Article:
Luksyte, A, Unsworth, KL and Avery, DR (2018) Innovative Work Behavior and Sex-Based Stereotypes: Examining Sex Differences in Perceptions and Evaluations of Innovative Work Behavior. Journal of Organizational Behavior, 39 (3). pp. 292-305. ISSN: 0894-3796

https://doi.org/10.1002/job.2219

Innovative Work Behavior and Sex-Based Stereotypes: Examining Sex Differences in Perceptions and Evaluations of Innovative Work Behavior

Author note

Aleksandra Luksyte, Business School, the University of Western Australia; Kerrie Unsworth, Leeds University, Business School; Derek Avery, Wake Forest University, School of Business.

Correspondence concerning this article should be addressed to Aleksandra Luksyte, Business school, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia. Email address: alex.luksyte@uwa.edu.au.

Abstract

Building on role congruity theory, we predict that innovative work behaviors are stereotypically ascribed to men more than to women. Because of this bias, women who innovate may not receive better performance evaluations than those who do not innovate, whereas engaging in innovative work behaviors is beneficial for men. These predictions were supported across three complementary field and experimental studies. The results of an experiment (Study 1; $N = 407$) revealed that innovative work behaviors are stereotypically associated with men more than women. In Studies 2 and 3, using multi-source employee evaluation data ($N = 153$) and by experimentally manipulating innovative work behaviors ($N = 232$), respectively, we found that favorable performance evaluations were associated with innovative work behaviors for men but not for women. These studies highlight a previously unidentified form of sex bias and are particularly important for those wishing to increase innovative behaviors in the workplace: We need to address this phenomenon of 'think-innovation-think-male'.

Keywords: Innovative work behavior; sex; performance evaluations

Given that roughly half of the world's population is female and women's labor force participation rates now rival men's in many developed and developing nations (e.g., in OECD countries labor force participation rates of women and men are 51% vs 69%, respectively; World Bank, 2014), the prospective value of women's workplace contributions are at an all-time high. If companies are to achieve their full potential, they must rely on their female employees' innovative work behaviors or intentional actions to produce or adopt novel ideas, products, and services that benefit multiple stakeholders (Janssen, 2000; Scott & Bruce, 1994). Unfortunately, women's contributions (including innovative work behaviors) may be viewed and valued differently than that of their male colleagues (Kabat-Farr & Cortina, 2012). Take, for instance, this quote from a female product development manager: "I can't tell you how many times I have made suggestions in meetings that were totally ignored" (Blank & Slipp, 1994, p. 153).

One reason to suspect that innovative work behaviors displayed by men and women may be viewed differently is based on role congruity theory (Eagly & Karau, 2002). The theory proposes that (a) many of the attributes of successful leaders (e.g., assertive, decisive) are incongruent with communal characteristics (e.g., nurturing, compassionate) traditionally ascribed to women (Ely, Ibarra, & Kolb, 2011), and (b) this discrepancy between leader and sex roles often results in women being perceived as ineffective leaders (Johnson, Murphy, Zewdie, & Reichard, 2008). It is also worth noting that lack of fit theory involves very similar predictions and has received sound empirical support (Heilman, 2012; Lyness & Heilman, 2006). The theory posits that it is not the negativity of descriptive sex stereotypes (generalized beliefs about what women and men *are* like) that explain why women fare worse in masculine jobs and when engaging in masculine work behaviors. Rather, these stereotypes give rise to a perceived lack of fit between communal attributes ascribed to women and characteristics (usually agentic) needed

to succeed in masculine jobs and work behaviors (Heilman, 2012). We contend that innovative work behavior is viewed as a prototypically masculine activity because, by definition, it is a risky endeavor (Janssen, van de Vliert, & West, 2004) that requires taking initiative (Parker & Collins, 2010) and embracing and championing change (Wu, Parker, & de Jong, 2014) – actions that are likely to be associated with men. Further, prescriptive stereotypes (those delineating how men and women *should* behave) establish normative behavioral expectations for men and women, resulting in negative evaluations of stereotype-inconsistent behaviors (Heilman, 2012). We argue that because innovative work behavior is expected more from men than women, women are perceived to violate this sex stereotype when they innovate. Hence, innovative work behaviors of women will not be rewarded to the same extent as those demonstrated by men.

This lack of recognition of innovative work behaviors displayed by women may explain why employees are often reluctant to innovate (Unsworth & Clegg, 2010), despite the growing importance of individual innovation in the workplace and organizational efforts to promote such behaviors (Anderson, Potocnik, & Zhou, 2014). Similar to creativity, innovative work behavior concerns the generation of novel and useful ideas; yet, it differs from creativity because being innovative means successful implementation of these ideas (Shalley, Zhou, & Oldham, 2004). Innovative work behavior also has been conceptualized as risk-taking because "it may lead to unintended costs for the innovators involved despite their intention to produce anticipated benefits" (Janssen et al., 2004, p. 130). Differently from risk taking however, which encompasses both work and non-work behaviors that "could lead to negative consequences" (Byrnes, Miller, & Schafer, 1999, p. 367), innovative work behavior is displayed only at work and is "explicitly intended to provide some kind of benefit" (de Jong & den Hartog, 2010, p. 24).

Generally speaking, innovative work behavior is likely to be an organizationally-valued behavior because it is intended to bring about improvements (Janssen et al., 2004). Given these positive changes and the centrality of technological innovation in most markets (Kafouros & Forsans, 2012), appraisers are likely to reward employee innovative actions to encourage these behaviors. Yet, innovation also can come with a transition cost, wherein it may take time, effort, and disruption of work routines to accept and implement an innovation (e.g., Choi & Chang, 2009; Janssen et al., 2004). And thus, innovative actions may be ignored or even punished by appraisers who may resist such inconveniences to their work habits. This is an interesting paradox, yet, research is lacking in this domain. Further, in light of sex differences in enactment and outcomes of conceptually similar behavior that, like innovative work behaviors, are potentially risky and challenge the status quo (e.g., creativity, speaking up in professional meetings, providing input in mixed-sex teams; Heilman & Haynes, 2005; Karpowitz, Mendelberg, & Shaker, 2012; Mendelberg, Karpowitz, & Oliphant, 2014; Proudfoot, Kay, & Koval, 2015), and the susceptibility of performance ratings to biases (DeNisi & Sonesh, 2011), we ask whether innovative work behavior is evaluated differently depending on whether a man or a woman innovates. Such differential evaluation of innovative work behaviors for men and women may represent a hidden form of discrimination, which is as detrimental for employees' wellbeing and productivity as overt discrimination (Jones, Peddie, Gilrane, King, & Gray, 2016).

Given the harmful effects of discrimination and the increasing participation rates of women in the workforce (e.g., 59% in Australia and 47% in USA; Australian Bureau of Statistics, 2011; Bureau of Labor Statistics, 2011), examining how and why men and women who innovate are evaluated differently should help managers reduce potential biases in performance evaluation and encourage innovative work behaviors among talented men and

women. In the remainder of the introduction, we propose that innovative work behavior (a) is a masculine type of activity, (b) tends to generally be rewarded in organizations, but (c) is evaluated differently depending on whether a man or a woman innovates because of theorized lack of fit.

