

Stereotypes of Aggression and their Influence on Performance Ratings

Jason D. Way
ACT, Inc.

Two studies sought to examine effects of gender and aggression on job performance ratings. Specifically, it was thought that negative performance episodes, such as aggressive behavior, might have less of an effect on performance ratings for males compared to females because males have a stereotype of being more aggressive. This hypothesis was tested in two studies: one with college undergraduate students and one with working adults. Results in both indicated that gender does not make a difference in performance ratings of aggressive behavior at work, suggesting that there may be less potential for discrimination in the evaluation of aggressive behavior.

INTRODUCTION

Stereotypes are a powerful influence on attitudes and behavior, and have a long history of study in the social psychological literature. The consensus on stereotypes is that they are the cognitive influence on the development of one's attitudes, as compared to prejudices, which are more affective in nature (Devine, 1989; Fiske, 1998). Given the long history of research on stereotypes, it is unsurprising that researchers whose focus is on employees and organizations have begun to pay more attention to the effects that stereotypes can have in a workplace context. Some specific stereotypes that have been examined in a workplace setting are those related to age (e.g. Henkens, 2005; Posthuma & Campion, 2009), ethnicity (e.g. Cocchiara & Quick, 2004), and gender (e.g. Gorman, 2005; Heilman & Chen, 2005).

Researchers examining gender stereotypes in the workplace have found profound effects on employee outcomes. Most commonly, this results in a disadvantage for female employees (Heilman, 2012). Specifically, female employees who exhibited helping behavior outside of their typical job duties received lower performance evaluations than male employees who exhibited the same behavior (Heilman & Chen, 2005). The authors concluded that the female employees were acting according to their gender norms and prescriptions by being helpful, and thus their performance evaluations did not benefit from their extra-role behavior. Because the male gender norm does not include helpfulness, the helpful male employees were seen as truly going beyond their expected behavior, and thus received higher performance evaluations. However, stereotype effects are not exclusive to female employees. Heilman and her fellow researchers have repeatedly found that both men (Heilman & Wallen, 2010) and women (Heilman & Okimoto, 2007; Heilman, Wallen, Fuchs, & Tamkins, 2004) suffer negative consequences as a result of their success on tasks inconsistent with the norms of their gender.

The two studies here examine the relationship between gender stereotypes of aggression. Specifically, this study will involve evaluations of behaviors performed by hypothetical employees. Performance evaluations can be viewed as person perception in a specific context (i.e. at work). Thus, because

stereotypes are an important aspect of attitude formation (Fiske, 1998), gender stereotypes should influence evaluations of employee performance.

Gender and Aggression

Aggressiveness is a common trait associated with masculinity. It shows up as a primary component of the male stereotype in western (particularly American) societies (Frodi, Macaulay, & Thome, 1977; Kaukiainen et al., 2001; Oswald & Lindstedt 2006; Williams & Bennett, 1975), in addition to non-western societies and primitive societies (Gilmore, 1990). Unlike most demographic-based stereotypes, there is an abundant research literature that shows that males are, in general, more aggressive than females (e.g. Eagly & Steffen, 1986; Lindeman, Harakka, & Keltikangas-Jarvinen, 1997). This is especially true for instances of physical aggression and scenarios that require aggressive behavior. This literature reports episodes of objective observations of aggressive behavior, and not subjective perceptions or attitudes regarding the aggressiveness of males. In the case of the latter, the consensus seems to be that male aggression is typically judged to be more aggressive than female aggression (e.g. Basow, Cahill, Phelan, Longshore, & McGillicuddy-DeLisi, 2007). However, at least one study using vignettes found that participants assigned equal levels of aggressiveness to males and females who performed the same aggressive acts (Stewart-Williams, 2002).

The context of the aggressive behavior must be a factor when comparing levels of aggression between men and women (Spector, 2012). In the context of intimate relationships, instances of initiating intimate partner violence (IPV; violence within intimate relationships) were much more common among women than among men, at least in a sample of newlyweds over the first 30 months of marriage (O'Leary et al., 1989). Additionally, instances where the male was the aggressor and the female was the target of the aggression were judged more harshly, considered more serious, and deemed more justifiable for reporting the act to the police than instances where the female was the aggressor and the male was the target, especially when the two were described in the vignette as being married (Felson & Feld, 2009).

The tendency for males to be more aggressive than females has also been reported in the organizational psychology literature in the context of workplace aggression (e.g. Baron et al., 1999; Hershcovis et al., 2007; McFarlin, Fals-Stewart, Major, & Justice, 2001). One meta-analysis concerning workplace aggression found that gender strongly predicted aggression directed at others as compared to aggression directed at the organization (Hershcovis et al., 2007). However, both types of aggression were related to gender when the researchers entered the meta-analytic correlations into a path analysis model. It should be noted that a low base rate in the reporting of physical aggression could account for the lack of gender difference in relation to this form of aggression.

