat different than that for fixed effects meta-analyses, the standard errors (*SE*s), degrees of freedom (*df*), and significance levels must be corrected (Hedges, 1994; Johnson, 1989). The corrected *SE* is calculated by dividing the *SE* derived from the computer output by the residual mean square (Hedges, 1994). This correction yields a *Z* statistic, which we computed by using DSTAT software (Johnson, 1989).

Consistent with the outcomes that our previous approaches to the data analysis have yielded, a univariate weighted regression analysis across the 105 observations showed that higher levels of provocation intensity, as assessed by female and male judges' pooled ratings, were related to smaller gender differences in aggression, $\beta = -.12, Z(102) = -2.49, p < .01$.[14] Similarly,

[12] As would be expected, partitioning the neutral conditions of studies in terms of type of instigation used in their provocation manipulations (i.e., physical attack, insult and negative evaluation, and self-esteem threat) yielded no differences among their mean weighted effect size, $Q_b = 3.69, p < .20$.

[13] The results for the unweighted analysis of the provocation difference effect size showed similar effects for type of instigation. For female participants, the provocation difference effect size was smaller for negative intelligence feedback (mean unweighted effect size = .23) than it was for physical (unweighted effect size = 1.86) and verbal (mean unweighted effect size = .23) provocations, $F(2, 26) = 5.38, p < .05$. For male participants, the provocation difference was somewhat smaller for verbal provocation (mean unweighted effect size = .23) than for physical provocation (mean unweighted effect size = 1.00) and negative intelligence feedback (mean unweighted effect size = .82), $F(2, 26) = 2.47, p = .11$. We could not perform contrasts with frustation manipulations because there was an insufficient number of provocation difference effect sizes in this category.

[14] In 1995, the dataset of effect sizes was updated to 1995 from a prior inclusion date of 1991. The judgments were made on the set of studies published in 1991 or earlier. Two studies (Lightdale & Prentice, 1994; Richardson, Hammock, Smith, Gardner, & Signo, 1994) published in 1994 were not replications of any paradigm previously included in the meta-analysis. Therefore, no judgments for these studies were available to enable their inclusion in the judgment analyses.

Table 7
Gender Difference Effect Size as a Function of Type of Instigation

Category	*k*	*M* unweighted effect size	*M* weighted effect size	95% CI	Q_w
Neutral	50	.43	.33	0.23–0.42	136.98***
Provocation					
Insult or negative evaluation	16	−.32	−.13	−0.28–0.03	22.87
Physical attack	13	.08	.12	−0.08–0.33	16.45
Frustration	15	.17	.24	0.10–0.37	23.47
Negative intelligence feedback	8	.59	.55	0.36–0.74	4.33

Note. The control condition effect size is reported to provide a baseline comparison. Positive effect sizes indicate a difference in the direction of greater aggression among men; negative effect sizes indicate a difference in the direction of greater aggression among women. *k* = number of effect sizes in the category; CI = confidence interval; Q_w = homogeneity within class. CIs and Q_ws are associated with the mean weighted effect size.
*** $p < .001$.

a separate regression analysis of the judges' pooled ratings of negative affect, again using all 105 observations, indicated that higher levels of negative affect were related to smaller gender differences in aggression, $\beta = -.13$, $Z(102) = -2.65$, $p < .01$. Two separate regression analyses conducted on only the data from the provocation conditions showed that both provocation intensity, $\beta = -.32$, $Z(53) = -3.55$, $p < .001$, and negative affect, $\beta = -.38$, $Z(53) = -3.98$, $p < .001$, predict gender differences in aggression under conditions of provocation. The results of the comparison between the average effect sizes for the provocation and neutral conditions, as well as the assessment of effects of the mean judged level of provocation and negative affect, support the prediction that provocation will reduce the magnitude of gender differences in aggression.

Gender Differences in Subjective Appraisals of Provocation, Negative Affect, and Danger

Eagly and Wood (1991) advised that, if differences between the social roles of men and women cause observed gender differences in aggressive behavior, variables that reflect the consequences of these roles should mediate these differences in aggression. Thus, men's and women's perceptions or cognitive appraisals of situations can be expected to differ. Such differences between the appraisals made by female and male participants in each experiment should be paralleled by differences between judgments about these same relevant situational variables that were made by our female and male judges. For instance, female judges' ratings of danger from retaliation should exceed those of the male judges (Eagly & Steffen, 1986). If so, the potential mediating role of this difference in predicting a gender difference in aggression can be tested. In sequence, then, we used paired sample *t* tests to examine the degree to which men and women differ in their cognitive appraisals of provocation intensity, negative affect, and potential danger of retaliation from the individual exposed to the participants' aggressive action.[15] We then assessed whether differences in these appraisals, as seen in the ratings made by our female and male judges, predicted gender differences in aggression.

As noted previously, female and male judges provided ratings both of the neutral and of the provocation conditions because the neutral conditions sometimes included features that could be interpreted by participants as somewhat aversive or provoking. A comparison of the average female and the average male ratings of provocation intensity, pooled across neutral and provocation conditions, showed that female judges, $M = 3.42$, were marginally less likely to rate conditions as provoking than were male judges, $M = 3.82$, $t(103) = 1.62$, $p < .06$. When the provocation intensity ratings were examined within only the provocation conditions, this tendency for the female judges to perceive less provocation was reliable, female $M = 4.89$, male $M = 5.58$, $t(55) = 2.33$, $p < .01$.

By contrast, female and male judges' ratings of negative affect did not differ, whether pooled across conditions, $t(103) = 1.93$, $p > .20$, or examined separately within the provocation conditions, $t(55) = 1.40$, $p > .10$. Because no analysis provided evidence of a reliable difference between female and male judges' ratings of negative affect, this variable is not considered further.

Finally, our judges' ratings of danger from retaliation showed that female judges, $M = 3.34$, estimated that they would face more danger (pooled across neutral–control and provocation conditions) than did male judges, $M = 3.00$, $t(103) = 2.42$, $p < .01$. Though it did not reach significance, analysis for only the provocation conditions showed that the female judges' mean rating for danger from retaliation, $M = 3.47$, when compared with that for male judges, $M = 3.45$, $t(55) = 1.00$, $p > .20$, showed a similar direction of effect.

If gender differences in these cognitive appraisals mediate gender differences in aggression, then the differences between the ratings of female and male judges are related to larger gender differences in aggression. A series of weighted least squares regression analyses, corrected to proper *SE*s and *df*, were conducted to assess the relationships between gender differences in aggression and gender differences both in ratings of provocation intensity and ratings of danger of retaliation. For the gender difference in ratings of provocation, the average of the two female ratings of provocation was subtracted from that of the two male ratings; for the gender difference in ratings of danger of

[15] All of the judgments were averaged across same gender judges and log transformed to normalize their distribution.

retaliation, the average of the male ratings of danger of retaliations was subtracted from that of the female ratings. As shown in Table 8, when entered as the sole predictor, gender differences in perceived provocation intensity were positively related to gender differences in aggression, $Z(102) = 5.98, p < .001$, indicating that when men perceive more provocation than women, the difference between male and female aggression is greater. In addition, when entered as a sole predictor, gender differences in danger of retaliation were positively related to gender differences in aggression, $Z(102) = 2.64, p < .01$, indicating that when women fear more danger of retaliation than men, the difference between female and male aggression is greater. When these two variables were entered simultaneously into a regression equation, gender differences in both provocation intensity, $Z(101) = 5.79, p < .001$, and perceptions of danger from retaliation, $Z(101) = 2.40, p < .05$, predicted gender differences in aggression.[16] The test of the model specification, however, indicated that there was a significant amount of variation left unexplained by the two predictors, $Q_E(101) = 216.02, p < .001$.

In a final regression analysis, these gender difference variables were simultaneously entered with type of condition (provocation vs. neutral). As can be seen in Table 8, type of condition, $Z(100) = -2.97, p < .01$, gender differences in provocation intensity, $Z(100) = 4.79, p < 1$, and gender differences in danger of retaliation, $Z(100) = 2.05, p < .05$, all remain predictors of gender differences in aggression.[17] Taken together, these results suggest that gender differences in perceived provocation intensity and perceived danger from retaliation partially mediate gender differences in aggression, but their contribution does not eliminate the independent contribution of experimental manipulations of provocation (or other, as yet unspecified mediators) in predicting when men and women will or will not differ in aggression. That is, the test of the model specification, once again, indicated that even with all three predictors in the model, substantial variation in gender differences remained, $Q_E(100) = 207.17, p < .01$.