Innovative Work Behavior and Sex-Based Stereotypes

As we argued above, the masculine nature of innovative work behavior suggests that it might be more expected from male than female employees. In short, by engaging in innovative work behaviors employees challenge the status quo because they provide an alternative proposal and they often do so in an assertive manner (Parker & Collins, 2010; Wu et al., 2014). Commonly held prescriptive stereotypes, however, prohibit women from "demonstrating the self-assertion, dominance and achievement orientation so celebrated in men" (Heilman, 2012, p. 123). This could result in innovative work behaviors being viewed differently for men and women. Though such a conclusion remains untested to date, studies on similar topics – differential assessment of female and male contributions to mixed-sex teamwork and speaking up in professional meetings – theorized and found similar effects. For example, in comparison to men, women tend to participate less and are more often interrupted during deliberation of important societal issues; yet, the opposite is true under an anonymous rule (when sex of deliberators is unknown) and when a group consists of many women and a majority rule is upheld (Karpowitz et al., 2012; Mendelberg et al., 2014). Female contributions also are devalued or ignored in team settings when working on masculine tasks (e.g., creating an investment portfolio) and, particularly, when there is an ambiguity regarding the input of male and female team members (Heilman & Haynes, 2005). Sex differences were found in self-reported innovative work behavior and creative output. For example, men reported engaging in more

innovative work behavior and producing greater innovative output than did women (de Jong & den Hartog, 2010). This may occur because members of each sex are conforming to expectation (Elsass & Graves, 1997). Sex descriptive stereotypes also might help account for the tendency for the same work to be seen as significantly more creative (d = .66) when signed by a man as opposed to a woman (Lebuda & Karwowski, 2013).

Extending this research, our study not only theorizes that innovative work behavior is a masculine activity, but empirically tests it. This is similar to the recent study on creativity, in which researchers demonstrated that men are perceived to be more creative than women; and men are viewed as agentic and thus more creative and more deserving a reward than women (Proudfoot et al., 2015). Our study goes beyond this research by comparing men and women's behavior and the evaluations thereof, not just perceptions of who they are; and by showing that prescriptive stereotypes influence the entire innovation process, which starts with generating novel and useful ideas (i.e., creativity) and ends with "implementing ideas toward better procedures, practices, or products" (i.e., innovation; Anderson et al., 2014, p. 1298). We argue that not only are women stereotypically described in terms of personal attributes (i.e., less creative; Proudfoot et al., 2015), but they will also not be expected to behave in an innovative way or challenge the status quo, disrupt work routines, and promote new work methods. We theorize that this is because in doing so they do not conform to the behavioral expectations of typical female work behaviors. This leads us to anticipate that innovative work behavior is a masculine type of activity and thus, to our first hypothesis that innovative work behavior is associated more with men than women. Moreover, this forms the theoretical foundation of our proposed differential model of innovative work behaviors for men and women.

Hypothesis 1: Innovative work behavior is ascribed more to men than to women.

Innovative Work Behaviors and Performance Evaluations

Research on innovative work behaviors has been criticized for a rather limited conceptualization of this important work behavior, wherein it has been viewed mainly as the dependent variable (Anderson, De Dreu, & Nijstad, 2004). This is unfortunate because it "leaves the reader with the inaccurate impression that innovations are the final end-product of previous processes which end abruptly at some predetermined point" (Anderson et al., 2004, p. 160). We advance our knowledge of innovative work behaviors by integrating research on performance evaluation to examine innovative work behavior as an independent variable.

Innovative work behavior is a highly desirable outcome for the organization because it is vital for team and organization financial performance (Kafouros & Forsans, 2012; Sung & Choi, 2012). It has been included in wider models of employee performance (Griffin, Neal, & Parker, 2007), such that good performance would incorporate innovative behavior. Accordingly, supervisors and peers likely promote innovative work behaviors by considering them when rating the overall performance of their coworkers who innovate. Yet, innovative work behavior also can create uncertainty (Mueller, Melwani, & Goncalo, 2012) that may yield negative reactions from others due to deeply ingrained resistance to change (Oreg, 2003) and this could offset any positive evaluations. Despite these possibilities, we argue that, *in general*, innovative work behavior will be rewarded by favorable performance evaluations because of the intention to improve existing processes and procedures. Performance appraisal refers to "evaluation of employees' performance in which an evaluator assesses the extent to which certain desired behaviors have been observed or achieved" (DeNisi & Sonesh, 2011, p. 255). Although the effects of task and contextual performance on the ratings of overall performance are well

documented (Rotundo & Sackett, 2002), whether innovative work behavior leads to favorable performance evaluation remains an open question.

We offer three reasons as to why innovative work behavior will likely be rewarded with favorable performance evaluations. First, innovative work behavior includes both task and contextual performance. Whereas task necessities appear to drive much innovative work behavior (Unsworth, Wall, & Carter, 2005), its other forms, such as that based around proactive creativity (Unsworth, 2001), are instead extra-role behaviors. Regarding the former, domain knowledge is required to identify problems and potential solutions (Amabile, 1996); thus innovative work behavior cannot be achieved successfully without a thorough understanding of tasks that contribute to core technical proficiency (Griffin et al., 2007). Employees can generate and implement innovative ideas only if they excel at their jobs and know the core processes enough to be able to improve them. Regarding the latter, and similar to contextual performance (e.g., VanScotter & Motowidlo, 1996), innovative work behavior likely exceeds core task responsibilities because it is often volitional and geared toward multiple beneficiaries (e.g., coworkers who may emulate a novel work method; Anderson et al., 2004).

Second, because of its focus on self-initiated change, innovative work behavior is considered a type of proactivity (Wu et al., 2014). Proactive work behavior is rewarded by positive performance ratings (Grant, Parker, & Collins, 2009). Hence, we argue that innovative work behavior generally will be rated positively by appraisers because of the greater organizational emphasis on employee adaptability and proactivity in the modern turbulent, fast-paced, and changing economy (Kafouros & Forsans, 2012). Third, we propose that the negative evaluations of employees who innovate identified by Mueller et al. (2012) occur only when the appraiser feels that the innovation is uncertain in its potential success. As we discuss in the next

section, such uncertainty may arise depending on whether a man or a woman engages in innovative actions but, in general, will not be relevant across the entire sample. Conceptualizing innovative work behavior as part of the broader performance domain, we hypothesize:

Hypothesis 2: Innovative work behavior is related positively to performance evaluations.

The Moderating Role of Sex in the Evaluation of Innovative Work Behaviors

Our research contributes to the literature by testing whether the hypothesized favorable evaluations of innovative work behavior are more pronounced for men who innovate than for women displaying the same behaviors. We argue that due to stereotypical sex roles, it is desirable and expected for men to engage in innovative work behavior. As such, they are rewarded for displaying these behaviors because in doing so they behave consistently with the stereotypes (e.g., Proudfoot et al., 2015). Conversely, women are not expected to innovate and, therefore, when they deviate from these prescriptive directives, they are not rewarded for engaging in innovative work behaviors because they violate behavioral expectations about who should innovate in organizational settings. It is well established that performance ratings are susceptible to a variety of inaccuracies and biases (Avery, 2011; DeNisi & Sonesh, 2011; Roberson, Galvin, & Charles, 2007). Yet, the role of sex stereotypes in these processes remains understudied, thereby encouraging researchers to "learn more about how [sex] stereotypes operate during the appraisal process and how they affect rater motivation to be accurate in actual field settings" (DeNisi & Sonesh, 2011, p. 271). We attempt to address this call by examining how the stereotypical expectation for men to innovate (as described in Hypothesis 1) explains the unfortunate tendency of innovative work behaviors displayed by men (as opposed to women) to be rewarded by more favorable performance evaluations.

Interestingly, Post and colleagues (2009) showed that women who innovated were rated as more promotable than men who engaged in innovative work behaviors. However, their study was conducted with scientists and engineers, for whom innovation is a core requirement and is likely to always be viewed positively, irrespective of whether a man or a woman innovates. In particular, women may not have been expected to be innovative but doing so would not be incongruent with the stereotype of a female engineer because of the job role. Consequently, women who acted positively in this way were rewarded more substantially than men because they exceeded (but did not violate) the expectations others hold for female engineers. Thus, the demonstrated support for expectancy-violation theory may be due to the underlying support for innovativeness within the industry they sampled. By contrast, our research examines occupations that vary in the extent to which being innovative is a core job requirement. We anticipate that in situations where innovative work behavior is not always viewed positively due to changing the status quo (Janssen, 2003, 2004), expectancy violation theory will not hold. In these more generic situations, we argue that the lack of fit between innovative work behavior and sex stereotypes will support role congruity theory (Eagly & Karau, 2002).