STUDY 1

The goal of this paper is to incorporate research on gender stereotypes into the literature regarding aggression by examining whether the cultural stereotype of male aggressiveness will influence the perceptions of and reactions to interpersonal aggressive behaviors performed by men. Although physically aggressive behavior is less common in the workplace than other forms of aggression, the male stereotype and idea of masculinity is more often associated with physical forms of aggression (see discussion above). Thus, these aggressive behaviors will be the focus of this study. While these behaviors are things that should be avoided, it could be the case that men who engage in such behaviors are simply following the stereotype of what men typically do, i.e. behaving aggressively. Men's ratings of performance may therefore suffer less for exhibiting such behaviors compared to women who engage in such behaviors. Women would not be expected to behave in such ways toward other employees, and in fact, women who do may even be violating norm prescriptions that they are supposed to be friendly and helpful toward others. Such violations are typically penalized (Cialdini & Trost, 1998; Heilman & Chen, 2005), giving further weight to the idea that there will be a differential reaction to physical aggression depending on whether the behavior is being performed by a man or a woman, in line with previous

findings showing that identical aggressive acts are judged differently depending on the gender of the actor (Felson & Feld, 2009).

Hypothesis

The following hypothesis is proposed:

H1: Engaging in physically aggressive behavior will result in lower ratings on performance evaluations and reward recommendations for women compared to men.

In order to determine whether any effects on the dependent variables are specifically due to physical aggression or a general reaction to poor workplace behavior, a condition will be included where the employee to be rated leaves work abruptly. Aggression and leaving work both qualify as instances of counterproductive work behavior (CWB; Rotundo & Sackett, 2002; Spector et al., 2006), so when these two conditions are collectively referred to in the presentation of Study 1, they will be referred to as such.

METHOD

Participants

Participants were recruited from an undergraduate introductory to psychology course at a large southeastern university. The students were offered a point of extra credit added to their overall course grade in exchange for participation in the study. Power analyses based on Cohen's set correlation were used to estimate the number of participants needed for a significant result. These analyses used the f^2 statistic for multivariate effect size (Cohen, 1988). In a study examining similar variables using similar methods, the results all exhibited at least a medium effect size, and some much higher (Heilman & Chen, 2005). Assuming a medium multivariate effect size in this study ($f^2 = 0.15$; Cohen, 1988, p. 478), the power analyses showed that even with 60 total participants, the power is still 0.88, above the minimum acceptable value of 0.8. Using a more conservative estimate of multivariate effect size ($f^2 = 0.07$), the power analyses showed that with 120 total participants, power would be at 0.87.

The initial sample was 140 students, but after accounting for missing data in some of the surveys, the final sample was 134 students (52 male, 82 female). The average age was 19.57 years ($SD = 1.65$), and the average amount of work experience was 2.87 years ($SD = 2.22$). Forty-eight point nine percent of the participants classified themselves as Caucasian, 6.7 percent African American, 29.6 percent Hispanic or Latino, 11.1 percent as Asian, and 3.7 percent as Other.

Procedure

Each participant was given a packet containing the research materials for the study. The packet contained the information sheet explaining the purpose of the study, supervisor comments about the employee's performance, a warning letter about the employee repeatedly being tardy for shifts, and finally the study questionnaires asking participants to evaluate the employee on job performance, make recommendations for organizational rewards, give attribute ratings for the employee, rate their gender ideal match, and finally report some demographic information about themselves. Employees in the leaving and aggression conditions received additional information about the employee, described below.

Independent Variable Manipulation

Gender of Target Employee

The gender of the employee to be rated were varied by the name and personal pronouns in the performance episodes and employee information sheet.

CWB

All conditions included basic information about the employee, supervisor comments about performance, and the warning letter about the employee's tardiness. In the control condition, this was all the information the participants received. In the other two conditions, there was also an employee incident report describing how the employee was faced with some critical comments about their state of

dress from a coworker, and the employee's reaction to those comments. In the aggression condition, the employee shoved the coworker against a wall and stormed off into the kitchen. In the leaving condition, the employee abruptly left work without telling their supervisor and did not return that evening.

Manipulation Checks

Three questions serve as manipulation checks for the study. They were "Has James ever behaved aggressively at work?", "Has James ever arrived late for work?", and "Has James ever left work in the middle of a shift?". Participants responded either "Yes", "No", or "Don't Know". They were instructed to select the third option if no information about that behavior was presented. The purpose of these was to make sure the participants were paying attention to the various types of information about the employee presented in the study.

Dependent Variables

Performance Evaluation

Performance was measured with the three items used by Heilman and Chen (2005). These items are: "Overall, how would you rate this employee's performance over the past year?", "In your opinion, how likely is it that this employee will advance in the company?", and "Give your assessment of the individual's likelihood of success." Each item was measured on a 7 point Likert scale with 1 indicating either poor (first item) or very unlikely, and 7 indicating either excellent (first item) or very likely. The reliability for this scale was $\alpha = .84$.

Reward Recommendations

This variable was also measured with items used by Heilman and Chen (2005). Participants were asked to give their recommendations for three types of common organizational rewards (salary increase, promotion, and bonus pay). They were assessed on a 7 point Likert scale ranging from 1 (*would definitely not recommend*) to 7 (*would definitely recommend*). The reliability for this scale was $\alpha = .86$.