Table 8
Regression Analyses of the Effect of Gender Differences in Appraisals

Variables in regression	Univariate model		Multivariate model		Multivariate model with categorical variable	
	B	β	*B*	β	*B*	β
Continuous						
Gender difference in provocation intensity	.62	.37**	.60	.36**	.71	.43**
Gender difference in danger of retaliation	.20	.18**	.17	.15*	.15	.13*
Categorical						
Type of experimental condition					−.09	−.20**

Note. *B* = unstandardized beta weight; β = standardized beta weight.
* $p < .05$. ** $p < .01$.

Contextual Variables

Analyses of the effect of the provocation dichotomy (neutral vs. provocation condition) were conducted separately for each of four categorical variables: (a) aggressive opportunity, (b) gender of target, (c) gender of experimenter, and (d) type of design (laboratory or field).[18] We discuss the effects of laboratory versus field paradigms on the gender difference effect sizes. The results for the provocation difference effect sizes are reported in each of these subsections when they help to clarify the pattern of results of the gender difference effect sizes.

Type of aggression. As shown in the lower panel of Table 9, whether provoked or not, men were more aggressive when the aggressive act consisted of manual delivery of shock or noise. Under neutral conditions, the average gender difference effect size was only slightly larger when the study required physical aggression, $d+ = 0.36$, as opposed to verbal aggression, $d+ = 0.30$, $\chi^2(1) = 0.35, p < .53$. In contrast, under provocation conditions, gender differences were larger when participants were physically aggressive, $d+ = 0.30$, compared with when they were verbally aggressive, $d+ = 0.05$, $\chi^2(1) = 7.39, p < .01$. Blaring a horn at a person who did not proceed through an intersection has been conceptualized as aggression because it involves bombarding the individual exposed to the unpleasant stimuli (Doob & Gross, 1968). When horn honking was the measure of aggression, the gender difference, $d+ = 0.30$, was equivalent to that for physical aggression and larger than verbal aggression, $\chi^2(1) = 4.60, p < .05$. An analysis comparing all three types of aggression under provocation yielded a significant between-class effect, $Q_b(2) = 8.78, p < .05$. The within-class homogeneity statistics show that the specified model fits well for the provocation conditions but not for the neutral conditions. In general, these patterns of differences are also revealed in the unweighted analyses, however, the unweighted mean effect sizes under the neutral conditions tend to be larger than those in the provocation conditions.

The results for the provocation difference effect sizes suggest that the gender differences in physical aggression and the ab-

[16] This same pattern of results emerged when the simultaneous regression was performed only for the provocation conditions. Gender differences both in perception of provocation intensity, $\beta = 39$, $Z(52) = 4.36, p < .01$, and danger from retaliation, $\beta = 27$, $Z(52) = 2.97, p < .01$, predicted gender differences in aggression.

[17] In addition, a stepwise regression supported the importance of the categorical provocation variable in predicting gender differences in aggression. In this analysis, gender differences between judges' ratings both of perception of danger of retaliation and provocation intensity were entered into the first step. To assess whether the experimental inductions of provocation predicted residual variance in gender differences in aggression, we entered experimental condition in a subsequent step. The results showed that this categorical variable accounted for significant additional variance in gender differences in aggression, $F_{change}(100) = 8.09, p < .001$. By implication, other unmeasured mediators affect gender differences in aggression, provocation has direct effects, or both are true.

[18] The effects of other contextual variables are not reported here. These variables did not interact with level of provocation or gender of participant, and the outcomes were generally consistent with those reported by Eagly & Steffen (1986).

Table 9
Weighted Gender Difference Effect Sizes as a Function of Type of Aggressive Response

Category	*k*	*M* unweighted effect size	*M* weighted effect size	95% CI	Q_w
Neutral					
Physical attack	35	.48	.36	0.25–0.47	105.87***
Verbal aggression	13	.36	.30	0.12–0.48	26.98
Provocation					
Physical attack	26	.21	.30	0.17–0.43	31.31
Verbal aggression	20	−.11	.05	−0.08–0.18	36.19*
Horn honking	8	.16	.30	0.11–0.49	19.90**

Note. Positive effect sizes indicate greater aggression among men. *k* = number of observations in the category; CI = confidence interval; Q_w = homogeneity within class. CIs and Q_ws are associated with the mean weighted effect size.
* $p < .05$. ** $p < .01$. *** $p < .001$.

sence of differences in verbal aggression may be due to the differential responsiveness of women. For female provocation difference effect sizes, verbal aggression, $d+ = 1.19$, exceeded that for physical aggression, $d+ = 0.72$, $\chi^2(1) = 10.31$, $p < .01$, whereas male provocation difference effect sizes for verbal, $d+ = 0.72$, and physical aggression, $d+ = 0.86$, $\chi^2(1) = 1.09$, $p < .30$, did not differ.[19] In summary, men are more apt to be physically aggressive than women even under conditions of provocation, but men and women do not differ in verbal aggression even under neutral conditions. Finally, provocation appears to have a greater effect on female verbal aggression than on female physical aggression.

Gender of target. There were three categories in the analysis of the effect of gender of target: female target, male target, and actor same gender as target. As shown in the upper panel of Table 10, type of target had no effect under neutral conditions, $Q_b(2) = 1.05$, $p < .70$, and the homogeneity statistic reveals variability within the classes. However, for the provocation conditions, there appears to be sufficient homogeneity within classes, but the difference between these classes was not significant, $Q_b(2) = 3.04$, $p < .30$. Looking at the overall pattern of effects, it appears that when the target and actor are the same gender, a relatively large gender difference emerges under neutral conditions, but provocation eliminates this difference. This pattern of results is stronger in the analyses of the unweighted effect sizes. For each of the five studies, the gender difference effect sizes could be calculated such that gender of target was either the same as, or opposite to, that of the actor. For example, the same gender target category included effect sizes that represented the female actors' aggression toward female targets, subtracted from male actors' aggression toward male targets. Analysis of these provocation conditions showed that when gender of target and actor differed, gender differences emerged, $d+ = 0.82$, CI = 0.49–1.15, but when they were of the same gender, there was no difference in aggression, $d+ = -0.01$, CI = −0.27–0.25). Additionally, the difference between these mean effect sizes was significant, $\chi^2(1) = 15.16$, $p < .0001$.

Provocation difference effect sizes also allow categorization of the effect sizes into same versus opposite gender targets. As depicted in the lower panel of Table 10, the provocation difference effect size for women is larger when their targets are female participants than when they are male participants, $\chi^2(1) = 19.92$, $p < .0001$; whereas for men, although the provocation difference effect size is directionally reversed as a function of target's gender, this difference does not approach significance, $\chi^2(1) = 1.08$, $p = .30$. Nevertheless, the heterogeneity of the effect sizes within each class suggests that other variables contribute to differences between the provocation difference effect sizes.

Taken together, the analyses of both types of effect sizes suggest that under neutral conditions, women may not be inclined to act aggressively toward female targets, but men, who are perhaps encouraged by their gender roles, appear to be more aggressive toward male targets. This gender difference in aggressive inclinations toward same gender target diminishes under provocation; both men and women behave aggressively toward same gender targets.

Gender of experimenter. As shown in the upper panel of Table 11, the gender difference weighted effect sizes, when categorized by gender of experimenter (female, male, or actor same gender as target), differed marginally under neutral conditions, $Q_b(2) = 4.79$, $p < .10$. Under provocation, however, gender differences were moderated by gender of experimenter, $Q_b(2) = 8.19$, $p < .05$, and the homogeneity statistic suggested that the effect sizes were sufficiently similar within classes. Simple contrast analyses showed that, for provocation conditions, the gender difference in aggression was smaller when the experimenter was the same gender as the participant than when the experimenter was a man, $\chi^2(1) = 8.00$, $p < .01$, but not when the experimenter was a woman, $\chi^2(1) = 2.19$, $p < .20$. Finally, the mean effect size for studies with only female experimenters did not differ from that with only male experimenters, $\chi^2(1) = 2.78$, $p < .10$. The analyses of the unweighted analyses showed a similar pattern of differences between mean effect sizes. Unfortunately, no studies allowed calculation of a gender difference effect size for conditions in which experimenters were the other gender of each group (female or male) of participants.