Although an overall positive relationship between innovative work behavior and performance evaluations is expected (see Hypothesis 2), we also propose that this link may not hold true (or may hold to a lesser extent) for women. Lack of fit theory (Heilman, 2012) suggests that descriptive stereotypes produce lack of fit between attributes traditionally ascribed to women and those needed to succeed in a male-typed job or position (e.g., leadership, executive role). When women engage in stereotype-inconsistent behaviors, they are likely to experience negative consequences because of the theorized lack of fit, which has “a profound effect on information processing, prompting cognitive distortions that form the basis for gender bias in performance

evaluations" (Heilman, 2012, p. 116). Such biased rater expectations lead to biased evaluations that, in turn, negatively affect impressions of a woman engaging in stereotype-inconsistent behavior. Because of the increased mental effort needed to process stereotype inconsistent behaviors (e.g., displaying innovative work behavior for women), these actions may be ignored or interpreted differently for men and women. Further, the detrimental effects of lack of fit are particularly pronounced in work behaviors that are perceived to be male (Heilman, 2012; Heilman & Okimoto, 2007; Lyness & Heilman, 2006). In our development of Hypothesis 1, we argued that innovative work behavior is one such masculine behavior. Therefore, building on lack of fit theory, we argue that women who display stereotypically masculine behaviors such as initiating, promoting, and championing innovative ideas will not reap the comparable rewards that men do after exhibiting these actions.

Hypothesis 3: Sex moderates the positive relationship between innovative work behavior and performance evaluations such that this relationship is stronger for men than for women.

Overview of the Studies

We conducted two experiments (Studies 1 and 3) and one field study (Study 2) to complement each other theoretically and methodologically. The purpose of Study 1 was to test empirically the theoretical premise of our research that men (as opposed to women) are expected to engage in innovative work behavior and thus it is stereotypically ascribed more to men than to women. After establishing this empirically, in Studies 2 and 3, we examine whether innovative work behavior is associated with favorable performance evaluations. We then determine whether the masculine nature of innovative work behavior (demonstrated in Study 1) explains why innovative work behavior displayed by women is not rewarded comparably to that displayed by

men. Methodologically, in the interest of triangulation (Grant et al., 2009), we employed different approaches to operationalize innovative work behavior. In Study 2, we measured it in the field with a survey and in Study 3, we experimentally manipulated it.

To increase the external generalizability of the findings, we tested the hypotheses across different organizational contexts: Study 1 is based on older and more experienced sample of working adults; Study 3 uses a sample of predominantly working students from a business school. Study 2 is based on a sample of public and private sector workers. Taken together, these three studies aimed to extend research on innovative work behavior, role congruity, and lack of fit theories by (a) empirically testing the theoretical notion that innovative work behavior is stereotypically perceived to be a masculine type of behavior, and (b) showing how innovative work behaviors are evaluated differently for men and women because, theoretically, this behavior is stereotype-inconsistent for women.

Study 1 – Method

Sample and Procedures

We recruited 407 participants through Amazon's Mechanical Turk (MTurk) website, a crowd-sourcing site used by researchers to solicit participants. Only users residing in the USA completed our survey for $0.25 per survey. Compensation rates do not tend to affect the quality of the data obtained via MTurk (Buhrmester, Kwang, & Gosling, 2011). The sample was balanced in terms of sex (51.5% were women), with the mean age of 36 years old (*SD* = 12.82); 76% of them worked (48% full-time, 15% part-time, and 13% were self-employed). In terms of race, 74% were White, followed by 8% Asian, 8% Black, 7% Hispanic, 2% biracial, and 1% Middle Eastern. Of those who worked, 26% reported supervising employees in their jobs.

The participants completed an on-line study in which they were asked to view three

pictures of a White man or a White woman and rate them on how attractive, likeable, warm, and competent they are. Collectively, the three male pictures did not significantly differ from the three female pictures on competence [M_{male} = 3.82, SE = .03 vs. M_{female} = 3.73, SE = .03; F (1, 1203) = 3.10, p = .08, η^2 = .003], likeability [M_{male} = 3.78, SE = .04 vs. M_{female} = 3.73, SE = .03; F (1, 1203) = .99, p = .32, η^2 = .001], or warmth [M_{male} = 3.66, SE = .04 vs. M_{female} = 3.60, SE = .04; F (1, 1203) = 1.41, p = .24, η^2 = .001]. The only significant difference was in attractiveness ratings, wherein female pictures were rated as more attractive than male pictures [M_{female} = 3.84, SE = .04 vs. M_{male} = 3.46, SE = .04; F (1, 1203) = 51.69, p < .001, η^2 = .04]. Because of this difference, in all our analyses we controlled for attractiveness to account for its potential influence on ratings of innovative work behaviors displayed by men and women. The participants were given these instructions, "Based on my first impression, I think, in general, this person is..." and were asked to rate these pictures on several behavioral examples including both innovative and non-innovative actions (e.g., task performance) to minimize a possible bias. To diminish the potential influence of order on participants' responses (e.g., Mcfarland, 1981), we randomized the order in which the pictures of three men or three women were presented.

Measures

Across the three studies, all measures, if not indicated otherwise, used a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Innovative work behavior. To capture participants' first impressions of the innovative work behaviors of men and women, we included a variety of innovative actions, which span the entire innovation process from the beginning (i.e., creativity) to the end of it (i.e., promotion and implementation of innovative ideas). In particular, we measured innovative work behaviors using two scales that assess the extent to which a person generates novel and useful ideas (i.e.,

creativity) as well as initiates, promotes, and champions these ideas (i.e., innovative behavior). These scales were: (a) a 13-item scale of creativity (e.g., "Comes up with creative solutions to problems;" Zhou & George, 2001); Cronbach reliabilities ranged from .94 to .96; and (b) a 3-item scale of innovative work behaviors (e.g., "Initiates better ways of doing things;" Axtell, Holman, Unsworth, Wall, & Waterson, 2000); the Cronbach reliabilities ranged from .81 to .89.

Study 1 – Results and Discussion

Because the participants viewed three pictures, the assumption of independence of observations was violated. To account for the nesting of the data (i.e., stimuli were nested within participants), we conducted multilevel modeling using MIXED procedure in SPSS (Peugh & Enders, 2005). We used maximum likelihood estimation and the REPEATED statement, which treats the ratings of stimuli as repeated measures within a person; we also controlled for attractiveness of the male and female faces. For our multilevel modelling, we used Rosenthal and Rubin's (2003) $r_{\text{equivalent}}$ as an indicator of effect sizes. In support of Hypothesis 1, which predicted that innovative work behavior is a masculine activity, the results showed that men were rated significantly higher than women on creativity (M_{male} = 3.57, SE = .03 vs. M_{female} = 3.42, SE = .03; $F(1, 412.37) = 10.20$, $p < .01$, $r_{equivalent}$ = .11) and innovative work behavior (M_{male} = 3.51, SE = .04 vs. M_{female} = 3.39, SE = .04; $F(1, 422.94) = 4.80$, $p = .03$, $r_{equivalent}$ = .16). Notably, we repeated our analyses while controlling for other attributes (i.e., competence, likeability, and warmth) and received similar results to those obtained with only attractiveness as a control. Specifically, when controlling for attractiveness, competence, likeability, and warmth, men were rated significantly higher than women on creativity (M_{male} = 3.53, SE = .03 vs. M_{female} = 3.40, SE = .03; $F(1, 407.16) = 11.87$, $p < .01$, $r_{equivalent}$ = .10) and innovative work behavior (M_{male} = 3.46, SE = .03 vs. M_{female} = 3.37, SE = .03; $F(1, 421.11) = 4.28$, $p = .04$, $r_{equivalent}$ = .17). Finally, to

rule out the possibility that these effects were due to the sex of raters, we repeated our analyses while controlling for sex of raters. The results stayed the same in that men were rated significantly higher than women on creativity (M_{male} = 3.53, SE = .03 vs. M_{female} = 3.40, SE = .03; $F(1, 404.71) = 12.11$, $p < .01$, $r_{equivalent} = .17$) and innovative work behaviors (M_{male} = 3.45, SE = .03 vs. M_{female} = 3.36, SE = .03; $F(1, 417.14) = 4.12$, $p = .04$, $r_{equivalent} = .10$).