RESULTS AND DISCUSSION

Manipulation Checks

The three questions that served as manipulation checks were intended to determine if the participants noted the instances of behavior that were presented in each condition. Across all conditions, the employee was described as having a history of arriving late to work. Accordingly, 131 out of the 134 participants correctly reported that the employee had previously arrived late to work. In the male and female leave conditions, the employee was described as leaving work abruptly in the middle of the shift, and 38 out of 45 participants (male and female conditions combined) correctly reported that the employee had done so. Finally, in the male and female aggress conditions, the employee was described as acting aggressively towards a coworker, and 39 out of 45 participants (male and female conditions combined) correctly reported that the employee had done so. These results raise some concerns regarding the quality of responses from those participants who did not answer the manipulation checks correctly. In order to examine this issue, all of the analyses described below were repeated after removing any participant who failed to respond correctly to any of the three manipulation check questions. The pattern of results was exactly the same for all of the statistical tests in the study, so the full sample was maintained in the analyses described below.

Demographic Effects

Because this study is concerned with gender effects, preliminary analyses were conducted with participant gender as an additional independent variable in order to determine if there are differences between men and women in their reactions to the vignettes. Some previous research has shown that men were more likely to attribute other men's physical aggression to situational factors (e.g. Weaver et al., 2010), so it is important to determine if participant gender should be included as an additional factor in

this study. In three way ANOVAs that included participant gender as a third variable (along with employee gender and performance condition), participant gender had no main effect or interaction effects on either the performance evaluation or reward recommendation variables. Additionally, all of the demographic variables were entered into a correlation matrix with all of the study variables (performance evaluation, reward recommendation, competence, interpersonal incivility, gender ideal, perceived aggressiveness). The only significant correlation in the matrix was that gender was weakly related to competence ($r = -.18, p < .05$), such females gave slightly lower ratings of competence than males. To further examine any possible participant gender effects, *t*-tests were conducted on all of the study variables with participant gender as the independent variable. Similar to the results of the correlational analysis, competence was the only variable to show an effect ($t(132) = 2.1, p < .05$), in the same direction as described above. Given these minor results, none of the demographic variables were entered as control variables or additional independent variables in the following analyses.

Tests of Study Hypotheses

Performance Rating Variables

Means, standard deviations, and correlations among the study variables are presented in Table 1.

TABLE 1
MEANS, STANDARD DEVIATIONS, AND CORRELATIONS AMONG DEPENDENT
VARIABLES IN STUDY 1

Variable	<i>M</i>	<i>SD</i>	1	2
1. Performance Evaluation	3.95	1.14	(.84)	
2. Reward Recommendation	3.25	1.26	.75*	(.86)

Note. * $p < .05$, ** $p < .01$. Values along the diagonal are reliability scores for that scale.

To examine Hypothesis 1, a multivariate analysis of variance was conducted to test for effects of employee gender and CWB performance on the combination of the two dependent variables (performance evaluations and reward recommendations). The Wilks' Lambda test statistic was not significant for either employee gender ($\Lambda = 0.99, F(1, 128) = 0.53, p > .05$), or the employee gender x CWB performance interaction ($\Lambda = 0.99, F(4, 256) = 0.36, p > .05$). However, the test statistic was significant for CWB performance ($\Lambda = 0.86, F(4, 256) = 5.01, p < .05$). Because employee gender had neither a main or interactive effect on the two dependent variables, Hypothesis 1 was not supported.

In spite of the lack of support for the first hypothesis, additional univariate ANOVAs were conducted on each dependent variable separately in order to determine the nature of the multivariate main effect of performance. A 2 x 3 ANOVA on the performance evaluation variable resulted in an identical pattern of results, such that employee gender did not have a significant main effect ($F(1, 129) = 1.04, p > .05$) or interaction effect ($F(2, 129) = 0.53, p > .05$). However, CWB performance still had a significant main effect ($F(2, 129) = 8.00, p < .05$). Post-hoc Tukey tests for the performance conditions revealed that the control groups had higher ratings of performance evaluation than either of the two CWB performance groups (leave, aggress; see Table 2). A 2 x 3 ANOVA on the reward recommendation variable also exhibited an identical pattern of results, such that employee gender did not have a significant main effect ($F(1, 129) = 0.43, p > .05$) or interaction effect ($F(2, 129) = 0.53, p > .05$). However, CWB performance again had a significant main effect ($F(2, 129) = 9.53, p < .05$). Post-hoc Tukey tests for the performance conditions revealed that the control groups had higher ratings of reward recommendations than either of the two CWB performance groups (leave, aggress; see Table 2).

TABLE 2
MEANS AND STANDARD DEVIATIONS OF STUDY VARIABLES BY CONDITION IN
STUDY 1

Condition		Performance Evaluation			Reward Recommendation	
Employee Gender	CWB Performance	<i>n</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Male	Control	23	4.35 _a	1.11	3.72 _a	1.18
	Leave	24	3.53 _b	0.55	2.78 _b	0.77
	Aggress	22	3.71 _b	0.90	3.06 _b	1.12
Female	Control	22	4.62 _a	1.17	4.05 _a	1.36
	Leave	21	3.90 _b	1.27	3.02 _b	1.42
	Aggress	23	3.63 _b	1.36	2.90 _b	1.23

Note. Higher means indicate higher scores on each of the study variables. Ratings were done on 7-point Likert scales. Means within a column with different subscripts differ significantly at $p < .05$.

The gender of the employee did not have any effect on the performance evaluation or reward recommendation ratings, either by itself or through an interaction with the level of CWB. There was a main effect for the CWB conditions, such that the employees in the two CWB conditions received lower ratings on the performance variables than the employees in the control condition, although the two CWB conditions were not significantly different from each other. This indicates that the participants in these conditions were taking the additional information into account when making their performance ratings, and that the manipulations between conditions did have some effect on the performance ratings, such that employees who performed more CWB were rated lower on the performance variables.