[19] The unweighted analyses suggest that type of aggression does not moderate the provocation difference effect size. For female participants, this effect size for physical aggression was 1.15 and for verbal aggression was 1.30, $F(1, 28) = .09$, $p < .80$. For male participants, the provocation difference effect size for physical aggression was .66 and for verbal aggression was .67, $F(1, 28) = 0.00$.

Table 10
Gender Difference Effect Sizes as a Function of Gender of Target

Category	*k*	*M* unweighted effect size	*M* weighted effect size	95% CI	Q_w
		Gender difference effect size			
Neutral condition					
Female target	8	0.27	0.23	−0.11–0.56	3.10
Male target	14	0.20	0.29	0.12–0.46	29.03**
Same as participant	20	0.74	0.38	0.24–0.52	98.05***
Provocation condition					
Female target	11	0.19	0.23	0.05–0.42	21.73*
Male target	18	0.18	0.24	0.10–0.38	20.40
Same as participant	18	−0.14	0.07	−0.08–0.22	50.99**
		Provocation difference effect size			
Female participants					
Female target	24	1.53	1.04	0.87–1.20	168.02***
Male target	6	0.15	0.33	0.07–0.59	38.99***
Male participants					
Female target	6	0.52	0.54	0.21–0.88	7.93
Male target	24	0.69	0.74	0.59–0.88	118.91***

Note. Positive gender difference effect sizes indicate greater aggression among men. Positive provocation difference effect sizes indicate greater aggression under provocation conditions than under control conditions. *k* = number of observations in the category; CI = confidence interval; Q_w = homogeneity within class.
* $p < .05$. ** $p < .01$. *** $p < .001$.

The results for the provocation difference effect sizes, shown in the lower panel of Table 11, confirmed that for female participants, the difference between aggression under provoking versus neutral conditions was much larger when the experimenter was a woman than a man, $Q_b(1) = 12.05, p < .001$. By contrast, the difference in male aggression as a function of provocation was not affected by the experimenter's gender, $Q_b(1) = 1.80, p < .20$.

Laboratory versus field studies. Several conditions compromised our comparison of the laboratory and field research.

Table 11
Gender Difference Effect Sizes as a Function of Gender of Experimenter

Category	*k*	*M* unweighted effect size	*M* weighted effect size	95% CI	Q_w
		Gender difference effect size			
Neutral condition					
Female experimenter	6	0.46	0.52	0.20–0.83	8.24
Male experimenter	13	0.30	0.32	0.14–0.50	18.70
Same as participant	7	0.85	0.64	0.41–0.87	28.57***
Provocation condition					
Female experimenter	9	0.01	0.01	−0.19–0.21	10.42
Male experimenter	9	0.16	0.27	0.05–0.48	5.21
Same as participant	5	−0.58	−0.25	−0.52–0.03	16.41**
		Provocation difference effect size			
Female participants					
Female experimenter	11	1.70	1.24	0.98–1.50	88.45***
Male experimenter	10	0.80	0.70	0.45–0.94	11.35
Male participants					
Female experimenter	7	0.69	0.50	0.15–0.85	5.17
Male experimenter	14	0.77	0.70	0.50–0.90	62.67***

Note. Positive effect sizes indicate greater aggression among men. Positive provocation difference effect sizes indicate greater aggression under provocation conditions than under control conditions. *k* = number of observations in the category; CI = confidence interval; Q_w = homogeneity within class.
** $p < .01$. *** $p < .001$.

First, within this dataset, none of the neutral conditions were derived from the field studies. Second, for the provocation conditions, the 11 observations derived from the field studies were limited to two types of instigations. Except for one study in which the provocation consisted of a confederate who adopted a rude tone, all of the instigations were manipulations of frustration—impeding an actor's progress while waiting either in a line or at an intersection. Keeping these constraints in mind, although the absolute magnitude of gender differences, as indicated by the weighted means, was somewhat larger under conditions of provocation in the field, $d+ = 0.22$, CI = 0.07–0.38, than in laboratory studies, $d+ = 0.15$, CI = 0.05–0.24, this difference was not reliable, $Q_b(1) = 0.67$, $p < .50$. Also, recognizing that the aggression measures were limited to either horn honking (elicited by blocking progress through intersection) or verbal aggression (in response to a confederate making a phone call or cutting into a line), the direction of effect of nonverbal versus verbal aggression for field studies was the same as that for the entire dataset. That is, the gender difference tends to be larger when aggression was measured as horn honking, $d+ = 0.30$, CI = 0.11–0.49, than when measured as negative spoken comments, $d+ = 0.07$, CI = −0.20–0.34, but this difference was not significant, $Q_b(1) = 1.84$, $p < .20$.[20]

Discussion

Moderating Effects of Provocation

The mean gender difference effect size of .22 for our sample of studies, $d+ = 0.24$, mean unweighted effect size = .23, was similar to that reported by Eagly and Steffen (1986), $d+ = 0.29$. These results suggest that, within the paradigms used to study aggression experimentally, men are generally more aggressive than women. This overall mean effect size, however, masks the moderating effect of provocation on gender differences in aggression. When the effect sizes are subdivided into those derived from provocation conditions and those derived from relatively neutral conditions, the moderating effect of provocation clearly can be seen—provocation greatly reduces gender differences in aggression. This outcome is supported further by the inverse relationship between judges' assessments of provocation intensity and the magnitude of gender differences in aggression. These results are all the more striking when one considers that the average level of perceived provocation for these studies was relatively moderate. The judges' mean rating of the provocation conditions was 5.25 on a 9-point scale, only slightly exceeding the scale midpoint. Thus, under truly high levels of provocation, there may be little reason to anticipate differential aggression on the part of men and women. This proposition is supported by research that shows that men and women are equally aggressive in their intimate relationships—a context in which provocation may often precede aggression (e.g., Malone, Tyree, & O'Leary, 1989; O'Leary et al., 1989). Our findings are also consistent with those reported by Lightdale and Prentice (1994), who found that men and women were equally aggressive when gender role considerations were diminished by means of a manipulation designed to create a feeling of deindividuation. Taken together, these results support a gender role approach, suggesting that provocation diminishes the impact of gender role norms on gender differences in aggression (Eagly & Steffen, 1986; Frodi et al., 1977).

Gender Differences in Appraisals of Danger, Provocation, and Negative Affect as Mediators of Gender Differences in Aggressive Responding

Our outcomes showed that gender differences in appraisals of danger are important. Beyond confirming Eagly and Steffen's (1986) findings that women differ from men in their assessment of the degree to which a situation might evoke dangerous retaliation, these gender differences in the appraisals of danger predict gender differences in aggression. It appears that the more women's fear of aggressive retaliation exceeds that of men, the larger the gender difference in aggression. The fact that differential appraisals of danger by male and female judges were related to differences in aggressive responding is especially interesting given that (a) experimental inductions were objectively identical for female and male participants and (b) with few exceptions, experimental paradigms provided little if any expectation that targets would have opportunities to retaliate.