These results demonstrated experimentally that creativity and innovative work behavior are ascribed more to men than to women. Having established empirically the masculine nature of innovative actions, we now present the results of the two studies in which we examine differential performance evaluations of innovative actions for men and women, wherein women do not benefit to the same extent as men for engaging in sex-role incongruent behaviors.

Study 2 - Method

Sample and Procedures

We recruited 153 employees to participate in an online Positive Self-Leadership program within six public and private organizations. This program was described to prospective participants as a personalized development tool to improve individual performance and well-being. The Human Resources Director or owner of each of these organizations decided which group of employees would benefit most from participating in this program and advertised it accordingly. The participants were employed in a range of jobs including academic roles, human resources, procurement, health scientists, laboratory assistants and maintenance engineering tradesmen – occupations that vary in the extent to which being innovative is a core job requirement. The majority of them were women (53%) with a mean age of 33 years old (M = 32.77, SD = 11.26). On average, they have been in paid employment (either full- or part-time) for 14 years (M = 14.35, SD = 10.86) and 2% were self-employed. At the start of the program, all

participants were requested to complete a self-report survey[1]. They also were asked to request performance evaluations from coworkers, supervisors, and other people they worked closely with – a minimum of three and a maximum of 12 responses were obtained, leading to 729 observations of participants' performance. These performance evaluations were used for self-development purposes in the self-leadership program by the participants and therefore the response rate for gathering the observations was 100%.

Measures

Innovative work behavior. We measured innovative work behavior with the full version of the scale used in Study 1 (Axtell et al., 2000). The scale has five items that assess the extent to which a person engages in innovation-related behaviors such as generating and suggesting changes across a wide variety of work contexts. The participants were instructed to indicate the extent to which they have proposed changes to the various aspects of their work (e.g., "Products or service that you provide;" $\alpha = .91$) and reported their opinions using the 5-point scale ranging from 1 (*not at all*) to 5 (*to a great deal*).

Performance evaluations. We measured performance evaluations with the scale first proposed by Goodale and Burke (1975); we asked each participant to contact their supervisor and several coworkers who were knowledgeable about their performance to complete the scale. Some participants contacted the Positive Self-Leadership Program Facilitator for help with identifying raters; in these cases the facilitator advised them to ask people that the participant worked closely with, even if they were outside the organization such as customers or suppliers. The scale has 12 items that cover a variety of performance dimensions such as organizing and planning, communication, productivity, and quality of work (e.g., "Understands the tasks, sets goals and establishes priorities;" $\alpha = .82$). The raters used a 7-point scale ranging from 1 (*never*)

to 7 (*very frequently*).

Controls. We controlled for several variables that are theoretically relevant to performance evaluation. First, performance evaluations are susceptible to a variety of biases including ageism wherein older employees tend to receive lower performance ratings than their younger counterparts (Posthuma & Campion, 2009; Roberson et al., 2007). Older people are stereotypically perceived as less flexible, adaptable, and innovative (Posthuma & Campion, 2009) – characteristics that are critical for good performance in today's dynamic work environment. Second, meta-analytic evidence suggests that supervisors tend to rate their highly experienced employees more favorably than their counterparts lacking experience because of the greater human capital of the former (Sturman, 2003). Because of differential relationships between age, work tenure, and performance evaluations, Sturman (2003) has concluded that "it would be valuable for future research to investigate the effects of these temporal variables simultaneously" (p. 627). Third, entrepreneurs, who exhibit low levels of innovative work behaviors, may be evaluated negatively for their performance (Koberg, Uhlenbruck, & Sarason, 1996). This is because entrepreneurs are expected to be innovative; failure to do so results in stereotype-inconsistent behavior that likely results in unfavorable performance evaluations.

Finally, women are often perceived as less competent than men especially when working on masculine tasks, and these sex differences in competence influence others' evaluations of employees (e.g., Heilman, Wallen, Fuchs, & Tamkins, 2004). Accordingly, we measured others' perceptions of employees' competence with three items ("Uses a clear and structured approach when completing a task," "Able to prioritize tasks according to level of importance," and "Establishes and maintains a systematic filing system on work files/meeting minutes/briefing notes/reports, etc."). We developed these items based on self-leadership theory, which

emphasizes the importance of engaging in behaviors that enhance feelings of one's competence (Manz, 1986; Prussia, Anderson, & Manz, 1998). There was sufficient interrater agreement (mean r_{wg} = .89, median = .94). As such, we aggregated raters' scores for each participant's competence items. Thus, we controlled for age, work tenure, whether people were self-employed (i.e., entrepreneurs) or not, and participants' competence in the subsequent analyses.

Study 2 – Results and Discussion

The descriptives and correlations are in Table 1. Out of all the control variables, only competence was significantly (and in the expected direction) related to performance evaluations (r = .74, p < .01). Although we hypothesized relationships between individual-level variables, participants' performance was rated by multiple observers, thereby creating a nested structure. Due to the nested nature of the data, we used an analysis of variance to calculate the intraclass correlation coefficient (ICC (1)) to determine whether this structure would affect the results. The ICC (1) was .35 (σ = .39, p < .01; τ_{00} = .21, p < .01), suggesting that 35% of variance in individual performance ratings was explained by the fact that multiple observers rated the focal employee's performance. Thus, we used multilevel modeling (SAS PROC MIXED; Singer, 1998) to test the hypotheses. We centered innovative work behavior to the grand mean and used maximum likelihood estimation and the REPEATED statement, which treats the raters' scores as repeated measures within a ratee (Littell, 2006). We specified the within-person variance-covariance matrix as autoregressive with a lag of one (AR(1); Singer, 1998).

Table 2 presents the results of multilevel modeling. Contrary to Hypothesis 2, innovative work behavior was not related to performance ratings overall (γ = .02, SE = .03, t = .66, p = .51). As predicted by Hypothesis 3, sex moderated the relationship between innovative work behaviors and performance ratings (γ = -.17, SE = .06, t = -2.68, p = .01). The simple slope test

revealed (Figure 1) that this relationship was significant for men ($\gamma = .13$, $SE = .05$, $p = .01$) but not significant for women ($\gamma = -.05$, $SE = .04$, $p = .27$). Thus, there was an increase in performance ratings for men when they engaged in innovative behavior but no similar beneficial effect for women. As such, Hypothesis 3 was supported. Further investigation, however, indicates that the story may be more complex – when examining low innovative work behavior, there was a significant difference between men and women ($\gamma = .28$, $SE = .10$, $p < .01$) but this slope was not significant for high innovative work behavior ($\gamma = -.08$, $SE = .09$, $p = .37$). This indicates that as well as women not being rewarded for innovative work behavior, men may be penalized by lowered performance ratings if they do not engage in the innovative work behavior, which is stereotypically expected more from men than women (as evidenced by Study 1).