However, the main manipulation that was hypothesized to have an effect but did not was employee gender. There are several reasons why this could have happened in this study. Although stereotypes have been found to have an effect in studies of task (e.g. Heilman & Wallen, 2010) and citizenship (e.g. Heilman & Chen, 2005) performance, it could be the case that stereotypes do not have an effect where negative work behaviors are concerned, such that employees are penalized equally regardless of their gender. However, this is unlikely, as Felson and Feld (2009) found significant gender differences in regard to how participants interpreted aggressive behavior coming from men compared to that coming from women, showing that stereotypes can indeed influence perceptions of negative behavior. It is more likely that the aggressive manipulation was not strong enough to produce the effect. In the descriptions of the incident in the aggress condition, the employee merely shoved the coworker who made the disparaging remark. However, stereotypes regarding men and physical aggression are more often associated with acts such as hitting or punching, and studies in which physical aggression was examined as a tool to regain perceived manhood use these sorts of stronger actions rather than shoving (e.g. Bosson et al., 2009). Thus, the hypothesized gender effects might require stronger aggressive actions in order to manifest than the actions described in this study. In addition, it would be interesting to expand these hypotheses to relational aggression, which is typically affiliated with femininity (see below).

STUDY 2

Several alterations were made to the design of Study 1 in an effort to conduct a more thorough test of its hypotheses. First, a larger sample was recruited, and the sample consisted of working adults rather than college students. Having more experience with the process of performance ratings could help draw out the hypothesized effects. Second, the leave work condition and rating recommendation dependent variable were removed, as they were extraneous to the main purpose of the study. There were no differences in results between the aggression and leave conditions in Study 1, nor between the two dependent variables (which were in fact highly correlated, $r = .75, p < .05$), suggesting that these additional aspects of the study design were redundant. Third, gender was specified for the target of the aggressive behavior and for the participant (despite no difference between males and females in Study 1) in case these would have significant effects in the larger sample. Some research suggests that reactions to aggressive behavior vary based on the gender of the target, with aggression against males being seen as less aggressive and more acceptable (e.g., Basow et al., 2007; Harris & Knight-Bahnhof, 1996), which could have impacts on how the aggressive behavior is rated. Fourth, the nature of the physically aggressive act was strengthened by presenting it as a “punch” rather than a “shove”, which would in theory make the act more salient for the participants. Finally, relational aggression was included as an additional aggression condition.

Relational aggression is one form of aggression that is more often associated with females than males, in part because females are stereotyped to be more focused on relationships with others than males, as well as the fact that they actually put more effort into developing and sustaining relationships (Golombok & Hines, 2002; Underwood, 2002). Relational aggression is characterized by a desire to harm relationships with others through behaviors such as threatening social exclusion to elicit a desired behavior, spreading rumors to elicit peer dislike, or outright excluding others from group activities (Crick, Casas, & Mosher, 1997). Some researchers have categorized relational aggression as indirect because it often does not involve a direct confrontation (e.g., Crick & Grotpeter, 1995, 1996). However, others (e.g., Archer, 2004; Spector, 2012) have pointed out that direct confrontation is often the means by which relationally aggressive behavior is performed, and the goal of such behavior is to cause direct harm to an individual, and relational aggression should therefore be categorized as direct aggression.

Relational aggression has been most commonly studied in the developmental literature, where studies have reported that young girls seem to exhibit more relational aggression than young boys (e.g., Archer, 2004; Lagerspetz, Bjorkqvist, & Peltonen, 1988). However, this finding has held true for older females as well (e.g., Archer & Coyne, 2005). Few studies have examined such behaviors in adults, but Spector (2012) reports one sample that showed male employees report performing more relational aggression-related CWB than female employees. However, he notes that social desirability may be skewing the reporting of the frequency of the behavior, and the work context may have further discouraged female employees from accurately reporting their behavior.

In addition to the stereotypes of females being more relationally oriented than males mentioned above, cultural stereotypes often associate females with relational aggression (Stewart-Williams, 2002). Further, a survey of college females showed that they see relational aggression as normal behavior for women (Miller-Ott & Kelly, 2013).

Hypotheses

The following hypotheses are proposed:

H1a: Engaging in physically aggressive CWB will result in lower performance ratings for women compared to men, because physical aggression is in line with a stereotype of men’s behavior.

H1b: Engaging in relationally aggressive CWB will result in lower performance ratings for men compared to women, because relational aggression is in line with a stereotype of women’s behavior.

Because these hypotheses are based on stereotypes of men and women’s typical or appropriate behavior, which should influence the rating given by the participant regardless of the participant’s gender or the employee who is not being rated (i.e., the target of the aggression), there are no hypothesized effects of participant gender and target gender on performance ratings.