In light of the strong relation between provocation and aggression, it should not be surprising that differences between men and women in their appraisals of provocation intensity also were associated with gender differences in aggression. In their interpretations of the degree to which instigations are provoking, the male judges rated instigations as more provoking than did the female judges. Moreover, these differences in appraisals at least partially mediated differences in the aggressive behavior of the male and female participants in the studies. Appraisals of negative affect, however, presented a different picture. As assessed by the pooled judgments of our male and female judges, higher levels of negative affect, such as higher provocation intensity, did attenuate gender differences in aggression. According to Berkowitz's (1989) theorizing, negative affect mediates aggression. Thus, these aspects of our results do support this view. However, as we argued previously, his theoretical perspective also seemingly requires that, if men and women differed in their aggressiveness, they must also have differed in the intensity of their subjective experience of negative affect. This latter expectation did not receive support. It is interesting, however, that although the female and male judges' ratings of negative affect did not differ reliably and that, consequently, gender differences in subjective negative affect did not affect gender differences in aggression, nevertheless, gender differences in subjective ratings of provocation intensity, like those for danger of retaliation, did predict gender differences in aggression. It is also interesting that gender differences in ratings of negative affect and provocation intensity were not strongly correlated, $r = .15$. These results, like that for differential appraisals of danger of retaliation, emphasize the cognitive or interpretive dimension of aversively instigative events as more important than their affective compo-

[20] The unweighted for due comparison between experimental (mean unweighted effect size = .04) and field studies (mean unweighted effect size = .15) also did not reveal differences between these methods, $F(1, 55) = .30$, $p < .60$. The comparison between verbal aggression (unweighted effect size = .13) and physical aggression (mean unweighted effect size = .16) was also nonsignificant, $F(1, 9) = .01$.

nent per se. As such, these outcomes do fit with the emphasis that Berkowitz placed on the cognitive aspect of negative emotional experience.

Finally, experimental manipulation of provocation had a residual effect of reducing gender differences in aggression, even when the effects of all of our measures of gender differences in cognitive appraisals were removed. This, too, seemingly counters Berkowitz's (1989) theorizing. That is, even if one interprets judged provocation intensity and judged fear of retaliation—cognitive appraisals—as counterparts of subjectively experienced negative affect, they do not fully account for the direct effect of manipulations of provocation intensity, as required by Berkowitz's mediational model.

Issues Concerning the Generalizability of Our Results

Although our meta-analytic integration of experimental studies suggests substantial equivalence between the aggressive behavior of men and women who are provoked, the ecological validity of the paradigms used in aggression research warrants consideration (Eagly & Steffen, 1986; Macaulay, 1985). More generally, Berkowitz and Donnerstein (1982) have argued that laboratory studies can have external validity and thus are generalizable, even if they lack ecological validity. Eagly and Steffen (1986) suggest that the conclusions of the meta-analytic findings in gender differences in aggression may be compromised by the tendency for experimental studies to involve aggression between strangers in relatively limited contexts (see also Bjorkqvist & Niemela, 1992; and Macaulay, 1985). In the sections that follow, we discuss issues relevant to the generalizability of the experimental findings we meta-analyzed.

Gender difference in aggression under relatively neutral conditions. Why are men more aggressive than women in the absence of provocation? Some theorists argue that gender differences in aggression arise from biological differences in aggressive readiness (e.g., Maccoby & Jacklin, 1974). The greater willingness of men to act aggressively in nonprovoking situations may reflect this difference in biological predisposition. Alternatively, or additionally, male gender role norms may encourage aggression under conditions of minimal justification. However, it may be that situations that are either ambiguous or devoid of obvious provocation are more likely to be interpreted as provoking by men than by women. This has been suggested by others (Crick & Dodge, 1994; Dodge & Coie, 1987) and is also seen in our own data. Berkowitz (1989) has argued that persons bound by social rules that strictly regulate their aggressive behavior may deny the possibility of provocation unless malicious intent is obvious. By extension, the female gender role may inhibit aggression when the situation is ambiguous or when the ill intention of another is less than obvious. If gender differences in the interpretation of situations, in gender roles, or both (perhaps in conjunction with biology) do explain gender differences in aggression, then efforts to reduce the aggressive behavior of men may be revealed by understanding the particular elements of the female gender role that discourage aggression, as well as those of the male gender role that encourage interpreting stimuli or actions as provoking. As Geen (1990) argued,

> [many] people, though instigated to aggress as much as others, do not possess the background characteristics of aggressors. If this is the case, then society's task is to discover the conditions under which . . . relatively unaggressive tendencies are fostered and to seek to implement them more widely. (p. 7)

Although our findings suggest that men are more aggressive than women under neutral conditions, in terms of Cohen's (1969) qualitative labeling of the magnitude of effect sizes, our obtained effect was "small." However, the magnitude of this gender difference may underestimate the gender differences in aggression observed in people's daily lives. There are several likely reasons for thinking so. First, by design, experimental studies explicitly hold constant or eliminate all variables that may influence aggression, other than those explicitly manipulated. By contrast, a confluence of cues that are absent from the neutral conditions of experimental studies may typically be present in natural settings. For example, experimental studies of aggression often control interpersonal interaction between the individual exposed to the aggression and the aggressor. As such, these studies may eliminate features of interaction that additionally promote differential interpretation of provocation by men and women in everyday circumstances (Crick & Dodge, 1994; Dodge & Coie, 1987). In addition, the absence of other cues that ordinarily characterize natural settings—such as the availability of weapons, status differences between the actor and target of aggression, or positive social or material consequences—may further account for the relatively small magnitude of gender differences found in the neutral conditions of the experimental studies we analyzed.

Finally, note that, as a result of media emphasis, the instances of aggression in everyday life that seem to receive much of the public's attention are those that involve physical attack and are physically damaging to the individual experiencing the attack—a type of aggressive response that our analyses show to be more prevalent among men than women. Consequently, the relatively moderate gender difference found for the neutral conditions may appear incongruent with the common view that men are considerably more aggressive than women because news media reports of aggression highlight vivid instances of physical aggression. At the same time, however, note that experimental paradigms contain procedures and elicit normative constraints that ordinarily proscribe strong displays of aggression.

Type of instigation. The types of provocations used in experimental studies of aggression are often encountered in everyday life. They include minor frustrations, such as having one's progress impeded by another not proceeding though an intersection, and direct attacks that are both verbal—as with an insult from a peer—and physical—as when electric shock is administered. Our findings, however, suggest that specific types of instigations differentially moderate the degree to which experimental provocations increase aggression among men and women. At the same time, experimental instigations may differ from those in natural settings by virtue of experimenters' explicit intent to provoke the participants. Conceivably, potential instigations in natural settings are more ambiguous. If so, men may be more aggressive than women in everyday life, in part because, as suggested previously, they are more apt to interpret ambiguous situations as provoking. In retrospect, it is unfortunate that we did not have our female and male judges rate the

ambiguity of the experimental provocations, enabling us to better explore this issue.

One laboratory procedure used to provoke participants exposed them to disparaging information about their intellectual capabilities. Although female and male participants respond similarly when either physically attacked or negatively evaluated by peers, when the provocation consisted instead of feedback from a person of authority who challenged the adequacy of their intellectual capacity, gender differences in aggression not only were considerably larger than those found in the absence of provocation but also exceeded those observed under other operationalizations of provocation. Interestingly, research suggests that women experience little affective reaction to this intended provocation (Berg, Stephan, & Dodson, 1981; Crandall, 1969; Deaux & Farris, 1977; Marini & Greenberger, 1978; Roberts, 1991). Indeed, when Pytkowicz, Wagner, and Sarason (1967) used negative intelligence feedback as a provocation, their female participants did not report greater hostility in response to negative performance feedback than to positive performance feedback, whereas male participants did. Deaux (1984) has argued that women often have low performance expectations, especially on tasks typically used in experimental manipulations of negative intelligence feedback. Not expecting to do well, they may experience little negative affect as a consequence of feedback indicating poor performance (Deaux, 1984). This suggestion is in line with Berkowitz's (1989) contention that people are apt to be aggressive when there is an expectation of a goal but not when there is merely a deprivation of a goal.

Alternatively, it may be that women respond to negative comments about their intellectual ability by feeling sad or shameful instead of angry (Van Goozen et al., 1994). Thus, in the face of disappointing feedback about their performance, they may see little reason to attack a figure of authority who provides this information. The effects of sadness on aggression are likely to differ considerably from other negative emotional states, such as anger or frustration, which direct attention outward. Manipulations that induce sadness not only elicit physiological responses that differ markedly from those associated with anger (Henry, 1986) but also evoke subjective self-reports about internal bodily states quite unlike those associated with anger (Shields, 1984). Although there have been some experimental circumstances in which those induced to feel sad behaved more aggressively, this primarily has appeared to be true of persons with impulsive personalities and because the situation required them to react quickly with little thought (Finman & Berkowitz, 1989; Hynan & Grush, 1986). In general, sadness is more likely to elicit an inner-directed rather than an other-directed focus (Carlson & Miller, 1988; Wood, Saltzberg, & Goldsamt, 1990). Self-focused ruminating may, in turn, reduce aggressive responding.

