

Abstract

Our experiment is aimed at understanding how employee reactions to negative feedback are received by the feedback provider, and how employee gender may play a role in the process. We focus specifically on the act of crying and, based on role congruity theory, argue that a male employee crying in response to negative performance feedback will be seen as atypical behavior by the feedback provider, which will bias evaluations of the employee on a number of different outcome variables, including performance evaluations, assessments of leadership capability, and written recommendations. That is, we expect an interactive effect between gender and crying on our outcomes; an effect that will be mediated by perceived typicality. We find support for our moderated mediation model in a sample of 169 adults, indicating that men who cry in response to negative performance feedback will experience biased evaluations from the feedback provider. Theoretical and practical implications are discussed.

Keywords: performance appraisal; feedback; role congruity; gender; crying

Boys, Don't Cry: Gender and Reactions to Negative Performance Feedback

Performance feedback reinforces job-relevant behavior and fosters awareness of performance goal discrepancies (Cianci, Klein, & Seijts, 2010; Klein, 1989). Receiving feedback often represents an affective event for employees (Weiss & Cropanzano, 1996), particularly when the feedback is negative (e.g., Ilies, De Pater, & Judge, 2007), and researchers have examined anger, disappointment, and frustration as typical responses to negative feedback (e.g., Belschak & Den Hartog, 2009; Johnson, 2005).

In this study we expand the potential emotional repertoire of employees and argue that the distressful nature of negative performance feedback can also lead to crying as a response. Crying is a frequent yet understudied behavioral reaction in the workplace (Bouskila-Yam & Kluger, 2011). We argue there is reason to believe that crying occurs in reaction to negative feedback. Grote (2002, p. 133) noted that “no manager should begin a performance appraisal discussion without a box of tissues handy”. In a study of performance appraisal systems, Björkman and Lu (1999) noted that some employees burst out crying when receiving negative feedback.

We believe that the act of crying in reaction to negative performance feedback can have significant implications for both the employee and the feedback provider (here referred to as the supervisor). Typically, studies have examined the effects of negative feedback on the employees themselves in terms of turnover intentions, job commitment, and performance (e.g., Boswell & Boudreau, 2000; Ellickson & Logsdon, 2002; Horvath & Andrews, 2007). We take a slightly different approach and argue that a crying employee can have a significant effect on the supervisor who provided the negative feedback. That is, watching an employee cry can affect a supervisor's perception of that employee.

While people are generally programmed to react to crying behavior, there are sex differences in the degree to which crying is viewed as a socially acceptable behavior (Fischer, Eagly, & Oosterwijk, 2013). Societal norms and pressures discourage men crying (especially once they are no longer young children) and allow or even encourage women crying (Hendriks, Croon, & Vingerhoets, 2008). This creates differences between men and women in both actual crying behavior and societally permitted crying behavior. Based on role congruity theory (Diekmann & Hirnisey, 2007), we argue that there will be an interactive effect between employee gender and employee crying on perceived typicality, which indicates whether something or someone is representative of a class or group (Wong, Steinfeldt, LaFollette, & Tsao, 2011), such that a man crying will be seen as least typical (or most atypical). We then argue that perceived typicality will mediate the effects of the interaction between employee gender and employee crying on supervisor perceptions and behaviors. Specifically, we examine supervisor evaluations of the employee's performance, an assessment of the employee's leadership capability, and the tone of a written recommendation for the employee. We believe that supervisor perceptions and behaviors toward the employee will be most negative when the employee is crying and male because that behavior is seen as most atypical, and atypical behavior is generally punished (Lehavot & Lambert, 2007).

Our study offers a number of contributions to the literature. First, while the vast majority of research on gender discrimination in the workplace has focused on prejudice towards women (Joshi, Neely, Emrich, Griffiths, & George, 2015), we identify a situation where role incongruous behavior can lead to prejudice towards *men*. In doing so, we provide evidence for the bi-directional nature of sexism. Second, we add to the feedback literature by identifying and exploring a novel expression of emotion in reaction to the receipt of negative feedback. This can

help expand our understanding of how bias can affect feedback providers and their evaluations of employees. Third, we expand the scope of role congruity theory, which has traditionally been constrained to the leadership literature (Eagly & Karau, 2002; Ritter & Yoder, 2004), to show that the theory can help explain a broader set of phenomena within organizations.

Role Congruity, Gender, and Crying

Women are expected to play communal roles in society by demonstrating kindness, gentleness, friendliness, concern, sympathy, and caring (Cialdini & Trost, 1998), while men are expected to play agentic roles by demonstrating competence, independence, decisiveness, dominance, power, and aggression (Rosette & Tost, 2010). We consistently reinforce these socially constructed gender roles in our culture (Dardenne, Dumont, & Bollier, 2007). In organizations, advertisements, language, politics, education, and works of art, women are regularly portrayed as benevolent and beautiful, while men are typically characterized as a strong source of authority (Goldenberg & Roberts, 2004; Jost & Kay, 2005). For instance, in picture books for young children, boys are often portrayed as active leaders and girls as passive followers (Anderson & Hamilton, 2005).

Role congruity theory proposes that people respond with prejudice toward others whom they perceive as acting contrary to their prescribed societal roles, often resulting in discriminatory behavior (Diekmann & Hirnisey, 2007). Failure to adhere to socially constructed gender roles is met with backlash and discrimination (Glick, Zion, & Nelson, 1988; Heilman & Chen, 2005). In essence, we perceive individuals who conform to their typical societal roles as more successful or effective than those who do not conform to their typical roles (Roberson, Galvin, & Charles, 2007).