METHOD

Participants

Participants were recruited from an online pool of adults through Amazon's Mechanical Turk (MTurk) tool. MTurk has proven to be an effective participant recruitment tool that results in more diverse samples than typical internet or college samples, as well as data with comparable reliability to those obtained with more traditional methods (e.g., Buhrmester, Kwang, & Gosling, 2011). The participants were offered \$0.50 in exchange for participation in the study. The MTurk task contained a link to an external online survey hosted by Qualtrics, which contained the study vignettes and scales described below. A power analysis prior to the data collection determined that at a small effect size ($f^2 = .02$), 480 participants would provide sufficient power (.86) for this study. This would have allowed for 30 participants per condition (16 conditions). In multivariate designs, a group of 10 to 20 is recommended in order to allow for the normal approximation of moderately non-normal distributions (Stevens, 2009, p.221). The smaller effect size was chosen in response to the null results in Study 1, in hope that the more conservative estimate would provide a sample size that would provide a more thorough test of the study hypotheses.

The initial sample collected was 649 participants. The oversampling was due to trying to get at least 30 usable participant responses per condition, which depended on getting enough male and female participants into the right conditions and those participants passing the manipulation check questions.

After removing participants who did not pass the manipulation checks from the sample, the final number of participants was 552. The gender of the participants was almost perfectly balanced, with 277 males (50.2%) and 275 females (49.8%). The average age of the sample was 36.1 years ($SD = 12.4$), although two participants chose not to respond to the age question. The ethnic makeup of the sample was as follows: 47 Asian/Pacific Islander participants (8.5%), 31 Black/African American participants (5.6%), 27 Hispanic/Latino participants (4.9%), 4 Native American participants (0.7%), 427 White/Caucasian participants (77.4%), 15 Other/More than one applies participants (2.7%). One participant chose not to respond to the ethnicity question. Participants came from a fairly even distribution in terms of regions of the United States: 138 participants (25.0%) were from the Northeast, 186 participants (33.7%) were from the South, 109 participants (19.7%) were from the Midwest, and 116 participants (21.0%) were from the West. The participants had an average of 15.6 years ($SD = 11.6$) of work experience. Of those participants that had supervisory experience ($n = 353$), the average amount of supervisory experience was 5.6 years ($SD = 5.7$).

Procedure

The online survey contained the same information as the survey in Study 1. The survey randomly assigned participants to one of eight versions of the vignettes, which varied by aggression type, gender of the aggressor, and aggressor of the target (2x2x2). All participants then saw the same set of items after the vignettes, with employee gender adjusted as appropriate. Following the data collection, the responses for the manipulation checks and the feasibility checks were screened as described below, which were followed by the main analyses of the study.

Independent Variable Manipulations

Gender of Participant

This was a subject variable.

Gender of Aggressor

Same as Study 1.

Gender of Target

The gender of the target of the aggression was varied by the name (Mike or Amanda) and personal pronouns in the performance vignettes and employee information sheet.

Aggression Type

There were two aggression conditions, a physical aggression condition and a relational aggression condition. Both conditions included basic information about the aggressor employee, supervisor comments about his or her performance, and an employee incident report describing an interpersonal conflict episode. The conflict episode consisted of the target employee mistakenly taking a meal intended for the aggressor employee's table while both employees were in the kitchen picking up meals during the dinner rush. An argument ensues about to whom the meal belongs, which escalates into the employees yelling at each other, and culminates in the aggressive behavior being performed by the aggressor employee. In the physical aggression condition, the aggressor employee punched the target employee in the arm after the argument. In the relational aggression condition, the aggressor employee left the kitchen and spread a rumor about the target employee.

Manipulation Checks

Three questions served as manipulation checks for the aggression conditions: "Was there an argument in the incident report?", "Did the employee hit someone?", and "Did the employee spread rumors about someone?". Participants responded either "Yes", "No", or "Don't Know". The purpose of these questions was to make sure the participants were paying attention to the various types of information presented in the study about the employee whom they rated. Participants in all conditions should have responded "Yes" to the first question. Participants in the physical aggression conditions should have responded "Yes" to the second question and "No" to the third question. Participants in the relational aggression conditions should have responded "No" to the second question and "Yes" to the third question. Participants who did not respond appropriately to these questions (given their randomly assigned condition) or who selected "Don't Know" for any of the questions were removed from the sample.

Feasibility checks

Two questions served to ensure that the interactions described in the vignettes were possible given the circumstances of the situations described in the vignettes: "Are these employee interactions possible in this context?" and "Is it conceivable that someone in this situation could act this way?". The response scales for both items were "Yes", "No", or "Don't Know". Participants in the final sample thought that the situations and conflicts described in the vignettes were feasible for the most part, with 512 participants (92.8%) responding "Yes" to the first question, and 509 participants (92.2%) responding "Yes" to the second question.

Dependent Variable

Performance Rating

Same items as Study 1. The reliability for this scale was $\alpha = .84$.

Other Dependent Variables

Because this study was part of a larger research effort, there were other dependent variables included that were not relevant to the hypotheses here. As such, the results below reflect a multivariate approach, but only results relevant to this study's hypotheses will be presented.

RESULTS AND DISCUSSION

Descriptive Statistics

Means, standard deviations, and intercorrelations for participant gender, age, work experience, supervisory experience, and the dependent variables can be found in Table 3. Cell sizes, means, and standard deviations of the dependent variable listed by condition can be found in Table 4. Participant responses on the dependent variable items were summed to form scale scores.