By contrast, when the intellect of men was challenged by the experimenter, they responded with relatively high levels of aggression in comparison with neutral conditions. Moreover, our judges' ratings indicated that men found negative intelligence feedback much more provoking than did women. In summary, this manipulation may lack validity as a provocation for women. At the same time, these findings highlight the importance of differences in perceptions of provocation intensity or interpretations of various incidents as provoking.[21]

Categorical predictors of gender differences in aggression. Consistent with other research (cf. Eagly & Steffen, 1986), larger gender differences were found when the response required physical as compared with verbal aggression. Moreover, for female participants, the effect of provocation was stronger when the experimental paradigm allowed verbal as opposed to physical aggression. Gender differences in physical aggression seem to emerge early in development and continue over time (Maccoby & Jacklin, 1974). Apparently, gender typing of acceptable behaviors acts to bias behavioral choices in aggression (Perry & Bussey, 1979). Bandura (1973) argued that girls typically display less physical aggression than boys because they learn that they are not supposed to be aggressive, whereas boys learn that physical aggressiveness is appropriate and even desirable. In this same vein, a survey of children (Huesmann, Guerra, Zelli, & Miller, 1992) showed that boys and girls hold different standards about the appropriateness of aggressive behavior; boys not only believed it is more acceptable but also reported being more physically aggressive than did girls.

Although many have argued that commonly used measures of aggression are valid, nevertheless, they may not adequately reflect the types of aggressive strategies most likely to be used by women. Bjorkqvist, Osterman, and Lagerspetz (1994) and Lagerspetz, Bjorkqvist, and Peltonen (1988) have shown that adolescent girls and women make greater use of indirect social aggression (e.g., gossiping or refusing to speak to another). To complement the understanding of gender differences in aggression, future studies also should examine gender differences in the use of indirect types of aggression (Bjorkqvist, Osterman, & Kaukiainen, 1992).

Deaux (1984) and Eagly (1983) have also argued that natural settings typically offer more behavioral choices than do laboratory settings. Consequently, the results from experimental studies of aggression may reflect aggressiveness that is found only when other behavioral choices are unavailable. When other behavioral choices are available, overall rates of aggressive responding may decline. How might an increased availability of behavioral options interact with provocation and gender of participant to modify the results we have reported?

From the gender role perspective, alternative response options are likely to appeal more to women than men. If so, behavioral choices are likely to attenuate the tendency of provocation to reduce gender differences in aggression. Among the few studies in the present sample that allowed participants a behavioral choice (e.g., a participant could either block or shock their opponent), only Yinon, Jaffe, and Feshbach (1975) comparatively reported use of the nonaggressive alternative by female and male participants. In support of the expectation derived from the gender role perspective, when given a choice between providing feedback by a red light or an electric shock, only 2 of 20 male participants (10%), but 5 of 14 female participants (36%), sometimes chose the nonaggressive red light signal. At

[21] Our discussion here is not meant to portray women as pathologically self-blaming. If anything, the response of women to this manipulation strikes us as more rational than that of men, who seemingly are blaming the messenger for their own shortfall.

the same time, note that women may have been aggressive when they were provoked, not only because direct aggression was the readily available response but also because danger of retaliation was unlikely. Only researchers of 22% of the studies included in this meta-analysis used an experimental paradigm that clearly led participants to expect a retaliation.

Gender differences in aggression may also be affected by gender of the target of aggression. Our results suggest that when provoked, women and men may behave more aggressively toward same gender targets. These findings are somewhat consistent with those that show that men are more aggressive toward male than female targets (Buss, 1966a; Harris, 1992a; Taylor & Epstein, 1967) and that chivalry norms discourage men from acting aggressively toward women (Bjorkqvist & Niemela, 1992; Eagly & Steffen, 1986). They are also in consonance with self-reports of female and male participants, which show that gender differences are less prevalent when the target is a woman (Harris, 1992b) and that women direct their aggression toward other women, whereas men primarily direct it toward other men (Harris, 1992a). To be aggressive toward persons of the same gender may be more culturally acceptable (Burbank, 1987). Yet, as previously noted, studies of intimate couples show that women, as well as men, can be quite violent with their opposite gender partners (e.g., Malone et al., 1989; O'Leary et al., 1989; Stets & Straus, 1989; Sugarman & Hotaling, 1989). Finally, because experimental studies often confound gender of target with other variables, firm conclusions about its effects are hazardous without further research (Eagly & Steffen, 1986).

Conclusion

Our primary goal was to provide a detailed analysis of the effect of provocation on gender differences in aggression. Our meta-analytic findings provide substantial, consistent, and strong evidence that provocation diminishes the commonly reported outcome of greater aggressiveness by men than by women. Consequently, they modify the previous meta-analytic conclusions of Eagly and Steffen (1986) with respect to this specific issue. Nevertheless, they do support the tenets of the social role approach, as elaborated by Eagly and Steffen and others (e.g., Deaux, 1971; Frodi et al., 1977).

Our outcomes with respect to our secondary goal of assessing mediators of the attenuating effect of provocation on gender differences in aggression were instructive. Contrary to our extrapolation from Berkowitz's (1989) theorizing, no support emerged for the expectation that gender differences in appraisals of levels of subjective negative affect would mediate the effects of provocation. At the same time, however, gender differences in subjective appraisals of provocation intensity and danger from retaliation, more cognitive representations of the negative aspects of provocation, did receive support as mediators of the effect of provocation. Finally, differences among types of instigation, gender configuration of interacting parties, type of aggressive response allowed by the setting, as well as other aspects of the setting, all moderated the degree to which provocation reduced gender differences in aggression.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Adams, D. (1992). Biology does not make men more aggressive than women. In K. Bjorkqvist & P. Niemela (Eds.), *Of mice and women: Aspects of female aggression* (pp. 17–25). San Diego, CA: Academic Press.

*Ahmed, S. M. S. (1982). Factors affecting frustration and aggression relationships. *Journal of Social Psychology, 116,* 173–177.

*Anderson, K. B. (1993). *The effects of provocation and uncomfortable temperatures on aggressive affect and behavior.* Unpublished master's thesis, University of Missouri—Columbia.

*Atkinson, C., & Polivy, J. (1976). Effects of delay, attack, and retaliation on state depression and hostility. *Journal of Abnormal Psychology, 85,* 570–576.

Axelrod, R. (1984). *The evolution of cooperation.* New York: Basic Books.

Bandura, A. (1973). *Aggression: A social learning analysis.* Englewood Cliffs, NJ: Prentice Hall.

Baron, B. A., & Richardson, D. R. (1994). *Human aggression* (2nd ed.). New York: Plenum Press.

*Baron, R. A. (1979). Heightened sexual arousal and physical aggression: An extension to females. *Journal of Research in Personality, 13,* 91–102.

*Baron, R. A., & Ball, R. L. (1974). The aggression-inhibiting influence of nonhostile humor. *Journal of Experimental Social Psychology, 10,* 23–33.

*Baron, R. A., & Bell, P. A. (1977). Sexual arousal and aggression by males: Effects of type of erotic stimuli and prior provocation. *Journal of Personality and Social Psychology, 35,* 79–87.

Benton, D. (1992). Hormones and human aggression. In K. Bjorkqvist & P. Niemela (Eds.), *Of mice and women: Aspects of female aggression* (pp. 37–46). San Diego, CA: Academic Press.

Berg, J. H., Stephan, W. G., & Dodson, M. (1981). Attributional modesty in women. *Psychology of Women Quarterly, 5,* 711–727.

*Berkowitz, L. (1960). Repeated frustrations and expectations in hostility arousal. *Journal of Abnormal and Social Psychology, 60,* 422–429.

Berkowitz, L. (1988). Frustrations, appraisals, and aversively stimulated aggression. *Aggressive Behavior, 106,* 59–73.

Berkowitz, L. (1989). Frustration-aggression hypothesis: Examination and reformulation. *Psychological Bulletin, 106,* 59–73.

Berkowitz, L., & Donnerstein, E. I. (1982). External validity is more than skin deep: Some answers to criticisms of laboratory experiments. *American Psychologist, 37,* 245–257.