Punishment for atypical behavior has been demonstrated across a number of areas. In the

field of adolescent psychology, researchers found that adolescent males who behave in a gender-atypical way (e.g., being gentle) experience more bullying, teasing, psychological disorders, social rejection, victimization, loneliness, and suicide (Young & Sweeting, 2004). In the field of management, research in leadership (Eagly & Karau, 2002; Ritter & Yoder, 2004), group dynamics (Carli, 2010; Heilman & Haynes, 2005; Lanaj & Hollenbeck, 2014), organizational citizenship behaviors (Heilman & Chen, 2005), and career experiences (Judiesch & Lyness, 1999; Ohlott, Ruderman, & McCauley, 1994) support the notion that we expect men and women to behave according to their typical societal roles. Negative effects are particularly strong when one gender acts in line with the other gender's role expectations, such as when women engage in autocratic or directive leadership behaviors (Butler & Geis, 1990; Eagly, Makhijani, & Klonsky, 1992) or when women work in male dominated professions (e.g., shipyards) which require agentic characteristics (e.g., physical strength; Fiske & Stevens, 1993).

In sum, according to role congruity theory atypical behavior is penalized. We believe that role expectations extend to performance feedback contexts, and male employees crying in response to receiving negative performance feedback will be viewed as engaging in role incongruous (i.e., atypical) behavior. While adult crying occurs across both sexes (Frijda, 1986; Young, 1937) we argue that crying is much more fitting for individuals who are perceived to possess communal (e.g., gentle, vulnerable) rather than agentic (e.g., powerful, dominant) characteristics. As a man is supposed to be powerful and dominant, the act of crying (which connotes helplessness, vulnerability, and weakness) violates our perceptions of his typical societal role. Research provides support for our arguments and the notion that a crying man is seen as atypical. Popular sayings like "boys don't cry" and "crying is for girls" (Good, Sherrod, & Dillon, 2000) reinforce the presence of societal norms regarding gender and crying behavior.

School-age boys who cry are labeled “fags” or “homos,” ranked as less popular than other classmates, and mocked in front of other students (Adler, Kless, & Adler, 1992). Jessor (1989) found that male crying triggered more confusion and avoidance among participants, especially men. Cretser, Lombardo, Lombardo, and Mathis (1982) reported the existence of a “double standard” in which crying women were more accepted than crying men, and that male crying was perceived as weak and atypical. Both men and women reported a tendency to comfort female over male criers with sympathy, acceptance, and willingness to help (Cretser et al., 1982).

Thus, we expect that a man crying in response to negative performance feedback will violate the assumptions of role congruity theory, leading to the following hypothesis:

Hypothesis 1. When an employee receives negative performance feedback, there will be an interaction between employee gender and employee crying on perceived typicality, such that perceived typicality is lowest when a male employee is crying.

In line with role congruity theory and our predisposition to punish atypical behavior, we propose that atypicality in the workplace will be discouraged. While the question of whether or not this tendency is justified in the modern world is certainly valid, it does not negate the fact that the majority of us demonstrate an automatic inclination to discourage and punish those who behave atypically, particularly when the behavior is more fitting of the opposite gender. In the context of performance feedback, we expect that supervisors who perceive atypical behavior in response to negative performance feedback will be biased against the employee. We examine how this bias manifests across a number of valued outcomes that can determine whether or not the employee is successful in the future, including performance evaluations (Arvey & Murphy, 1998; Greenberg, 1986; Lambert, 2003; Wombloh, 1996), assessments of leadership capability (Lord, Foti, & de Vader, 1984), and the tone of a written recommendations on behalf of the

employee by the supervisor. A letter of recommendation represents a type of performance evaluation in which the writer documents the employee's past behavior, reports positive and negative attributes, and assesses the probability of future success (Judge & Higgins, 1998; Knouse, 1983; Shea, O'Grady, Morrison, Wagner, & Morris, 2008).

We propose that the interactive effects of crying and gender on all three outcomes will be mediated by perceived typicality, leading to the following hypothesis:

Hypothesis 2. When an employee receives negative performance feedback, the indirect effect of employee crying on (a) performance evaluations, (b) assessments of leadership capability, and (c) tone of recommendation letter through perceived typicality is moderated by employee gender such that the indirect effect is significant only when the male employee is crying.

Method

Participants and Procedure

Our experiment was conducted using a sample of 169 adults recruited through the online research sampling tool Microworkers¹ (Nguyen, 2014). Management scholars argue that this is an ideal environment in which to conduct controlled, yet realistic studies utilizing a diverse group of participants (Aguinis & Lawal, 2012). Online workers provide honest and meaningful answers, and exert respectable levels of effort on tasks (Hossain & Kauranen, 2015; Suri, Goldstein, & Mason, 2011).

To participate in our experiment, workers had to be U.S. residents. The average age of participants was 32.19 and 58% were female. In terms of employment status, 36.3% of participants were full-time employees, 19.4% were part-time employees, 20% were self-employed, and 24.4% were unemployed at the time. Contributing to the generalizability of our

results, the employed participants worked across a wide variety of fields such as business management, marketing, information technology, health care, and education, with 66% having held their job for more than two years. Most participants reported an income that fell between \$25,000 and \$40,000. Among the participants, 58% had managed employees in the past, 41% had given a performance evaluation in the past, and 28% had been in a situation where an employee had cried during a performance evaluation.

To decrease the possibility of careless responding, participants had to pass a two-item quality control test in order to receive payment and be included in our data set. The items (*“please mark this item as strongly agree”* and *“please mark this item as strongly disagree”*) were randomly interjected in the survey. Of the 181 workers who finished the survey, 12 workers were excluded for failing the test, resulting in a final sample of 169 individuals. The excluded participants did not differ significantly from the included participants in age, gender, employment status, profession, tenure, or income.