The results of Study 2 showed that innovative work behavior corresponded with greater positivity of appraisals for men but not for women. It appears that the performance ratings of women are not related to whether they innovate or not. Though the interaction indicates that men but not women are rewarded for engaging in innovative work behavior, it also shows that sex differences in performance ratings tend to favor women more when the levels of innovative work behaviors levels are low. Whereas men appear to benefit from engaging in innovative work behaviors in ways that women do not, it may potentially be the case that men experience a penalty when they do not innovate. Given the nature of our non-experimental field data in Study 2, it is not possible to determine which of these statistically equivalent interpretations is most valid. Thus, we sought additional evidence by conducting an experimental Study 3.

Despite the strengths of multi-source data in Study 2, innovation scholars have noted that "it will in particular be useful to combine field-based surveys with experimentation [...] especially if researchers take up the challenge to study innovation as an independent rather than

dependent variable" (Anderson et al., 2004, p. 165). To address this call and to further understand and complement the results from our previous study, we conducted an experiment wherein we manipulated innovative work behaviors. By doing so, we offer stronger evidence of causality. We also support our model of differential evaluation of innovative work behavior by showing that even when this work behavior is standardized, manipulated and not measured via self-report, we would obtain similar results: Men benefit more than women from innovative work behaviors because it is a stereotypically masculine activity (as shown in Study 1).

Study 3 – Method

Participants

The participants were 232 undergraduate students from a business school in one of the largest universities of Australia. Because seven participants failed the sex manipulation check, we excluded them from subsequent analyses and retained the final sample of 225. The majority were men (54.5%), with the mean age of 19 (M = 18.86, SD = 2.06); 73% worked (37% casual (i.e., without paid benefits), 35% part-time, and 1% self-employed). They varied in race/ ethnicity with 53% Caucasian (European, American, and Australian), 31% Asian, 6% mixed background, 4% other, and 2% African.

Design and Procedure

We utilized a 2 (Employee's sex: man/ woman) x 2 (Employee's innovative work behavior: outstanding/ average) between-subjects factorial design to test the hypotheses. We manipulated the hypothetical employee's sex by using pretested pictures of a male and a female faces with stereotypical White ("Anne O'Connell" and "George Baker") names (Luksyte, Waite, Avery, & Roy, 2013). We also manipulated the employee's innovative work behaviors. We presented a supervisor's and a co-worker's account of the hypothetical employee's innovative

actions, which included a description of examples of innovative work behaviors (Axtell et al., 2000; Zhou & George, 2001). For instance, an employee who displays high levels of innovative work behaviors was described as one who "initiates better ways of doing things" or "promotes and champions ideas to others." In doing so, we drew from the 360-degree feedback literature where peers and supervisors provide anonymous feedback to managers about the focal employee's performance including innovative work behavior (e.g., Brett & Atwater, 2001). The Appendix[2] includes an example of a performance appraisal of a White man who is portrayed as displaying high levels of overall performance and innovative work behaviors. Participants also saw information about the prospective candidate's sales volume (which we used as a proxy for competence in this Study), yet, this information was held constant (high) across experimental conditions as a way to control for potential impact of competence on performance evaluations.

The study was administered via an online data collection website. Upon accessing the study link, participants were instructed to imagine that they were middle-level managers at "Creative Solutions Inc.," which specializes in selling computer software. The participants were also told about their primary responsibilities to supervise sales representatives and were asked to report on the performance of these representatives to upper-level managers. After reviewing performance information about one of the hypothetical employees, the participants were asked to evaluate their fictitious subordinates' performance.

Measures

Sex. We dummy coded the hypothetical employee's sex such that man was the referent group and assigned these codes: '0' = men, '1' = women (female dummy variable).

Performance evaluation. We created a composite of three items to assess hypothetical employees' performance appraisal. These items were adopted from prior research (Greenhaus,

Parasuraman, & Wormley, 1990); a sample item is: "This employee is skilled" ($\alpha = .73$).

Study 3 – Results and Discussion

Manipulation Checks

To gauge the participants' awareness of the hypothetical employees' sex, we asked them at the end of the study to recall the sex of the person in the performance appraisal that they had just reviewed. The majority of participants correctly recalled the sex of the employee (97%). We tested our hypotheses with and without the respondents who failed this manipulation check and obtained the same results. Consequently, we excluded these respondents from subsequent analyses to maximize internal validity given that sex is the focal variable. Participants were asked to recall whether the innovative work behavior of the candidate they reviewed was poor (1), average (2), or outstanding (3). As expected, those in the high innovative work behavior condition reported significantly higher innovative work behavior levels than those in the low condition ($M = 2.84$, $SD = .37$ vs. $M = 1.65$, $SD = .54$; $t(223) = 18.76$, $p < .001$, $d = 2.51$).

Hypothesis Testing

As predicted by Hypothesis 2, the ANOVA (used to test our research hypotheses) revealed a significant main effect of innovative work behaviors on performance evaluations ($F(1, 221) = 32.48$, $p < .001$, $\eta^2 = .13$), such that employees exhibiting high levels of innovative work behaviors received more favorable performance evaluations ($M = 4.30$, $SD = .60$) than those displaying low levels of innovative work behaviors ($M = 3.86$, $SD = .48$). Supporting Hypothesis 3, the sex of hypothetical employees significantly interacted with innovative work behaviors in predicting performance evaluations ($F(1, 221) = 4.16$, $p = .04$, $\eta^2 = .02$). The independent samples t-tests for men and women showed (Figure 2) that the effect of innovative behaviors on performance evaluations is 2.5 times larger for men ($t(122) = -6.26$, $SE_{difference} = .10$, $p < .01$, $d =$

1.13) than for women ($t(99) = -2.27$, $SE_{difference} = .13$, $p = .03$, $d = .46$). Further, research has suggested that demographic dissimilarity may influence raters' evaluation of their (dis)similar ratees. To account for this possibility, we controlled for sex similarity between a hypothetical employee and participants. The inclusion of sex similarity as a control variable did not change the results. The sex of stimuli significantly interacted with innovative work behaviors in predicting performance evaluations when controlling for sex dissimilarity ($F(1, 219) = 3.95$, $p = .048$, $\eta^2 = .02$). The independent samples t-tests for men and women showed that the effect of innovative work behaviors on performance evaluations is nearly three times larger for men ($t(121) = -6.18$, $p < .01$, $d = 1.18$) than for women ($t(99) = -2.11$, $p = .04$, $d = .46$). The three-way interaction between sex of stimuli, sex similarity between participants and stimuli, and innovative work behaviors was not significant ($F(1, 216) = .11$, $p = .74$, $\eta^2 = .00$). This further supports the interpretation that sex dissimilarity did not affect the results. Next, to rule out the possibility that raters' sex influenced the results, we repeated our analyses while controlling for this variable. Notably, the results stayed the same in that hypothetical employees' sex moderated the link between innovative work behavior and performance evaluations ($F(1, 221) = 4.33$, $p = .04$, $\eta^2 = .02$). Finally, to test whether men might be penalized for displaying low innovative work behavior, we examined differences between men and women at low and high levels of innovative behavior. There were no significant differences in either the low innovative behavior ($t(119) = 1.21$, $SE_{difference} = .10$, $p = .23$) or high innovative behavior conditions ($t(122) = -1.62$, $SE_{difference} = .12$, $p = .11$) indicating no penalty for men who do not engage in this behavior.