*Berkowitz, L., & Holmes, D. S. (1959). The generalization of hostility to disliked objects. *Journal of Personality, 27,* 565–577.

*Berkowitz, L., & Knurek, D. A. (1969). Label-mediated hostility generalization. *Journal of Personality and Social Psychology, 13,* 364–368.

Bjorkqvist, K., & Niemela, P. (1992). New trends in the study of female aggression. In K. Bjorkqvist & P. Niemela (Eds.), *Of mice and women: Aspects of female aggression* (pp. 3–16). San Diego, CA: Academic Press.

Bjorkqvist, K., Osterman, K., & Kaukiainen, A. (1992). Indirect aggression: Conceptions and misconceptions. In K. Bjorkqvist & P. Niemela (Eds.), *Of mice and women: Aspects of female aggression* (pp. 51–53). San Diego, CA: Academic Press.

Bjorkqvist, K., Osterman, K., & Lagerspetz, M. J. (1994). Sex differences in covert aggression among adults. *Aggressive Behavior, 20,* 27–33.

Bond, M. H., & Dutton, D. G. (1975). The effect of interaction anticipation and experience as a victim of aggressive behavior. *Journal of Personality, 43,* 515–527.

Borden, R. J. (1975). Witnessed aggression: Influence of an observer's sex and values on aggressive responding. *Journal of Personality and Social Psychology, 31,* 567–573.

Burbank, V. K. (1987). Female aggression in cross-cultural perspective. *Behavior Science Research, 21,* 70–100.

*Buss, A. H. (1963). Physical aggression in relation to different frustrations. *Journal of Abnormal and Social Psychology, 67,* 1–7.

Buss, A. H. (1966a). Instrumentality of aggression, feedback, and frustration as determinants of physical aggression. *Journal of Personality and Social Psychology, 3,* 153–162.

*Buss, A. H. (1966b). The effect of harm on subsequent aggression. *Journal of Experimental Research in Personality, 1,* 249–255.

*Buvinic, M. L. (1975). Sex differences in the aggressive consequences of expressing feelings or opinions. *Dissertation Abstracts International, 36,* 1498B.

*Caprara, G. V. (1982). A comparison of the frustration–aggression and emotional susceptibility hypotheses. *Aggressive Behavior, 8,* 234–236.

*Caprara, G. V., Renzi, P., Alcini, P., D'Imperio, G., & Travaglia, G. (1983). Instigation to aggress and escalation of aggression examined from a personalogical perspective: The role of irritability and emotional susceptibility. *Aggressive Behavior, 9,* 345–351.

*Caprara, G. V., Renzi, P., Amolini, P., D'Imperio, G., & Travaglia, G. (1984). The eliciting cue value of aggressive slides reconsidered in a personalogical perspective: The weapons effect on irritability. *Journal of Social Psychology, 14,* 313–322.

Carlson, M. (1988). *Toward understanding the relationship between unpleasant stimulation and aggressive behavior in humans.* Unpublished doctoral dissertation, University of Southern California.

Carlson, M., Marcus-Newhall, A., & Miller, N. (1989). Evidence for a general construct of aggression. *Personality and Social Psychology Bulletin, 15,* 377–389.

Carlson, M., Marcus-Newhall., A., & Miller, N. (1990). Effects of situational aggression cues: A quantitative review. *Journal of Personality and Social Psychology, 58,* 622–633.

Carlson, M., & Miller, N. (1988). The differential effects of social and nonsocial negative events on aggressiveness. *Sociology and Social Research, 72,* 155–158.

*Carver, C. S. (1975). Physical aggression as a function of objective self-awareness and attitudes toward punishment. *Journal of Experimental Social Psychology, 11,* 510–519.

Chaikin, A. L., Derlega, V. J., Yoder, J., & Phillips, D. (1974). The effects of appearance on compliance. *Journal of Social Psychology, 92,* 199–200.

*Chase, L. J., & Mills, N. H. (1973). Status of frustrator as facilitator of aggression: A brief note. *Journal of Psychology, 84,* 225–226.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

Cooper, H. M., & Lemke, K. M. (1991). On the role of meta-analysis in personality and social psychology. *Personality and Social Psychology Bulletin, 17,* 245–251.

Crandall, V. C. (1969). Sex differences in expectancy of intellectual and academic reinforcement. In C. P. Smith (Ed.), *Achievement-related motives in children.* New York: Russell Sage Foundation.

Craparo, J. S., Hines, R. P., & Kayson, W. A. (1981). Effects of experienced success or failure on self-esteem and problem-solving ability. *Psychological Reports, 49,* 295–300.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115,* 74–101.

*DaGloria, J., & DeRidder, R. (1979). Sex differences in aggression: Are current notions misleading? *European Journal of Social Psychology, 9,* 49–66.

*DaGloria, J., & DeRidder, R. (1981). Aggression in dyadic interaction. *European Journal of Social Psychology, 7,* 189–219.

*Deaux, K. K. (1971). Honking at the intersection: A replication and extension. *Journal of Social Psychology, 84,* 159–160.

Deaux, K. K. (1984). From individual differences to social categories: Analysis of a decade's research on gender. *American Psychologist, 39,* 105–116.

Deaux, K., & Farris, E. (1977). Attributing causes for one's own performance: The effects of sex, norms, and outcome. *Journal of Research in Personality, 11,* 59–72.

Deaux, K. K., & Major, B. (1987). Putting gender into context: An interactive model of gender-related behavior. *Psychological Review, 94,* 369–389.

Deturck, M. A. (1987). When communication fails: Physical aggression as a compliance-gaining strategy. *Communication Monographs, 54,* 106–112.

Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology, 53,* 1146–1158.

Dollard, J., Doob, L., Miller, N., Mowrer, O. H., & Sears, R. R. (1939). *Frustration and aggression.* New Haven, CT: Yale University Press.

Donnerstein, E., & Hallam, J. (1978). Facilitating effects of erotica on aggression against women. *Journal of Personality and Social Psychology, 36,* 1270–1277.

*Donnerwerth, G. V., & Foa, U. G. (1974). Effect of resource class on retaliation to injustice in interpersonal exchange. *Journal of Personality and Social Psychology, 29,* 785–793.

*Doob, A. N., & Gross, A. E. (1968). Status of frustrator as an inhibitor of horn-honking responses. *Journal of Social Psychology, 76,* 213–218.

*Dor-Shav, N. K., & Dolgin, N. (1981). Sexual identification as a facilitator of aggression in a simulated shock-delivery paradigm. *Journal of Social Psychology, 113,* 287–288.

Doyle, M. A., & Biaggio, M. K. (1981). Expression of anger as a function of assertiveness and sex. *Journal of Clinical Psychology, 37,* 154–157.

Eagly, A. H. (1983). Gender and social influence: A social psychological analysis. *American Psychologist, 38,* 971–981.

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation.* Hillsdale, NJ: Erlbaum.

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. (1991). Explaining sex differences in social behavior: A meta-analytic perspective. *Personality and Social Psychology Bulletin, 17,* 306–315.

Eagly, A. H., & Wood, W. (1994). Using research synthesis to plan future research. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 457–483). New York: Russell Sage Foundation.

*Edwards, N. L. (1968). Aggressive expression under threat of retaliation. *Dissertation Abstracts International, 28,* 3470B.

*Epstein, R. (1965). Authoritarianism, displaced aggression, and social status of the target. *Journal of Personality and Social Psychology, 2,* 585–589.

ERIC [Database]. (1966–1994, April). Washington, DC: Educational Resources Information Center [Producer]. Available from: SilverPlatter.

Finman, R., & Berkowitz, L. (1989). Some factors influencing the effect of depressed mood on anger and overt hostility toward another. *Journal of Research in Personality, 23,* 70–84.

*Fischer, D. G., Kelm, H., & Rose, A. (1969). Knives as aggression-eliciting stimuli. *Psychological Reports, 24,* 755–760.

*Fisher, J. L., & Harris, M. B. (1976). Modeling, arousal, and aggression. *Journal of Social Psychology, 100,* 219–226.

*Frodi, A. (1978). Experiential and physiological responses associated with anger and aggression in women and men. *Journal of Research in Personality, 12,* 335–349.