In the experiment, participants were randomly assigned in a 2 (employee crying: yes vs. no) \times 2 (employee gender: male vs. female) between-subjects factorial design. Participants were told that they were going to watch a video recording of a performance evaluation. It was made clear that the employee in the video, Pat, was a grocery store manager² who was receiving performance feedback from his or her direct supervisor. In reality, Pat was an actor hired for the study. Participants were told to imagine that they were the direct supervisor. Before viewing the video, all participants read through Pat’s fictional resume, which described Pat as “talented, knowledgeable, and resourceful” to increase psychological realism. It also listed qualifications, professional experience, awards, and education. Participants then viewed the video, answered manipulation check questions, and provided perceptions of typicality. Participants then provided

performance evaluations, an assessment of leadership capability, and wrote a letter of recommendation (outcomes were counterbalanced), before completing a set of control questions (i.e., attractiveness) and demographic information.

Manipulations and Measures

Gender and crying. Two actors auditioned and were selected to play the role of Pat (the employee was named Pat in all conditions). Both actors were majoring in drama and had ample experience acting and performing, were selected from a pool of candidates to resemble each other as close as possible, and were white and in their early twenties. Nevertheless, we controlled for participants' perceptions of Pat's attractiveness in all analyses. The same female and the same male played Pat for both the crying and non-crying videos.

The four videos, filmed by a professional videographer, were choreographed in such a way that the camera was always focused on Pat, and only a portion of the back of the direct supervisor's head was visible (i.e., participants never saw the direct supervisor). The direct supervisor, whose name was not given but was always the same male, gave the performance feedback. Throughout the video, Pat was sitting across a table from the direct supervisor. The table and background were completely white to prevent any visual distractions. Participants saw one of the four videos. Each video was approximately six minutes long.

Pat and the direct supervisor followed a script that was exactly the same across all videos. It was strongly emphasized that with the exception of crying, the actors should keep their performance, tone, hand gestures, and facial expressions exactly the same across both videos. The videos began with a short introduction in which the direct supervisor explained to Pat how performance evaluations work. He then gave Pat the opportunity to judge his or her own performance over the past six months. After this, the direct supervisor described Pat's

performance as unsatisfactory, and explicated that Pat had not achieved the goals that upper management had prescribed. More specifically, he told Pat that he or she was often late, discourteous, rude to staff members, and that sales had gone down 1% while stores in the surrounding area had not seen a similar decrease. It was at this point that Pat began to cry in the crying conditions. Crying occurred at the same point in time and lasted for the same amount of time across videos. The actors shed tears, occasionally wiped their eyes with their hands, and sobbed while interacting with the direct supervisor. At the end of the videos the direct supervisor stated that he would have to have a serious discussion about Pat's employment status at the company in the future. The video ended after the direct supervisor dismissed Pat from the room.³

Manipulation checks. In order to verify that participants watched the video we asked three yes or no questions: "*Was Pat male or female?*", "*Was the direct supervisor male or female?*", and "*Was Pat crying in the video?*" All participants in each condition answered the questions correctly. If participants responded yes to the last question, they were also asked "*How intense was Pat's crying?*" on a 7-point Likert scale from (1) *not intense at all* to (7) *very intense*. With regards to crying intensity, the responses when Pat was male ($M = 3.38$, $SD = 1.67$) did not differ significantly from when Pat was female ($M = 3.41$, $SD = 1.54$), $F(1, 77) = .005$, *ns*.

Perceptions of typicality. Participants completed a five question perceptions of typicality scale adapted from Wong et al. (2011). All questions were answered on a 7-point Likert scale from (1) *strongly disagree* to (7) *strongly agree*. Example questions included "*In this situation, Pat's expression of emotion is typical of most employees*" and "*Most employees would not have displayed emotion the way that Pat did in this situation*" (reverse coded). The scale showed good internal reliability ($\alpha = .87$).

Performance evaluations. Participants completed a three item performance evaluation

scale adapted from Heilman and Chen (2005). The first question, “*Overall, how would you rate this employee’s performance over the past year?*” was rated on a 7-point Likert scale from (1) *poor* to (7) *excellent*. The other two questions, “*In your opinion, how likely is it that this employee will advance in the company?*” and “*Give your assessment of the employee’s likelihood of success*” were rated on a 7-point Likert scale from (1) *very unlikely* to (7) *very likely*. Together, these items showed good internal reliability ($\alpha = .88$).

Leadership capability. Participants completed the same four item scale used in Rosette, Leonardelli, and Phillips (2008) based on capabilities associated with effective leadership (e.g., Epitropaki & Martin, 2004; Lord et al., 1984). Participants evaluated Pat on his or her intelligence, competence, confidence, and competitiveness on a 7-point Likert scale from (1) *strongly disagree* to (7) *strongly agree*. Example questions included, “*I think that Pat is intelligent,*” and “*I think that Pat is competent.*” The scale showed internal good reliability ($\alpha = .81$).

Recommendation tone. Participants read the following instructions: “*Pat is planning to move next month for family reasons. Pat has asked you, the direct supervisor, to write a letter of recommendation to help in finding a new job. Given what you know about Pat (CV and performance evaluation video), please write a recommendation for Pat.*” Two independent raters, blind to condition, then read and rated the recommendations on a scale from (1) *very negative* to (7) *very positive*. The raters were instructed to base their ratings on the hedonic tone of the recommendation (De Dreu, Baas, & Nijstad, 2008). An example of a negative recommendation was “*Well, I have to be honest, I’m not putting my name on the line for a slacker like Pat, I suggest that he get to the nearest McDonald’s and start working his way up from the bottom. Based on the reports I received I can’t in good conscience recommend that anywhere hire him.*”

With that being said, good luck, if all else fails maybe he can get a gig flipping signs. All the best.” An example of a positive recommendation was “*Pat was an exemplary worker and has always been maximally committed to work and I have nothing but praise for him. I guarantee that he will always try more than most employees.*” Inter-rater reliability was excellent. The average ICC(2) for the two raters was .85 (Fleiss, 1986; Fraquelli et al., 2007; Rico, Sánchez-Manzanares, Antino, & Lau, 2012). Therefore, we averaged the ratings from both raters to form the overall measure of recommendation tone.