The results of the experimental Study 3 informed and constructively replicated the field findings (Study 2) about differential evaluation of innovative work behaviors displayed by men and women. Innovative behavior was more beneficial to men than to women, ostensibly because

innovative work behaviors are stereotypically ascribed to men more than to women (as evidenced by Study 1). Although high innovative work behaviors were rewarded with higher performance ratings than low innovative actions in general, this difference was significantly smaller for women displaying such behaviors than the increase in performance ratings given to innovative actions for their male counterparts. It appears that the controlled experimental conditions of Study 3 made it easier to notice innovative work behaviors displayed by women (in contrast to a field Study 2, wherein innovative actions by women may have been obscured by other factors). Yet, even in situations where innovative actions by women are noticed by raters, they still do not receive the same positive performance evaluations as those displayed by men.

General Discussion

This study bridges literature from two areas - innovative work behaviors (Janssen & Van Yperen, 2004) and biases in performance appraisal (Roberson et al., 2007) - to uncover another form of sex discrimination in the workplace. Drawing on the prototypically masculine perceptions of other work behaviors that, like innovative behaviors, require challenging the status quo (e.g., creativity, speaking up in meetings, contributing in mixed-sex teams; Heilman & Haynes, 2005; Karpowitz et al., 2012; Proudfoot et al., 2015), we empirically demonstrated (Study 1) that innovative work behavior is stereotypically ascribed more to men than to women. Building on this finding, our study showed that men experienced greater returns than women with respect to performance appraisals for engaging in innovative work behavior (Studies 2 and 3). Unexpectedly, we did not find support for the main effects of innovative work behaviors on performance evaluations in the field study (Study 2); yet, this behavior was related positively to performance evaluation when manipulated experimentally (Study 3).

Theoretical Implications

The current findings augment our understanding of innovative work behavior (Anderson et al., 2004; Janssen & Van Yperen, 2004) by positioning it as an independent variable and examining its differential effect on the work experience of men and women who innovate. The majority of studies have explored drivers of innovative work behaviors (Madrid, Patterson, Birdi, Leiva, & Kausel, 2014; Scott & Bruce, 1994) and ignored its consequences (for exceptions see Janssen, 2003, 2004; Post, DiTomaso, Lowe, Farris, & Cordero, 2009); or have measured those consequences solely through self-reports (Anderson et al., 2004), thereby limiting our knowledge about the nomological network of this critical work behavior in the modern marketplace. Consistent with our theorizing about innovative work behaviors being an indication of high in-role and extra-role behaviors, we showed experimentally (Study 3) that these effects are supported through generally positive consequences for displaying innovative work behavior. The results of Study 3 revealed that innovative actions are valued work behaviors that play a critical role in the appraisal of performance. By positioning innovative behaviors at the intersection of task and contextual performance, we address calls to expand the job performance domain to include behaviors necessary to succeed in changing work environments (Griffin et al., 2007).

Importantly, our results shed light on the stereotypical nature of innovative work behavior by showing that, similar to other prototypically masculine behaviors such as leadership (Ely et al., 2011), it is stereotypically ascribed to men more than to women. Recently it has been empirically demonstrated that men are perceived to be more creative than women (Proudfoot et al., 2015). Building on and extending this finding, we showed that these biases apply to not just descriptions of men and women, but their displayed innovative work behavior, and has consequences for their performance appraisal at work. This stereotypical ‘think innovation-think male’ notion likely triggers deeply ingrained societal biases toward prescriptive expectations for

a man to innovate (as evidenced by our Study 1), which may explain why innovative work behaviors displayed by men were rewarded but the same actions demonstrated by women were ignored or not recognized to the same extent . This may also be the reason behind our finding in Study 2 that men who did not engage in innovative behaviors were penalized by lowered performance ratings.

Building on the empirical demonstration of innovative work behavior being a masculine behavior (as evidenced by Study 1), our research also might extend role congruity (Eagly & Karau, 2002) and lack of fit (Heilman, 2012) theories by applying their tenets to the innovative work behavior domain. Consistent with these theories, we showed that women's innovative work behaviors are ignored or downplayed (Study 2) or not recognized to the same extent as those displayed by men (Study 3). In contrast, innovative work behaviors are rewarded for men in terms of favorable evaluation of their overall performance. Women receive similar performance ratings irrespective of whether they innovate or not (Study 2). Further, even when their innovative work behaviors are rewarded in terms of positive performance evaluations, the returns are substantially less than those received by men displaying these behaviors (Study 3). Theoretically, this is because of the perceived misfit between sex behavioral expectations (i.e., women are categorized as nice, but less competent, whereas the opposite is true for men; Bauer & Baltes, 2002) and the prototypical person who innovates (i.e., change agent; Wu et al., 2014). It appears that descriptive stereotypes created expectations of women as being communal and, thus, when they innovate (which exemplifies a masculine and thus stereotype inconsistent behavior), they were not rewarded to the same extent as their male counterparts.

These findings are consistent with prior research on differential return on investment, wherein women are not rewarded for behaving consistently with sex stereotypes (i.e., being

helpful; Heilman & Chen, 2005) because of perceived deficiencies of stereotypical female attributes (Heilman & Okimoto, 2007). Women receive reduced promotion opportunities and unfavorable performance ratings (Lyness & Heilman, 2006) for violating these behavioral expectations. Prior research on sex differences in other prototypically masculine behaviors such as speaking up and contributing to mixed-sex teamwork (Heilman & Haynes, 2005; Karpowitz et al., 2012) uncovered some boundary conditions that might minimize the sex bias in evaluation of sex stereotype-inconsistent behaviors. When sex identity is concealed (unanimous rule) or it is not salient (women are majority), sex bias seems to not influence their participation in deliberation of important societal issues (Mendelberg et al., 2014). Alternatively, female team members were not devalued when their contributions to the joint male-typed tasks were unambiguously clear (Heilman & Haynes, 2005). Similar to this research, it appears that, in our sample, the extent to which being innovative was required or not for jobs was an important contextual feature of our findings. Post et al. (2009) has found that in jobs (e.g., engineers) in which innovation was integral to success, women who displayed innovative actions were rewarded because they exceeded positive expectations of engineers (and did not violate sex stereotypes). The participants in our sample varied in the extent to which they were expected to innovate and thus in such situations prescriptive sex stereotypes (and not positive expectations of innovative workers) were likely to operate when evaluating innovative actions of women.

Interestingly and unexpectedly, there was the significant difference in performance ratings between men and women who engage in innovative behaviors (Study 2). Although this was not replicated in Study 3, it is in line with stereotype violations and suggests that alongside women who do innovate, men who do not innovate may also be viewed differently. Given the ambiguity of our findings we do not pose this as a definitive conclusion from our research;

however we do urge researchers to examine how the stereotypically masculine nature of innovative work behavior may differentially affect men as well as women.

Notably, we did not find the main effects of innovative work behaviors on performance ratings in a field setting (Study 2). In comparison to experimental Study 3, in Study 2, several (as opposed to one rater in Study 3) raters evaluated innovative work behaviors of the same employee. It could be that some of these raters successfully adopted an innovation; whereas other raters might have lost time and self-confidence while trying to learn a new system, process, or product. The former likely gave positive ratings to their colleagues demonstrating innovative work behaviors; the latter might evaluate innovative actions by the same employee negatively. Together, these favorable and unfavorable ratings may cancel each other out, resulting in non-significant relationships between innovative work behaviors and performance ratings. Future studies could test these possibilities by including open-ended questions in a survey asking to justify and explain evaluators' ratings of their coworkers' innovative work behaviors.

Practical Implications

Our results have implications for organizations striving to promote innovative work behaviors among all employees irrespective of sex. Workers demonstrating innovative behaviors and recipients of these innovations should be made aware of the stereotypical 'think innovation-think male' bias and, consequently, unfortunate asymmetrical consequences of innovative actions for men and women. Similar to our results, performance appraisals have received criticism for being discriminatory (Roberson et al., 2007). Managers and coworkers should be cognizant of their own biases when evaluating innovative actions of their subordinates and peers. This can be achieved through formal training in which appraisers are educated about the unfortunate role of sex (and associated stereotypes) in their performance ratings of men and

women exhibiting the same innovative work behaviors. In doing so, organizations will encourage innovative work behaviors and promote equal treatment of men and women engaging in same innovative actions. Finally, the non-significant relationship between innovative work behavior and performance appraisal in Study 2, which may have occurred because we used multiple raters, could suggest that multi-rater appraisals be used as a strategy toward ameliorating the evaluation bias for women engaging in innovative work behaviors.