TABLE 3
DEMOGRAPHIC AND DEPENDENT VARIABLE MEANS, STANDARD DEVIATIONS, AND
INTERCORRELATIONS FOR STUDY 2

Variable	N	Mean	SD	1	2	3	4	5
1. Gender	552	--	--	--				
2. Age (years)	550	36.1	12.4	.09*	--			
3. Work Experience (years)	551	15.6	11.6	.07	.91**	--		
4. Supervisory Experience (years)	550	3.6	5.3	-.07	.49**	.55**	--	
5. Performance Rating	552	9.5	3.2	-.03	.02	.06	.02	(.84)

Note. * $p < .05$. ** $p < .01$. Values in parentheses on the diagonal are alpha coefficients of the dependent variable scales.

TABLE 4
CELL MEANS AND STANDARD DEVIATIONS OF PERFORMANCE IN STUDY 2

Cell	Participant Gender	Aggression Type	Aggressor Gender	Target Gender	N	Perf	
						Mean	(SD)
1	F	P	F	F	43	8.5	(3.3)
2	F	P	F	M	33	10.5	(3.4)
3	F	P	M	F	35	8.9	(3.6)
4	F	P	M	M	30	9.7	(3.4)
5	F	R	F	F	35	9.2	(2.6)
6	F	R	F	M	33	9.4	(2.7)
7	F	R	M	F	36	9.7	(3.2)
8	F	R	M	M	30	9.8	(2.9)
9	M	P	F	F	34	9.7	(3.4)
10	M	P	F	M	33	10.1	(3.6)
11	M	P	M	F	31	7.6	(3.0)
12	M	P	M	M	37	9.7	(3.8)
13	M	R	F	F	31	9.7	(2.6)
14	M	R	F	M	30	9.8	(3.0)
15	M	R	M	F	40	9.8	(3.4)
16	M	R	M	M	41	10.0	(3.3)

Note. Total $N = 552$. F=Female, M=Male, P=Physical Aggression, R=Relational Aggression, Perf=Performance Rating, Accep=Acceptability of Aggression, AggAct=Perceived Aggressiveness of Act, AggAgg=Perceived Aggressiveness of Aggressor, Masc=Perceived Masculinity, Fem=Perceived Femininity, Ideal=Perceived Gender Ideal.

MANOVAs and Follow-Up Tests

MANOVA

A multivariate analysis of variance was conducted using Wilk's Lambda test statistic to test the effects of the independent variables of participant gender, aggression type, aggressor gender, target gender, and their interactions on the set of the performance rating and other dependent variables. The MANOVA showed significant main effects for aggression type ($\Lambda = 0.99$, $F(4, 533) = 0.73$, $p < .05$) and target gender ($\Lambda = 0.95$, $F(4, 533) = 6.83$, $p < .05$), and a significant interaction effect for aggression type x aggressor gender ($\Lambda = 0.97$, $F(4, 533) = 3.65$, $p < .05$). All other main effects and interaction effects for this MANOVA were not significant. A full set of test statistics for this analysis can be found in Table 4.

The univariate ANOVA of performance rating ($F(2, 549) = 4.57$, $p < .0125$) was significant. Follow up Tukey tests for performance rating showed no mean difference for aggression type, but target gender had a mean difference such that aggression against male targets resulted in significantly higher performance ratings than aggression against female targets (mean difference = .76). However, the significant main effects need to be examined in the context of the significant interaction, and the hypothesized effects in the study are interaction effects, so the interaction effects will be the main focus of analysis and interpretation.

Follow-Up Test for Significant Interaction Effects

In order to assess the nature of the univariate effect that the aggression type x aggressor gender interaction (the only interaction showing a significant multivariate effect) was having on the dependent variable, the study conditions were dummy coded to the four combinations of the levels of these two variables (i.e., physical aggression from a male, physical aggression from a female, relational aggression from a male, or relational aggression from a female). This dummy-coded variable was then used as the independent variable in a one-way factorial ANOVA on each dependent variable (using the Bonferroni-adjusted alpha level), and follow up Tukey tests identified the conditions with means significantly different from one another. The one-way ANOVA on performance rating ($F(3, 548) = 1.57$, $p > .05$) was not significant.

Study Hypotheses

Because the aggressor type x aggressor gender interaction did not have a significant effect on the performance rating variable, *H1a* and *H1b* were not supported. These results mean that employees of different genders are not rated differently in terms of their performance, regardless of the type of aggression they are exhibiting. Thus, in terms of rating aggressive forms of counterproductive work behaviors, it appears that stereotypes of gender-appropriate forms of aggression are not related to performance ratings of these behaviors, and that male and female employees are rated equally when performing equivalent examples of aggressive behavior.

GENERAL DISCUSSION

These two studies attempted to show that the gender of an employee would influence that employee's performance ratings. It was hypothesized that men would receive better performance ratings than women in instances where the employees were exhibiting physical aggression, and that women would receive better performance ratings than men in instances where the employees were exhibiting relational aggression, because these behaviors were in line with the male and female stereotype, respectively, and thus the employees were less likely to be penalized for exhibiting them. This hypothesis was not supported in Study 1. In Study 2, the hypothesized effect should have resulted in a significant interactive effect of aggressor gender and aggression type on the performance ratings variable, but this effect was not significant, meaning that employees of different genders are not rated differently in terms of their performance, regardless of the type of aggression they are exhibiting. Thus, in terms of rating aggressive forms of counterproductive work behaviors, it appears that stereotypes of gender-appropriate forms of

aggression are not related to performance ratings of these behaviors, and that male and female employees are rated equally when performing equivalent examples of aggressive behavior.