Control variables and demographic information. Participants rated Pat’s attractiveness on a 7-point Likert scale from (1) *very unattractive* to (7) *very attractive*. The average level of attractiveness was 4.23 ($SD = 2.63$) for the male actor and 5.79 ($SD = 2.19$) for the female actor. We controlled for attractiveness in every analysis. We also controlled for the gender of the participant (0 = male, 1 = female) in every analysis. There is a tendency to evaluate attractive people more generously (Marlowe, Schneider, & Nelson, 1996) and there may be an in-group gender bias (Ashforth & Mael, 1989).

Results

See Table 1 for all means and standard deviations. See Table 2 for correlations among all study variables.

To test Hypothesis 1, we performed a 2×2 between-subjects ANOVA on perceived typicality, controlling for employee attractiveness and participant gender. While there was no main effect of gender, $F(1, 154) = 1.54$, *ns*, there was a main effect of crying, $F(1, 154) = 13.19$, $p < .001$, which was qualified by a significant interaction, $F(1, 154) = 5.05$, $p < .05$. Separate pairwise comparisons indicated that the interaction was driven by the male crying condition. Perceived typicality in the male crying condition was significantly lower than in the female

crying condition $t(165) = 2.04, p < .05$, the male no crying condition $t(165) = 3.97, p < .001$, and the female no crying condition, $t(165) = 3.53, p < .001$, supporting Hypothesis 1. Perceived typicality in the female crying condition was not significantly different from the female no crying condition $t(165) = 1.44, ns$, though it was lower than the male no crying condition, $t(165) = 1.90, p < .10$. Perceived typicality in the female no crying condition was not significantly different from the male no crying condition, $t(165) = .52, ns$.⁴ See Figure 1 for a graph of the interaction.

Our second hypothesis proposed moderated mediation, which signifies that the strength of an indirect effect is contingent upon the moderator (Preacher, Rucker, & Hayes, 2007). Specifically, we proposed that when receiving negative performance feedback, the indirect effect of employee crying on our outcomes through perceived typicality is observed only when a male employee is crying. In order to provide evidence for moderated mediation, the index of moderated mediation must be significantly different from zero (Hayes, 2015). If the bootstrap confidence interval generated for the index does not include zero, this leads to the inference that moderated mediation is present. Identifying which indirect effect at which particular level of the moderator is significant involves generating bootstrap confidence intervals for each indirect effect at each particular level of the moderator. We used the Monte Carlo approach for constructing bootstrap confidence intervals in estimating the effects (Bauer, Preacher, & Gil, 2006; Shrout & Bolger, 2002).

To test Hypothesis 2a, we entered emotion expression (crying vs. no crying) as the independent variable, perceived typicality as the mediator, performance evaluations as the dependent variable, and employee gender (male vs. female) as the moderator. The estimates, standard errors, and bias-corrected 95 percent confidence intervals for the conditional indirect

effects using 10,000 bootstrap resamples indicated that the index of moderated mediation was significantly different from zero, coefficient = .12, SE = .09, CI [.01, .36]. The indirect effect of crying on performance evaluations through perceived typicality was significant when the employee was male, coefficient = -.16, SE = .09, CI [-.36, -.02], though the indirect effect was not significant when the employee was female, coefficient = -.04, SE = .04, CI [-.16, .02]. Thus, Hypothesis 2a was supported.

We used the same strategy for testing Hypothesis 2b, entering assessments of leadership capability as the dependent variable. The index of moderated mediation was significantly different from zero, coefficient = .19, SE = .12, CI [.02, .51]. The indirect effect of crying on assessments of leadership capability through perceived typicality was significant when the employee was male, coefficient = -.24, SE = .12, CI [-.52, -.05], though not significant when the employee was female, coefficient = -.06, SE = .06, CI [-.24, .03], supporting Hypothesis 2b.

For Hypothesis 2c, we entered recommendation tone as the dependent variable. The index of moderated mediation was significantly different from zero, coefficient = .25, SE = .15, CI [.03, .66]. The indirect effect of crying through perceived typicality was significant when the employee was male, coefficient = -.33, SE = .15, CI [-.69, -.09] but not when the employee was female, coefficient = -.08, SE = .08, CI [-.30, .05], supporting Hypothesis 2c.

Discussion

“These foolish drops do something drown my manly spirit”

-Shakespeare, *The Merchant of Venice*

The belief that crying is unmanly has been ingrained in our culture and is often referenced in literature, film, and music. Men who cry are considered abnormal because they deviate significantly from models of typical behavior (Labott & Martin, 1990). In this study, we

found that cultural bias bleeds into organizational life. Although crying does not represent an everyday occurrence in organizations, we argue that it represents a potential reaction to negative performance feedback due to the distressing nature of the situation. For men who are expected to exhibit agentic characteristics and act in a dominant manner, crying is not an option. When men cry, our results suggest that they are labeled as atypical, which has consequences for how others evaluate them at work, leading to lower performance evaluations and assessments of leadership capability, and a more negative tone in letters of recommendation. Our results are consistent with the tenets of role congruity theory and indicate that incongruous behavior can result in significant bias against male employees in certain contexts, particularly when that behavior is more fitting of female role expectations.

As noted by Joshi et al. (2015, p. 1459), “sexism both overt and subtle remains pervasive in many professional domains”. Although we agree with the statement in general, we feel it is important to note that the overwhelming majority of research in the field of gender and diversity has focused on sexism in one direction – sexism that contributes to the plight of women. Without question there are valid reasons for doing so. Nevertheless, ignoring sexism towards men, which is a form of bias in and of itself, prevents scholars from exploring the full-range effects of gender in the workplace. While there may be fewer situations where men experience sexism, they do exist, as indicated by our results. We need to acknowledge that gender bias can harm men just as it can harm women. In doing so, researchers may find theories that have been traditionally used to investigate bias against women, such as role congruity theory, can also operate to the detriment of men.