Frodi, A., Macaulay, J., & Thome, P. R. (1977). Are women always less aggressive than men? A review of the experimental literature. *Psychological Bulletin, 84,* 634–660.

*Gaebelein, J. W. (1977). Sex differences in instigative aggression. *Journal of Research in Personality, 11,* 466–474.

*Geen, R. G. (1968). Effects of frustration, attack, and prior training in aggressiveness upon aggressive behavior. *Journal of Personality and Social Psychology, 9,* 316–321.

Geen, R. G. (1990). *Human aggression.* Pacific Grove, CA: Brooks/Cole.

*Geen, R. G., Stonner, D., & Shope, G. L. (1975). The facilitation of aggression by aggression: Evidence against the catharsis hypothesis. *Journal of Personality and Social Psychology, 31,* 721–726.

*Gentry, W. D. (1972). Biracial aggression: Effect of verbal attack on sex of victim. *Journal of Social Psychology, 88,* 75–82.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

*Golin, S., & Romanowski, M. A. (1977). Verbal aggression as a function of sex of subject and sex of target. *Journal of Psychology, 97,* 141–149.

*Goodwin, A. R. (1973). The effects of locus of control and feedback cues on behavioral aggression. *Dissertation Abstracts International, 34,* 2387A.

Gouldner, A. W. (1960). The norm of reciprocity. *American Sociological Review, 25,* 161–178.

Gustafson, R., Hemlin, K., & Soderberg, M. (1987). Does a single exposure to filmed violence affect aggressive tendencies? A case of serendipity. *Psychological Reports, 61,* 724–726.

Hall, J. A., & Rosenthal, R. (1991). Chautauqua: Issues in meta-analysis, testing for moderator variables in meta-analysis issues and methods. *Communication Monographs, 58,* 437–448.

Hall, J. A., Rosenthal, R., Tickle-Degnen, L., & Mosteller, F. (1994). Hypotheses and problems in research synthesis. In H. Cooper & L. Hedges (Eds.), *The handbook of research synthesis* (pp. 17–28). New York: Russell Sage Foundation.

*Hammock, G. S., & Richardson, D. R. (1992). Predictors of aggressive behavior. *Aggressive Behavior, 18,* 219–229.

*Harris, M. B. (1973). Field studies of modeled aggression. *Journal of Social Psychology, 89,* 131–139.

Harris, M. B. (1974). Aggressive reactions to a frustrating phone call. *Journal of Social Psychology, 92,* 193–198.

*Harris, M. B. (1976). Instigators and inhibitors of aggression in a field experiment. *Journal of Social Psychology, 98,* 27–38.

Harris, M. B. (1991). Effects of sex of aggressor, sex of target, and relationship on evaluations of physical aggression. *Journal of Interpersonal Violence, 6,* 174–186.

Harris, M. B. (1992a). Sex and ethnic differences in past aggressive behaviors. *Journal of Family Violence, 7,* 85–102.

Harris, M. B. (1992b). Sex, race, and experiences of aggression. *Aggressive Behavior, 18,* 201–217.

Harris, M. B. (1993). How provoking! What makes men and women angry. *Aggressive Behavior, 19,* 199–213.

*Harris, M. B., & Huang, L. C. (1974). Aggression and the attribution process. *Journal of Social Psychology, 92,* 209–216.

*Harris, M. B., & Klingbeil, D. (1976). The effects of ethnicity of subject and accent and dependency of confederate on aggressiveness and altruism. *Journal of Social Psychology, 98,* 47–53.

Harris, M. B., Liguori, R., & Joniak, A. (1973). Aggression, altruism, and models. *Journal of Social Psychology, 91,* 343–344.

Harris, M. B., & Samerotte, G. (1975). The effects of aggressive and altruistic modeling on subsequent behavior. *Journal of Social Psychology, 95,* 173–182.

Hedges, L. V. (1994). Fixed effects models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 286–299). New York: Russell Sage Foundation.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

*Hedrick, T. E. (1976). Behavioral and autonomic responses to aggression by traditional and nontraditional men and women. *Dissertation Abstracts International, 37,* 4755B.

Henry, J. P. (1986). Neuroendocrine patterns of emotional response. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience* (pp. 37–60). Orlando, FL: Academic Press.

*Hoppe, C. M. (1977). Interpersonal aggression as a function of subject's sex, subject's sex role identification, opponent's sex, and degree of provocation. *Dissertation Abstracts International, 38,* 2939B.

Huesmann, L. R., Guerra, N. G., Zelli, A., & Miller, L. S. (1992). The role of social norms in the development of aggressive behavior. In A. Fraczek & H. Zumkley (Eds.), *Socialization and aggression.* New York and Heidelberg: Springer-Verlag.

Hyde, J. S. (1984). How large are gender differences in aggression? A developmental meta-analysis. *Developmental Psychology, 20,* 722–736.

*Hynan, M. T. (1982). Aggression in a competitive task. *Psychological Reports, 50,* 663–672.

*Hynan, M. T., & Esselman, J. A. (1981). Victims and aggression. *Bulletin of Psychonomic Society, 18,* 169–172.

Hynan, D. J., & Grush, J. E. (1986). Effects of impulsivity, depression, provocation, and time on aggressive behavior. *Journal of Research in Personality, 20,* 157–171.

*Hynan, M., Harper, S., Wood, C., & Kallas, C. (1980). Parametric effects of blocking and winning in a competition paradigm of human aggression. *Bulletin of Psychonomic Society, 16,* 295–298.

Infante, D. A. (1989). Response to high argumentatives: Message and sex differences. *Southern Communication Journal, 54,* 159–170.

Johnson, B. T. (1989). *DSTAT: Software for the meta-analytic review of research literatures.* Hillsdale, NJ: Erlbaum.

*Juhnke, R., Goldman, M., & Baer, R. (1977). Predispositional frustration responses and the expression of evaluations. *Journal of Social Psychology, 102,* 275–282.

Kanekar, S., Bulsara, R. M., Duarte, N. T., & Kolsawalla, M. B. (1981). Perception of an aggressor and his victim as a function of friendship and retaliation. *Journal of Social Psychology, 113,* 241–246.

Knott, P. D., & Drost, B. A. (1970). Sex-role identification, interpersonal agression, and anger. *Psychological Reports, 27,* 154.

*Koerner, F. E. (1977). The effects of depression and sex on aggressive affect and behavior toward the self and toward others. *Dissertation Abstracts International, 38,* 1887B–1888B.

*Konecni, V. J. (1975). The mediation of aggressive behavior: Arousal level versus anger and cognitive labeling. *Journal of Personality and Social Psychology, 32,* 706–712.

*Konecni, V. J., & Ebbesen, E. B. (1976). Disinhibition versus the cathartic effect: Artifact and substance. *Journal of Personality and Social Psychology, 34,* 352–365.

Lagerspetz, K. M. J., Bjorkqvist, K., & Peltonen, T. (1988). Is indirect aggression typical of females? Gender differences in aggressiveness in 11- to 12-year-old children. *Aggressive Behavior, 14,* 403–414.

Lahey, M. A., Downey, R. G., & Saal, F. E. (1983) Intraclass correlations: There's more there than meets the eye. *Psychological Bulletin, 93,* 586–595.

*Lando, H. A., & Donnerstein, E. I. (1978). The effects of a model's success or failure on subsequent aggressive behavior. *Journal of Research in Personality, 12,* 225–234.

*Lando, H. A., Johnson-Payne, E., Gilbert, L. A., & Deutsch, C. J. (1977). Sex differences in response to physical and nonphysical instigators. *Sex Roles, 3,* 523–535.

*Landy, D., & Mettee, D. (1969). Evaluation of an aggressor as a func-

tion of exposure to cartoon humor. *Journal of Personality and Social Psychology, 12,* 66–71.

*Larsen, K. S., Coleman, D., Forbes, J., & Johnson, R. (1972). Is the subject's personality or the experimental situation a better predictor of a subject's willingness to administer shock to a victim? *Journal of Personality and Social Psychology, 22,* 287–295.

*Leventhal, D. B., & Shemberg, K. M. (1969). Sex role adjustment and nonsanctioned aggression. *Journal of Experimental Research in Personality, 3,* 283–286.