Theoretical Implications

In contrast to previous research on gender stereotypes and their effect on evaluations of positive workplace behavior (e.g., Heilman & Chen, 2005), these stereotypes do not appear to influence performance ratings of aggressive forms of counterproductive work behavior. Although there are ways that these effects could be teased out given some changes to the design of the study (see below), it could be that the negative aspect of aggressive behavior at work is substantial enough to override any effect that stereotypes would have on such performance ratings.

Practical Implications

The lack of significant results in terms of the performance ratings of the aggressive behavior actually have several positive practical implications for performance appraisal situations, as it appears that gender stereotypes of aggression do not influence appraisals of aggressive behavior. Indeed, aggressive behavior is seen as equally undesirable regardless of the gender of the aggressor or the type of aggression. If these results hold true, it would mean there are fewer chances that a manager would discriminate based on gender when rating the performance of employees who have exhibited aggression at work. Also, the lack of significant effects of participant gender on performance rating means that the gender of the manager or the rater would not make a difference or have an effect on the ratings, so that, all else being equal, managers are more likely to rate such behavior the same regardless of their gender.

Limitations

Like any study, this one had several limitations that could have impacted the results. A wider variety of examples of physical and relational aggression, as in the Basow et al. (2007) study, could have elicited a greater range of differences among the primary dependent variables, leading to more significant results for the performance, acceptability, and perceived aggressiveness variables. Portraying the aggressive behavior as being more severe and having more of an impact on the target of the aggression in terms of the consequences of the aggressive behavior could also have resulted in greater significant effects on these variables.

Future Research

Future research should attempt to address some of the limitations of this study in order to have a greater chance of eliciting the hypothesized effects. For example, including more examples of physical and relational aggression as well as portraying more severe consequences of the aggression for the target may get a stronger reaction out of the participants to the aggressive behavior described in the vignettes. Also, the vignettes in this study provided some justification for the aggression in terms of an escalating argument over a mistake made by the target. This may have taken some of the responsibility for the aggressive behavior off the aggressor and placed it onto the target in the minds of the participants, so future studies could include this as an additional variable to determine if it made a difference in the results. Additionally, it would be interesting to conduct a priming study where participants are primed with either manager training videos for rating employees or videos showing stereotyped depictions of physical and relational aggression in order to see if these two conditions resulted in different levels of ratings based on the aggressive behavior.

CONCLUSIONS

The results of this study suggest that gender stereotypes of aggression may not influence performance ratings of the aggressors, however, further research is needed to clarify these findings.

REFERENCES

- Archer, J. (2004). Sex differences in aggression in real-world settings: A meta-analytic review. *Review of General Psychology, 8*, 291-322. <http://dx.doi.org/10.1037/1089-2680.8.4.291>
- Archer, J., & Coyne, S. M. (2005). An integrated review of indirect, relational, and social aggression. *Personality and Social Psychology Review, 9*, 212-230. http://dx.doi.org/10.1207/s15327957pspr0903_2
- Baron, R. A. & Richardson, D. (1994.) *Human aggression*. New York: Plenum. <http://dx.doi.org/10.1007/978-1-4615-7195-7>
- Basow, S. A., Cahill, K. F., Phelan, J. E., Longshore, K., & McGillicuddy-DeLisi, A. (2007). Perceptions of relational and physical aggression among college students: Effects of gender of perpetrator, target, and perceiver. *Psychology of Women Quarterly, 31*, 85-95. <http://dx.doi.org/10.1111/j.1471-6402.2007.00333.x>
- Bosson, J. K., Vandello, J. A., Burnaford, R. M., Weaver, J. R., & Wasti, S. A. (2009). Precarious manhood and displays of physical aggression. *Personality and Social Psychology Bulletin, 35*, 623-634. <http://dx.doi.org/10.1177/014616720831161>
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*, 3-5.
- Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity, and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 151-192). Boston: McGraw-Hill. <http://dx.doi.org/10.4135/9781412956321.n65>
- Cocchiara, F. K. & Quick, J. C. (2004). The negative effects of positive stereotypes: Ethnicity-related stressors and implications on organizational health. *Journal of Organizational Behavior, 25*, 781-785. <http://dx.doi.org/10.1002/job.263>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Crick, N. R., Casas, J. F., & Mosher, M. (1997). Relational and overt aggression in preschool. *Developmental Psychology, 33*, 579-588. <http://dx.doi.org/10.1037/0012-1649.33.4.579>
- Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender and social-psychological adjustment. *Child Development, 66*, 710-722. <http://dx.doi.org/10.2307/1131945>
- Crick, N. R., & Grotpeter, J. K. (1996). Children's treatment by peers: Victims of relational and overt aggression. *Development and Psychopathology, 8*, 367-380.
- Devine, P. G. (1989). Stereotypes and Prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5-18. <http://dx.doi.org/10.1037/0022-3514.56.1.5>
- Eagly, A. H. & Steffen, V. J. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin, 100*, 309-330. <http://dx.doi.org/10.1037/0033-2909.100.3.309>
- Felson, R. B., & Feld, S. L. (2009). When a man hits a woman: Moral evaluations and reporting violence to the police. *Aggressive Behavior, 35*, 477-488. <http://dx.doi.org/10.1002/ab.20323>
- 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. <http://dx.doi.org/10.1037/0033-2909.84.4.634>
- Gilmore, D. D. (1990). *Manhood in the Making*. New Haven: Yale University Press.
- Golombok, S. & Hines, M. (2002). Sex differences in social behavior. In P. K. Smith & C.H. Hart (Eds.), *Blackwell handbook of childhood social development* (pp. 117-136). Malden, MA: Blackwell.
- Gorman, E. H. (2005). Gender stereotypes, same-gender preferences, and organizational variation in the hiring of women: Evidence from law firms. *American Sociological Review, 70*, 702-728. <http://dx.doi.org/10.1177/000312240507000408>
- Harris, M. B. & Knight-Bohnhoff, K. (1996). Gender and aggression II: Personal aggressiveness. *Sex roles, 35*, 27-42. <http://dx.doi.org/10.1007/BF01548173>

- Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32, 113-135. <http://dx.doi.org/10.1016/j.riob.2012.11.003>
- Heilman, M. E. & Chen, J. J. (2005). Same behavior, different consequences: Reactions to men's and women's altruistic citizenship behavior. *Journal of Applied Psychology*, 90, 431-441. <http://dx.doi.org/10.1037/0021-9010.90.3.431>
- Heilman, M. E. & Okimoto, T. G. (2007). Why are women penalized for success at male targeted tasks?: The implied communality deficit. *Journal of Applied Psychology*, 92, 81-92.
- Heilman, M. E. & Wallen, A. S. (2010). Wimpy and undeserving of respect: Penalties for men's gender-inconsistent success. *Journal of Experimental Social Psychology*, 46, 664-667. <http://dx.doi.org/10.1016/j.jesp.2010.01.008>
- Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed at male gender-typed tasks. *Journal of Applied Psychology*, 89, 416-427. <http://dx.doi.org/10.1037/0021-9010.89.3.416>
- Henkens, K. (2005). Stereotyping older workers and retirement: The managers' point of view. *Canadian Journal on Aging*, 24, 353-366. <http://dx.doi.org/10.1353/cja.2006.0011>
- Hershcovis, M. S., Turner, N., Barling, J., Arnold, K. A., Dupre, K. E., Inness, M., LeBlanc, M. M., & Sivanathan, N. (2007). Predicting workplace aggression: A meta-analysis. *Journal of Applied Psychology*, 92, 228-238. <http://dx.doi.org/10.1037/0021-9010.92.1.228>
- Lagerspetz, K. M. J., Björkqvist, K., & Peltonen, T. (1988). Is indirect aggression typical of females? *Aggressive Behavior*, 14, 403-414.
- Lindeman, M., Hara, T., & Keltikangas-Jarvinen, L. (1997). Age and gender differences in adolescents' reactions to conflict situations: Aggression, prosociality, and withdrawal. *Journal of Youth and Adolescence*, 26, 339-351.
- McFarlin, S. K., Fals-Stewart, W., Major, D. A., & Justice, E. M. (2001). Alcohol use and workplace aggression: An examination of perpetration and victimization. *Journal of Substance Abuse*, 13, 303-321. [http://dx.doi.org/10.1016/S0899-3289\(01\)00080-3](http://dx.doi.org/10.1016/S0899-3289(01)00080-3)
- Miller-Ott, A. E. & Kelly L. (2013). Mean Girls in College: An Analysis of How College Women Communicatively Construct and Account for Relational Aggression. *Women's Studies in Communication*, 36, 330-347. <http://dx.doi.org/10.1080/07491409.2013.829792>
- O'Leary, K., 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. <http://dx.doi.org/10.1037/0022-006X.57.2.263>
- Oswald, D. L. & Lindstedt, K. (2006). The content and function of gender self-stereotypes: An exploratory investigation. *Sex Roles*, 54, 447-458. <http://dx.doi.org/10.1007/s11199-006-9026-y>
- Posthuma, R. A. & Campion, M. A. (2009). Age stereotypes in the workplace: Common stereotypes, moderators, and future research directions. *Journal of Management*, 35, 158-188. <http://dx.doi.org/10.1177/0149206308318617>
- Rotundo, M. & Sackett, P. R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy capturing approach. *Journal of Applied Psychology*, 87, 66-80. <http://dx.doi.org/10.1037/0021-9010.87.1.66>
- Spector, P. E. (2012). Gender differences in aggression and counterproductive work behavior. In S. Fox & T. R. Lituchy (Eds.), *Gender and the dysfunctional workplace*. Cheltenham Glos, UK: Edward Elgar. <http://dx.doi.org/10.4337/9780857932600.00009>
- Stevens, J. (2009). *Applied multivariate statistics for the social sciences* (5th Ed.). Mahwah, NJ: Erlbaum.
- Stewart-Williams, S. (2002). Gender, the perception of aggression, and the overestimation of gender bias. *Sex Roles*, 46, 177-189.
- Underwood, M. K. (2002). Sticks and stones and social exclusion: Aggression among girls and boys. In P. K. Smith & C. H. Hart (Eds.), *Blackwell handbook of childhood social development* (pp. 533-548). Malden, MA: Blackwell.
- Williams, J. E. & Bennett, S. M. (1975). The definition of sex stereotypes via the adjective check list. *Sex Roles*, 1, 327-337. <http://dx.doi.org/10.1007/BF0028722>