Our findings also contribute to the literature on role congruity theory. Specifically, we show that crying creates perceived incongruence for males, which disadvantages them in a

number of different ways, including how others perceive them as leaders. Our study is one of the first to show that men get punished for behaving in a manner opposite of our agentic expectations. We believe this can help expand our understanding of the role of gender in the workplace and may help to identify other contexts where men are at a disadvantage for behaving in role incongruous ways. For example, in responding to crises, researchers argue that leaders need to be strong and decisive (Stewart, 1967, 1976; Yukl, 2002) and more directive, autocratic, and goal-focused in their behavior (Mulder, Ritsema van Eck, & de Jong, 1970). As noted by Pillai and Meindl (1998, p. 647), “crises provide leaders with opportunities to take bold purposeful action”. If male leaders violate our assumption of agentic action, they will likely be evaluated more negatively.

Our results add to research examining the expression of emotion at work. Although crying has not received much attention from organizational researchers, a recent article by Sinaceur, Kopelman, Vasiljevic, and Haag (2015) examined whether expressions of sadness, such as sighing or saying things like “this almost brings tears to my eyes,” induced concessions during negotiations. The authors argued and found that expressions of sadness send the message that the expresser needs help and support, which stimulates empathy and compassion, leading to increased concessionary behavior. We did not find direct effects for crying on concessionary behavior in our study, which would be reflected in our outcome variables. First, crying is a much more intense behavior than sighing and may, in general, be seen as less appropriate in business contexts. As evidence, we found direct effects for crying on perceptions of typicality. Regardless of gender, crying was seen as less typical. It would be interesting for future research to compare crying and more subtle expressions of sadness in negotiation settings. Second, our results likely highlight the importance of social aspects of the situation and, in a performance feedback setting,

the supervisor is less likely to be motivated to act on any sort of other-concern. In other situations (e.g., group work), where motivations are different, crying could stimulate a strong helping response from others. Finally, we note that Sinaceur et al. did not test whether the gender of the expresser moderated the effects of sadness expression. Our results suggest that this might be an interesting direction for future research in negotiations as men may be much less likely to benefit from expressions of sadness because such behavior runs counter to agentic expectations.

Our results inform research on reactions to negative performance feedback. Thus far, researchers have focused on emotions such as anger (e.g., Belschak & Den Hartog, 2009; Johnson, 2005). Generally, researchers have examined whether experiencing anger has implications for outcomes such as attitudes at work and turnover intentions. Our work suggests that the *expression* of anger by employees may also matter. Based on role congruity theory, expressing anger in an aggressive or hostile manner, such as threatening the supervisor, would likely be seen as incongruous behavior for women because it violates our expectations of their communal nature. As a result, the expression of different emotions (aside from crying) in reaction to negative performance feedback could handicap women rather than men.

We feel our findings have significant practical implications as well. We advise that men be cautious of the emotions they display during a negative performance appraisal. We also suggest that supervisors be wary of potential discrimination. Studies like this can be presented to performance appraisers as evidence of a biased tendency to rate crying men worse even though their performance might be exactly the same as that of a crying female. In the worst cases, evidence of such discrimination could wind up in a court of law, where employees argue that they were discriminated against because of atypical displays of emotion. Recent evidence suggests that discrimination against men is not as unusual as it may seem. In January 2015, The

Equal Opportunity Employment Commission filed a lawsuit against casual restaurant chain Ruby Tuesday for overt gender discrimination against male job applicants (Philip, 2015).

Finally, our work is not without its limitations. We examined immediate reactions to crying during performance feedback. Normally there is a lag between performance feedback and, for example, writing a letter of recommendation. The question remains as to whether reactions would continue to exert an effect over time. Another limitation involves the fact that participants were not face-to-face with the crying employee. In organizations, performance feedback is typically given in person, though it is increasingly common for organizations to provide performance feedback virtually (Farr, Fairchild, & Cassidy, 2013; Weisband & Atwater, 1999). Accordingly, a potential avenue for future research would be to examine how being in the same room with a crying employee affects supervisor reactions, although we believe it would likely strengthen the results reported here.

Conclusion

The expression of emotion at work is an impactful event, particularly when the expression violates our expectations of appropriate behavior. Unfortunately our expectations are not without bias and, when it comes to crying, we feel it is less appropriate for men to engage in such behavior. When they do, we label them as atypical and downgrade their standing on a number of outcomes that are personally and professionally significant for the employee. Our findings contribute to the literature on gender, role congruity, and the expression of emotion at work; findings that highlight potential discrimination against men in performance feedback settings.

References

Adler, P. A., Kless, S. J., & Adler, P. (1992). Socialization to gender roles: Popularity among elementary school boys and girls. *Sociology of Education*, *65*, 169-187.

Aguinis, H., & Lawal, S. O. (2012). Conducting field experiments using eLancing's natural environment. *Journal of Business Venturing*, *27*, 493-505.

Anderson, D. A., & Hamilton, M. (2005). Gender role stereotyping of parents in children's picture books: The invisible father. *Sex Roles*, *52*, 145-151.

Arvey, R. D., & Murphy, K. R. (1998). Performance evaluation in work settings. *Annual Review of Psychology*, *49*, 141-168.

Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. *Academy of Management Review*, *14*, 20-39.

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods*, *11*, 142-163.

Belschak, F. D., & Den Hartog, D. N. (2009). Consequences of positive and negative feedback: The impact on emotions and extra-role behaviors. *Applied Psychology*, *58*, 274-303.

Björkman, I., & Lu, Y. (1999). The management of human resources in Chinese-Western joint ventures. *Journal of World Business*, *34*, 306-324.

Boswell, W. R., & Boudreau, J. W. (2000). Employee satisfaction with performance appraisals and appraisers: The role of perceived appraisal use. *Human Resource Development Quarterly*, *11*, 283-299.