It also bears mentioning that our results aid in explaining the disturbing trend involving women leaving the corporate world to start their own businesses at a much higher rate than in the past (Mattis, 2004). In fact, a chief reason cited by these female defectors is that they feel their contributions are not recognized by their employers, resulting in their employment opportunities being restricted relative to their male counterparts (Mattis, 2004). Our findings provide some empirical support for this contention, as women's innovative work behaviors were somewhat lost on their peers and supervisors when assessing their performance. If organizations are to hold on to this valuable source of human capital, it is imperative that the differential evaluation of men and women engaging in innovative work behaviors documented in this research be redressed.

One potential means of minimizing voluntary turnover of disappointed women who innovate might be to systematically address the stereotypical associations of men as opposed to women to innovative work behavior. Stereotypes are often a function of ignorance, as employees could be unaware that women were responsible for innovations such as the circular saw, liquid paper, and windshield wipers, among others. Educational initiatives that inform participants about the ability of both sexes to innovate could prove beneficial. For instance, exposing employees to information about noteworthy female innovators and, thus, eliminating their "blind spot" on this issue, may help attenuate the male-innovator association (Pronin & Kugler, 2007).

Limitations and Future Research

Despite the strengths of our three studies (i.e., experimental design, field data, multi-source, manipulated innovative work behavior), we should note several limitations. The data for Studies 2 and 3 were collected in Australia, which has one of the highest proportions of innovative workers in the world (i.e., 30%; Florida, 2004), thereby potentially limiting the generalizability of the results to other countries where innovation is not such a wide-spread workplace phenomenon. These concerns were somewhat alleviated by Study 1, which was conducted with participants residing in the USA, which has fewer innovative workers (i.e., 23.6%; Florida, 2004) than Australia. Further, we conducted Study 3 with a student sample. Notably, 73% of them worked and we replicated the moderating effects of sex in the relationship between innovative work behavior and performance evaluations in both an experimental Study 3 and in a field setting (Study 2). Yet, we encourage future researchers to test our model with other working populations in other Western and non-Western countries. In addition, in Study 2 we measured innovative work behaviors via self-reports, thereby introducing a chance of this behavior being under- or over-reported by both men and women displaying innovative actions. Some evidence demonstrated convergence of both self-ratings and supervisor evaluations of innovative performance (Axtell et al., 2000); and perceived innovative performance of an organization has been found to correlate with the number of patents filed in that organization (Sung & Choi, 2014). Yet, to address a potential limitation of self-reported innovative work behavior (Anderson et al., 2004), we replicated the Study 2's findings by experimentally manipulating employees' innovative actions and detecting the main and moderator effects (Study 3) of this work behavior and sex in predicting performance evaluation. Finally, though performance of employees' innovative behaviors in Study 2 was rated by both peers and

supervisors, we could not test our hypotheses separately for each source of rating. We expect our findings about differential evaluation of men and women who innovate to hold across both peers and supervisors, yet, there may be differences in the magnitude of bias between the two sources.

Based on the current findings, we offer several avenues for future research. First, scholars should expand the criterion domain of innovative work behaviors, focusing both on positive and negative outcomes. For instance, what are other consequences of innovative work behaviors for those who engage in these actions? Are they rewarded with higher salaries and heightened expert reputation? Is there a difference in the consequences for employees demonstrating innovative work behaviors depending on the degree to which these actions are truly supported? What happens when employees are not rewarded in a timely and consistent fashion for their innovative behaviors – do they voluntarily quit or engage in counterproductive work behaviors to compensate for this lack of recognition? Will poor treatment of men and women who innovate have a detrimental effect on their well-being and the quality of both their work and family lives? Second, we examined only one social marker – sex– in the relationship between innovative work behaviors and performance ratings. Research on biases in performance appraisals (DeNisi & Sonesh, 2011) suggests that, like the sex of employees who innovate, other demographics (e.g., race, age) may have a similar biasing effect on the assessment of innovative behaviors. It will be informative to not merely demonstrate the existence of such biases but also explain their prevalence and persistence. Finally, some research showed that women tended to undervalue their contributions in collaborative work (Haynes & Heilman, 2013). Will such self-handicapping biases apply to reporting innovative work behaviors by men and women, particularly when working on joint innovation-related tasks, and their subsequent performance evaluations?

References

Amabile, T. M. (1996). *Creativity in context*. Boulder, Colorado: Westview Press.

Anderson, N., De Dreu, C. K. W., & Nijstad, B. A. (2004). The routinization of innovation research: a constructively critical review of the state-of-the-science. *Journal of Organizational Behavior, 25*, 147-173. doi:10.1002/job.236

Anderson, N., Potocnik, K., & Zhou, J. (2014). Innovation and creativity in organizations: A state-of-the-science review, prospective commentary, and guiding framework. *Journal of Management, 40*, 1297-1333. doi:10.1177/0149206314527128

Australian Bureau of Statistics. (2011). Labor force, Australia [Press release]. Retrieved from http://www.abs.gov.au/ausstats/abs@.nsf/mf/6202.0

Avery, D. R. (2011). Why the playing field remains uneven: Impediments to promotions in organizations. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology: Maintaining, expanding, and contracting the organization* (Vol. 3, pp. 577-613). Washington, DC, US: American Psychological Association.

Axtell, C. M., Holman, D. J., Unsworth, K. L., Wall, T. D., & Waterson, P. E. (2000). Shopfloor innovation: Facilitating the suggestion and implementation of ideas. *Journal of Occupational and Organizational Psychology, 73*, 265-285. doi:10.1348/096317900167029

Bauer, C. C., & Baltes, B. B. (2002). Reducing the effects of gender stereotypes on performance evaluations. *Sex Roles, 47*, 465-476. doi:10.1023/A:1021652527696

Blank, R., & Slipp, S. (1994). *Voices of diversity: Real people talk about problems and solutions in a workplace where everyone is not alike*: Amacom.

Brett, J. F., & Atwater, L. E. (2001). 360 degrees feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Applied Psychology, 86*, 930-942. doi:10.1037/0021-9010.86.5.930

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. doi:10.1177/1745691610393980

Bureau of Labor Statistics. (2011). Women at work [Press release]. Retrieved from http://www.bls.gov/spotlight/(2011)/women/

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin, 125*, 367-383. doi:10.1037//0033-2909.125.3.367

Choi, J. N., & Chang, J. Y. (2009). Innovation implementation in the public sector: An integration of institutional and collective dynamics. *Journal of Applied Psychology, 94*, 245-253. doi:10.1037/a0012994

de Jong, J., & den Hartog, D. (2010). Measuring innovative work behaviour. *Creativity and Innovation Management, 19*, 23-36. doi:10.1111/j.1467-8691.2010.00547.x

DeNisi, A. S., & Sonesh, S. (2011). The appraisal and management of performance at work. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology, Vol. 3, Maintaining, expanding, and contracting the organization* (pp. 255-279). Washington, DC, US: American Psychological Association.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*, 573-598. doi:10.1037//0033-295X.109.3.573

Elsass, P. M., & Graves, L. M. (1997). Demographic diversity in decision-making groups: The experiences of women and people of color. *Academy of Management Review, 22*, 946-973. doi:10.2307/259250

Ely, R. J., Ibarra, H., & Kolb, D. M. (2011). Taking gender into account: Theory and design for women's leadership development programs. *Academy of Management Learning & Education, 10*, 474-493. doi:10.5465/amle.2010.0046

Florida, R. (2004). America's looming creativity crisis. *Harvard Business Review, 82*, 122-136.