*Leventhal, D. B., Shemberg, K. M., & Van Schoelandt, S. K. (1968). Effects of sex-role adjustment upon the expression of aggression. *Journal of Personality and Social Psychology, 8,* 393–396.

*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.

Lohr, J. M., Hamberger, L. K., & Bonge, D. (1988). The relationship of factorially validated measures of anger-proneness and irrational beliefs. *Motivation and Emotion, 12,* 171–183.

Macaulay, J. (1985). Adding gender to aggression research: Incremental or revolutionary change? In V. E. O'Leary, R. K. Unger, & B. S. Wallston (Eds.), *Women, gender, and social psychology* (pp. 191–224). Hillsdale, NJ: Erlbaum.

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

Malone, J., Tyree, A., & O'Leary, K. D. (1989). Generalization and containment: Different effects of past aggression for husbands and wives. *Journal of Marriage and the Family, 51,* 687–697.

Marini, M. M., & Greenberger, E. (1978). Sex differences in educational aspirations and expectations. *American Educational Research Journal, 15,* 67–79.

Miller, N., & Carlson, M. (1990). Valid theory-testing meta-analyses further question the negative state relief model of helping. *Psychological Bulletin, 107,* 215–225.

Miller, N., Lee, J., & Carlson, M. (1991). The validity of inferential judgments when used in theory-testing meta-analysis. *Personality and Social Psychology Bulletin, 17,* 335–343.

Miller, N., & Pollock, V. E. (1994). Meta-analytic synthesis for theory development. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 457–483). New York: Russell Sage Foundation.

Miller, N., & Pollock, V. (in press). Use of meta-analysis for testing theory. *Evaluation and the Health Professions.*

*Mueller, C., & Donnerstein, E. I. (1977). The effects of humor-induced arousal upon aggressive behavior. *Journal of Research in Personality, 11,* 73–82.

Mullen, B., Salas, E., & Miller, N. (1991). Using meta-analysis to test theoretical hypotheses in social psychology. *Personality of Social Psychology Bulletin, 17,* 259–264.

O'Leary, K. D., Barling, J., Arias, I., Rosenbaum, A., Malone, J., & Tyree, A. (1989). Prevalence and stability of physical aggression between spouses: A longitudinal analysis. *Journal of Consulting and Clinical Psychology, 57,* 263–268.

Perry, D. F., & Bussey, K. (1979). The social learning theory of sex differences: Imitation is alive and well. *Journal of Personality and Social Psychology, 37,* 1699–1712.

PsycLIT [Database]. (1974–1994, April). Washington, DC: American Psychological Association [Producer]. Available from: SilverPlatter.

*Pytkowicz, A. R., Wagner, M. H., & Sarason, I. G. (1967). An experimental study of the reduction of hostility through fantasy. *Journal of Personality and Social Psychology, 5,* 295–305.

*Raden, D. (1974). The relationships between prejudice and overt hostile behavior. *Dissertation Abstracts International, 34,* 6747B-6748B.

Reinisch, J. M., & Sanders, S. A. (1986). A test of sex differences in aggressive response to hypothetical conflict situations. *Journal of Personality and Social Psychology, 50,* 1045–1049.

*Richardson, D. R., Hammock, G. S., Smith, S. M., Gardner, W., & Signo, M. (1994). Empathy as a cognitive inhibitor of interpersonal aggression. *Aggressive Behavior, 20,* 275–289.

Roberts, T. (1991). Gender and the influence of evaluations of self-assessments in achievement settings. *Psychological Bulletin, 109,* 297–308.

*Rogers, R. W. (1980). Expressions of aggression: Aggression-inhibiting effect of anonymity to authority and threatened retaliation. *Personality and Social Psychology Bulletin, 6,* 315–320.

*Rohsenow, D. J., & Bachorowski, J. (1984). Effects of alcohol and expectancies on verbal aggression in men and women. *Journal of Abnormal Psychology, 93,* 418–432.

Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills, CA: Sage.

*Rothaus, P., & Worchel, P. (1960). Ego support, communication, catharsis, and hostility. *Journal of Personality, 32,* 296–312.

*Scharff, W. H., & Schlottmann, R. S. (1973). The effects of verbal reports of violence on aggression. *Journal of Psychology, 84,* 283–290.

*Schuck, S. Z., Schuck, A., Hallam, E., Mancini, F., & Wells, R. (1971). Sex differences in aggressive behavior subsequent to listening to a radio broadcast of violence. *Psychological Reports, 28,* 931–936.

*Shemberg, K. M., Leventhal, D. B., & Allman, L. (1968). Aggression machine performance and rated aggression. *Journal of Experimental Research in Personality, 3,* 117–119.

Shields, S. A. (1984). Reports of bodily changes in anxiety, sadness, and anger. *Motivation and Emotion, 8,* 1–21.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.

Siegel, S. (1956). *Nonparametric statistics.* New York: McGraw-Hill.

Siegman, A. W., & Dintzer, L. (1977). The catharsis of aggression and hostility. *Psychological Reports, 41,* 399–410.

*Stagner, R., & Gongdon, C. S. (1975). Another failure to demonstrate displacement of aggression. *Journal of Abnormal and Social Psychology, 51,* 695–696.

Stets, J. E., & Straus, M. A. (1989). The marriage license as a hitting license: A comparison of assaults in dating, cohabiting, and married couples. *Journal of Family Violence, 41,* 33–52.

Sugarman, D. B., & Hotaling, G. T. (1989). Violent men in intimate relationships: An analysis of risk markers. *Journal of Applied Social Psychology, 19,* 1034–1048.

*Taylor, S. P., & Epstein, S. (1967). Aggression as a function of the interaction of the sex of the aggressor and the sex of the victim. *Journal of Personality, 35,* 474–486.

Terman, L. M., & Tyler, L. E. (1954). Psychological sex differences. In L. Carmichael (Ed.), *Manual of child psychology* (2nd ed., pp. 1064–1114). New York: Wiley.

*Titley, R. W., & Viney, W. (1969). Expression of aggression toward the physically handicapped. *Perceptual & Motor Skills, 29,* 51–56.

*Turner, C. W., Layton, J. F., & Simons, L. S. (1975). Naturalistic studies of aggressive behavior: Aggressive stimuli, victim visibility, and horn honking. *Journal of Personality and Social Psychology, 31,* 1098–1107.

*Unger, R. K., Raymond, B. J., & Levine, S. M. (1974). Are women a "minority" group? Sometimes! *International Journal of Group Tensions, 4,* 71–81.

Van Goozen, S. H. M., Frijda, N. H., Kindt, M., & Poll, N. E. van de. (1994). Anger proneness in women: Development and validation of the anger situation questionnaire. *Aggressive Behavior, 20,* 79–100.

White, J. W. (1983). Sex and gender issues in aggression research. In R. G. Green & I. Donnerstein (Eds.), *Aggression: Theoretical and empirical reviews* (Vol. 2, pp. 1–26). New York: Academic Press.

*Wilson, L., & Rogers, R. W. (1975). The fire this time: Effects of race of target, insult, and potential retaliation on Black aggression. *Journal of Personality and Social Psychology, 32,* 857–864.

Wood, J. V., Saltzberg, J. A., & Goldsamt, L. A. (1990). Does affect induce self-focused attention? *Journal of Personality and Social Psychology, 58,* 899–908.

*Worchel, P. (1966). Displacement and the summation of frustration. *Journal of Experimental Research in Personality, 1,* 256–261.

*Yinon, Y. M., Jaffe, Y. M., & Feshbach, S. (1975). Risky aggression in individuals and groups. *Journal of Personality and Social Psychology, 31,* 808–815.

*Yousseff, Z. I. (1968). The role of race, sex, hostility, and verbal stimulus in inflicting punishment. *Psychonomic Science, 12,* 285–286.

Zillman, D. (1979). *Hostility and aggression.* Hillsdale, NJ: Erlbaum.

*Zillman, D., Baron, R. A., & Tambourini, R. (1981). Social costs of smoking: Effect of tobacco smoke on hostile behavior. *Journal of Applied Social Psychology, 11,* 548–561.

Received July 18, 1994
Revision received July 21, 1995
Accepted July 31, 1995 ■