- Bouskila-Yam, O., & Kluger, A. N. (2011). Strength-based performance appraisal and goal setting. *Human Resource Management Review, 21*, 137-147.
- Butler, D., & Geis, F.L. (1990). Nonverbal affect responses to male and female leaders: Implications for leadership evaluations. *Journal of Personality and Social Psychology, 58*, 48-59.
- Carli, L. L. (2010). Gender and group behavior. In J. C. Chrisler, & D. R. McCreary (Eds.), *Handbook of gender research in psychology* (pp. 337-358). New York, NY: Springer.
- 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, MA: McGraw-Hill.
- Cianci, A. M., Klein, H. J., & Seijts, G. H. (2010). The effect of negative feedback on tension and subsequent performance: The main and interactive effects of goal content and conscientiousness. *Journal of Applied Psychology, 95*, 618-630.
- Cretser, G. A., Lombardo, W. K., Lombardo, B., & Mathis, S. (1982). Reactions to men and women who cry: A study of sex differences in perceived societal attitudes versus personal attitudes. *Perceptual and Motor Skills, 55*, 479-486.
- Dardenne, B., Dumont, M., & Bollier, T. (2007). Insidious dangers of benevolent sexism: Consequences for women's performance. *Journal of Personality and Social Psychology, 93*, 764-779.

- De Dreu, C. K. W., Baas, M., & Nijstad, B. A. (2008). Hedonic tone and activation level in the mood-creativity link: Toward a dual pathway to creativity model. *Journal of Personality and Social Psychology, 94*, 739–756.
- Diekmann, A. B., & Hirnisey, L. (2007). The effect of context on the silver ceiling: A role congruity perspective on prejudiced responses. *Personality and Social Psychology Bulletin, 33*, 1353-1366.
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*, 573-598.
- Eagly, A.H., Makhijani, M.G., & Klonsky, B.G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin, 111*, 3-22.
- Ellickson, M. C., & Logsdon, K. (2002). Determinants of job satisfaction of municipal government employees. *Public Personnel Management, 31*, 343-358.
- Epitropaki, O., & Martin, R. (2004). Implicit leadership theories in applied settings: Factor structure, generalizability, and stability over time. *Journal of Applied Psychology, 89*, 293-310.
- Farr, J. L., Fairchild, J., & Cassidy, S. E. (2013). Technology and performance appraisal. In M. D. Covert, & L. F. Thompson (Eds.), *The psychology of workplace technology* (pp. 77-98). New York, NY: Routledge.
- Fischer, A. H., Eagly, A. H., & Oosterwijk, S. (2013). The meaning of tears: Which sex seems emotional depends on the social context. *European Journal of Social Psychology, 43*, 505-515.

- Fiske, S. T., & Stevens, L. E. (1993). *What's so special about sex? Gender stereotyping and discrimination*. Thousand Oaks, CA: Sage Publications, Inc.
- Fleiss, J. L. (1986). Reliability of measurement. In J. L. Fleiss (Ed.), *The design and analysis of clinical experiments* (pp. 1-32). New York, NY: John Wiley & Sons.
- Fraquelli, M., Rigamonti, C., Casazza, G., Conte, D., Donato, M. F., Ronchi, G., & Colombo, M. (2007). Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. *Gut*, *56*, 968-973.
- Frijda, N. H. (1986). *The emotions: Studies in emotion and social interaction*. New York, NY: Cambridge University Press.
- Glick, P., Zion, C., & Nelson, C. (1988). What mediates sex discrimination in hiring decisions?. *Journal of Personality and Social Psychology*, *55*, 178-186.
- Goldenberg, J. L., & Roberts, T. A. (2004). The beast within the beauty: An existential perspective on the objectification and condemnation of women. In J. Greenberg, S. L. Koole, & T. Pyszczynski (Eds.), *Handbook of experimental existential psychology* (pp. 71-85). New York, NY: Guilford Press.
- Good, G. E., Sherrod, N. B., & Dillon, M. G. (2000). Masculine gender role stressors and men's health. In R. M. Eisler, & M. Hersen (Eds.), *Handbook of gender, culture, and health* (pp. 63-81). Mahwah, NJ: Lawrence Erlbaum Associates.

- Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26*, 213-224.
- Greenberg, J. (1986). Determinants of perceived fairness of performance evaluations. *Journal of Applied Psychology, 71*, 340-342.
- Grote, R. C. (2002). *The performance appraisal question and answer book: A survival guide for managers*. New York, NY: AMACOM Division of American Management Association.
- Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research, 50*, 1-22.
- 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.
- Heilman, M. E., & Haynes, M. C. (2005). No credit where credit is due: Attributional rationalization of women's success in male-female teams. *Journal of Applied Psychology, 90*, 905-916.
- Hendriks, M. C., Croon, M. A., & Vingerhoets, A. J. (2008). Social reactions to adult crying: The help-soliciting function of tears. *The Journal of Social Psychology, 148*, 22-42.
- Hirth, M., Hoßfeld, T., & Tran-Gia, P. (2011, June). Anatomy of a crowdsourcing platform – using the example of microworkers.com. In *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2011 Fifth International Conference on* (pp. 322-329). IEEE.

- Horvath, M., & Andrews, S. B. (2007). The role of fairness perceptions and accountability attributions in predicting reactions to organizational events. *The Journal of Psychology, 141*, 203-222.
- Hossain, M., & Kauranen, I. (2015). Crowdsourcing: A comprehensive literature review. *Strategic Outsourcing: An International Journal, 8*, 2-22.
- Ilies, R., De Pater, I. E., & Judge, T. (2007). Differential affective reactions to negative and positive feedback, and the role of self-esteem. *Journal of Managerial Psychology, 22*, 590-609.
- Jesser, C. J. (1989). Men and crying. *Changing Men, 20*, 12-15.
- Johnson, J. (2005). Employees' justice perceptions of performance appraisal systems: Attitudinal, behavioral, and performance consequences. *Journal of Applied Psychology, 79*, 937-939.
- Joshi, A., Neely, B., Emrich, C., Griffiths, D., & George, G. (2015). Gender research in AMJ: An overview of five decades of empirical research and calls to action. *Academy of Management Journal, 58*, 1459-1475.
- Jost, J. T., & Kay, A. C. (2005). Exposure to benevolent sexism and complementary gender stereotypes: consequences for specific and diffuse forms of system justification. *Journal of Personality and Social Psychology, 88*, 498-509.
- Judge, T. A., & Higgins, C. A. (1998). Affective disposition and the letter of reference. *Organizational Behavior and Human Decision Processes, 75*, 207-221.
- Judiesch, M. K., & Lyness, K. S. (1999). Left behind? The impact of leaves of absence on managers' career success. *Academy of Management Journal, 42*, 641-651.

- Klein, H. J. (1989). An integrated control theory model of work motivation. *Academy of Management Review*, *14*, 150-172.
- Knouse, S. B. (1983). The letter of recommendation: Specificity and favorability of information. *Personnel Psychology*, *36*, 331-341.
- Labott, S. M., & Martin, B. M. (1990). Emotional coping, age, and physical disorder. *Behavioral Medicine*, *16*, 53-61.
- Lambert, E. (2003). The impact of organizational justice on correctional staff. *Journal of Criminal Justice*, *31*, 155-168.
- Lanaj, K., & Hollenbeck, J. (2014). Leadership over-emergence in self-managing teams: The role of gender and countervailing biases. *Academy of Management Journal*, *58*, 1476-1494.
- Lehavot, K., & Lambert, A. J. (2007). Toward a greater understanding of antigay prejudice: On the role of sexual orientation and gender role violation. *Basic and Applied Social Psychology*, *29*, 279-292.
- Lord, R. G., Foti, R. J., & de Vader, C. D. (1984). A test of leadership categorization theory: Internal structure, information processing, and leadership perception. *Organizational Behavior and Human Decision Making Processes*, *34*, 343-378.
- Marlowe, C. M., Schneider, S. L., & Nelson, C. E. (1996). Gender and attractiveness biases in hiring decisions: Are more experienced managers less biased? *Journal of Applied Psychology*, *81*, 11-21.

Mulder, M., Ritsema van Eck, J. R., & de Jong, R. D. (1970). An organization in crisis and non-crisis conditions. *Human Relations, 16*, 317–334.

Nguyen, N. (2014, November). Microworkers crowdsourcing approach, challenges and solutions. In *Proceedings of the 2014 International ACM Workshop on Crowdsourcing for Multimedia* (pp. 1-1). doi: 10.1145/2660114.2660128

Ohlott, P. J., Ruderman, M. N., & McCauley, C. D. (1994). Gender differences in managers' developmental job experiences. *Academy of Management Journal, 37*, 46-67.

Philip, A. (2015, January 26). Ruby Tuesday accused of employment discrimination. Against men. *The Washington Post*. Retrieved from: <https://www.washingtonpost.com>.

Pillai, R., & Meindl, J. R. (1998). Context and charisma: A “meso” level examination of the relationship of organic structure, collectivism, and crisis to charismatic leadership. *Journal of Management, 24*, 643-671.

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42*, 185-227.

Rico, R., Sánchez-Manzanares, M., Antino, M., & Lau, D. (2012). Bridging team faultlines by combining task role assignment and goal structure strategies. *Journal of Applied Psychology, 97*, 407-420.

Ritter, B. A., & Yoder, J. D. (2004). Gender differences in leader emergence persist even for dominant women: An updated confirmation of role congruity theory. *Psychology of Women Quarterly, 28*, 187-193.

- Roberson, L., Galvin, B. M., & Charles, A. C. (2007). When group identities matter: Bias in performance appraisal. *The Academy of Management Annals*, *1*, 617-650.
- Rosette, A. S., Leonardelli, G. J., & Phillips, K. W. (2008). The White standard: Racial bias in leader categorization. *Journal of Applied Psychology*, *93*, 758-777.
- Rosette, A. S., & Tost, L. P. (2010). Agentic women and communal leadership: How role prescriptions confer advantage to top women leaders. *Journal of Applied Psychology*, *95*, 221-235.
- Shea, J. A., O'Grady, E., Morrison, G., Wagner, B. R., & Morris, J. B. (2008). Medical student performance evaluations in 2005: An improvement over the former dean's letter? *Academic Medicine*, *83*, 284-291.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422-445.
- Sinaceur, M., Kopelman, S., Vasiljevic, D., & Haag, C. (2015). Weep and get more: When and why sadness expression is effective in negotiations. *Journal of Applied Psychology*. Advance online publication. <http://dx.doi.org/10.1037/a0038783>
- Stewart, R. (1967). *Managers and their jobs: A study of the similarities and differences in the way managers spend their time*. London, England: MacMillan.
- Stewart, R. (1976). *Contrasts in management*. Berkshire, England: McGraw-Hill.

- Suri, S., Goldstein, D. G., & Mason, W. A. (2011, August). Honesty in an online labor market. In L. von Ahn, & P. G. Ipeirotis (Eds.), *Papers from the 2011 AAI Workshop*. Menlo Park, CA: AAI Press.
- Weisband, S., & Atwater, L. (1999). Evaluating self and others in electronic and face-to-face groups. *Journal of Applied Psychology, 84*, 632-639.
- Weiss, H. M., & Cropanzano, R. (1996). Affective events theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. *Research in Organizational Behavior, 19*, 1-74.
- Womboh, B. S. (1996). The management implications of performance discrepancy in a library organization. *Library Management, 17*, 10-19.
- Wong, Y. J., Steinfeldt, J. A., LaFollette, J. R., & Tsao, S. C. (2011). Men's tears: Football players' evaluations of crying behavior. *Psychology of Men & Masculinity, 12*, 297-310.
- Young, P. T. (1937). Laughing and weeping, cheerfulness and depression: A study of moods among college students. *The Journal of Social Psychology, 8*, 311-334.
- Young, R., & Sweeting, H. (2004). Adolescent bullying, relationships, psychological well-being, and gender-atypical behavior: A gender diagnosticity approach. *Sex Roles, 50*, 525-537.
- Yukl, G. (2002). *Leadership in organizations* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Footnotes

¹ Microworkers has a very similar platform and participant pool as Amazon Mechanical Turk (Hirth, Hoßfeld, & Tran-Gia, 2011).

² In order to determine whether a grocery store manager was considered a gender-neutral profession we conducted a study using Amazon Mechanical Turk (Goodman, Cryder, & Cheema, 2013), $N = 40$, 48% female, $M_{\text{age}} = 36.33$. We instructed participants to rate a grocery store manager on 10 adjectives, including five traditionally feminine characteristics (e.g., gentle, communal, kind), $\alpha = .89$, and five traditionally masculine characteristics (e.g., aggressive, agentic, dominant), $\alpha = .76$, (Rosette & Tost, 2010). We asked, “*To what extent do you think a grocery store manager exhibits the following characteristics?*” on a 7-point Likert scale from (1) *not characteristic at all* to (7) *very characteristic*. A within-subjects ANOVA, controlling for participant gender, showed no significant difference between feminine characteristics ($M = 4.32$, $SD = 1.15$) and masculine characteristics ($M = 4.05$, $SD = 1.07$), $F(1, 38) = .12$, *ns*.

³ Videos are available from the first author upon request.

⁴ The effect of participant gender did not qualify the interactive effects found in our study, $F(1, 152) = 1.01$, *ns*.

Table 1.
Means and standard deviations for study variables by condition.

<i>Employee Crying</i> <i>Employee Gender</i>	<i>Yes</i>		<i>No</i>	
	<i>Male</i>	<i>Female</i>	<i>Male</i>	<i>Female</i>
Perceived Typicality	3.90 (1.26)	4.44 (1.13)	4.98 (1.37)	4.84 (1.27)
Performance Evaluations	1.95 (1.09)	2.50 (1.19)	2.33 (1.34)	2.62 (1.14)
Leader Capabilities	2.85 (1.31)	3.57 (1.30)	3.36 (1.36)	3.65 (1.23)
Recommendation Tone	3.60 (1.78)	4.18 (1.49)	3.88 (1.74)	3.99 (1.70)
Employee Attractiveness	4.23 (2.47)	6.03 (2.33)	4.22 (2.81)	5.60 (2.07)

Note. N = 169. Standard deviations are in parentheses. All variables were measured on a scale from 1-7.

Table 2.
Descriptive statistics and correlations among study variables.

Variable	<i>M</i>	<i>SD</i>	1.	2.	3.	4.	5.	6.	7.	8.
1. Employee Gender	.50	.50	-							
2. Employee Crying	.47	.50	-.05	-						
3. Perceived Typicality	4.56	1.32	.08	-.29**	-					
4. Performance Evaluations	2.35	1.21	.18*	-.11	.16*	-				
5. Leader Capabilities	3.36	1.33	.19*	-.12	.20**	.54**	-			
6. Recommendation Tone	3.91	1.68	.10	-.02	.20**	.47**	.25**	-		
7. Employee Attractiveness	5.01	2.53	.31**	.03	-.06	.10	.19*	.05	-	
8. Participant Gender	.58	.50	.05	.01	.07	.02	.09	-.02	-.03	-

Note. N = 169. Employee gender was coded as 0 = Male, 1 = Female. Employee crying was coded as 0 = not crying, 1 = crying. Participant gender was coded as 0 = Male, 1 = Female. All other variables were measured on a scale from 1-7.

* $p < .05$.

** $p < .01$.

Figure Caption.

Figure 1. Interaction between gender and crying on perceived typicality.

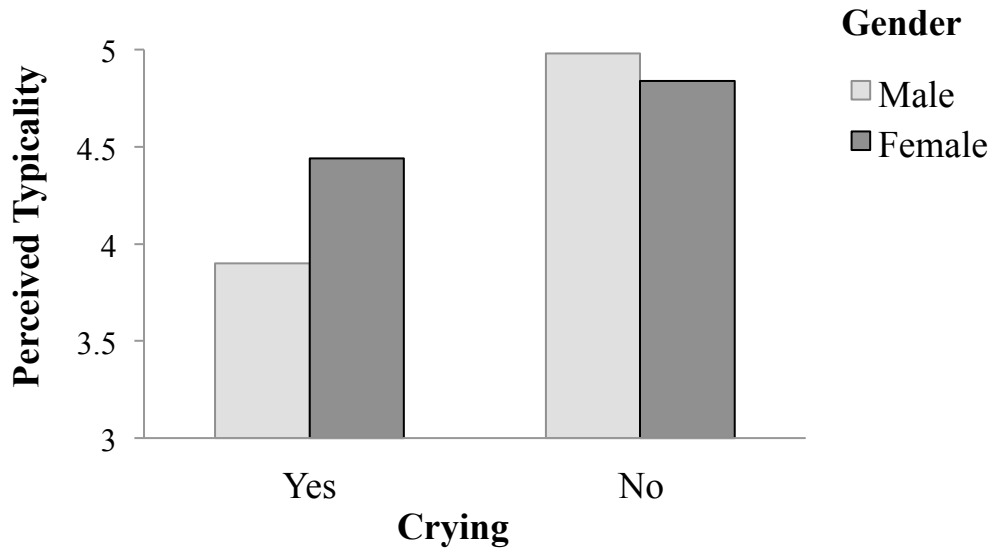


Figure 1.