Goodale, J. G., & Burke, R. J. (1975). Behaviorally based rating scales need not be job specific. *Journal of Applied Psychology, 60*, 389-391. doi:10.1037/h0076629

Grant, A. M., Parker, S., & Collins, C. (2009). Getting credit for proactive behavior: Supervisor reactions depend on what you value and how you feel. *Personnel Psychology, 62*, 31-55. doi:10.1111/j.1744-6570.2008.01128.x

Greenhaus, J. H., Parasuraman, S., & Wormley, W. M. (1990). Effects of race on organizational experiences, job-performance evaluations, and career outcomes. *Academy of Management Journal, 33*, 64-86. doi:10.2307/256352

Griffin, M. A., Neal, A., & Parker, S. K. (2007). A new model of work role performance: Positive behavior in uncertain and interdependent contexts. *Academy of Management Journal, 50*, 327-347. doi:10.5465/AMJ.2007.24634438

Haynes, M. C., & Heilman, M. E. (2013). It had to be you (not me)! Women's attributional rationalization of their contribution to successful joint work outcomes. *Personality and Social Psychology Bulletin, 39*, 956-969. doi:10.1177/0146167213486358

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior: An Annual Series of Analytical Essays and Critical Reviews, 32*, 113-135. doi: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. doi:10.1037/0021-9010.90.3.431

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.

Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks?: The implied communality deficit. *Journal of Applied Psychology, 92*, 81-92. doi:10.1037/0021-9010.92.1.81

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*(3), 416-427. doi:Doi 10.1037/0021-9010.89.3.416

Janssen, O. (2000). Job demands, perceptions of effort-reward fairness and innovative work behaviour. *Journal of Occupational and Organizational Psychology, 73*, 287-302. doi:10.1348/096317900167038

Janssen, O. (2003). Innovative behaviour and job involvement at the price of conflict and less satisfactory relations with co-workers. *Journal of Occupational and Organizational Psychology, 76*, 347-364. doi:10.1348/096317903769647210

Janssen, O. (2004). How fairness perceptions make innovative behavior more or less stressful. *Journal of Organizational Behavior, 25*, 201-215. doi:10.1002/job.238

Janssen, O., van de Vliert, E., & West, M. (2004). The bright and dark sides of individual and group innovation: a Special Issue introduction. *Journal of Organizational Behavior, 25*, 129-145. doi:10.1002/job.242

Janssen, O., & Van Yperen, N. W. (2004). Employees' goal orientations, the quality of leader-member exchange, and the outcomes of job performance and job satisfaction. *Academy of Management Journal, 47*, 368-384. doi:10.2307/20159587

Johnson, S. K., Murphy, S. E., Zewdie, S., & Reichard, R. J. (2008). The strong, sensitive type: Effects of gender stereotypes and leadership prototypes on the evaluation of male and female leaders. *Organizational Behavior and Human Decision Processes, 106*, 39-60. doi:10.1016/j.obhdp.2007.12.002

Jones, K. P., Peddie, C. I., Gilrane, V. L., King, E. B., & Gray, A. L. (2016). Not so subtle: A meta-analytic investigation of the correlates of subtle and overt discrimination. *Journal of Management, 42*, 1588-1613. doi:10.1177/0149206313506466

Kabat-Farr, D., & Cortina, L. M. (2012). Selective incivility: Gender, race, and the discriminatory workplace. In D. Kabat-Farr & L. M. Cortina (Eds.), *Gender and the Dysfunctional Workplace* (pp. 120–134). Northampton, MA: Edward Elgar Publishing.

Kafouros, M. I., & Forsans, N. (2012). The role of open innovation in emerging economies: Do companies profit from the scientific knowledge of others? *Journal of World Business, 47*, 362-370. doi:10.1016/j.jwb.2011.05.004

Karpowitz, C. F., Mendelberg, T., & Shaker, L. (2012). Gender inequality in deliberative participation. *American Political Science Review, 106*, 533-547. doi:10.1017/S0003055412000329

Koberg, C. S., Uhlenbruck, N., & Sarason, Y. (1996). Facilitators of organizational innovation: The role of life-cycle stage. *Journal of Business Venturing, 11*, 133-149. doi:10.1016/0883-9026(95)00107-7

Lebuda, I., & Karwowski, M. (2013). Tell me your name and I'll tell you how creative your work is: Author's name and gender as factors influencing assessment of products' creativity in four different domains. *Creativity Research Journal, 25*, 137-142. doi:10.1080/10400419.2013.752297

Littell, C. R. (2006). *SAS system for mixed models (2nd ed.)*. Cary, NC: SAS Institute Inc.

Luksyte, A., Waite, E., Avery, D. R., & Roy, R. (2013). Held to a different standard: Racial differences in the impact of lateness on advancement opportunity. *Journal of Occupational and Organizational Psychology, 86*, 142-165. doi:10.1111/joop.12010

Lyness, K. S., & Heilman, M. E. (2006). When fit is fundamental: Performance evaluations and promotions of upper-level female and male managers. *Journal of Applied Psychology, 91*, 777-785. doi:10.1037/0021-9010.91.4.777

Madrid, H. P., Patterson, M. G., Birdi, K. S., Leiva, P. I., & Kausel, E. E. (2014). The role of weekly high-activated positive mood, context, and personality in innovative work behavior: A multilevel and interactional model. *Journal of Organizational Behavior, 35*, 234-256. doi:10.1002/job.1867

Manz, C. C. (1986). Self-leadership - toward an expanded theory of self-influence processes in organizations. *Academy of Management Review, 11*(3), 585-600. doi:10.5465/AMR.1986.4306232

Mattis, M. C. (2004). Women entrepreneurs: Out from under the glass ceiling. *Women in Management Review, 19*, 154-163. doi:10.1108/09649420410529861

Mcfarland, S. G. (1981). Effects of question order on survey responses. *Public Opinion Quarterly, 45*, 208-215. doi:10.1086/268651

Mendelberg, T., Karpowitz, C. F., & Oliphant, J. B. (2014). Gender inequality in deliberation: Unpacking the black box of interaction. *Perspectives on Politics, 12*, 18-44. doi:10.1017/S1537592713003691

Mueller, J. S., Melwani, S., & Goncalo, J. A. (2012). The bias against creativity: Why people desire but reject creative ideas. *Psychological Science, 23*, 13-17. doi:10.1177/0956797611421018

Oreg, S. (2003). Resistance to change: Developing an individual differences measure. *Journal of Applied Psychology, 88*, 680-693. doi:10.1037/0021-9010.88.4.680

Parker, S. K., & Collins, C. G. (2010). Taking stock: Integrating and differentiating multiple proactive behaviors. *Journal of Management, 36*, 633-662. doi:10.1177/0149206308321554

Peugh, J. L., & Enders, C. K. (2005). Using the SPSS mixed procedure to fit cross-sectional and longitudinal multilevel models. *Educational and Psychological Measurement, 65*, 811-835. doi:10.1177/0013164405278558

Post, C., DiTomaso, N., Lowe, S. R., Farris, G. F., & Cordero, R. (2009). A few good women Gender differences in evaluations of promotability in industrial research and development. *Journal of Managerial Psychology, 24*, 348-371. doi:10.1108/02683940910952723

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. doi:10.1177/0149206308318617

Pronin, E., & Kugler, M. B. (2007). Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. *Journal of Experimental Social Psychology, 43*, 565-578. doi:10.1016/j.jesp.2006.05.011

Proudfoot, D., Kay, A. C., & Koval, C. Z. (2015). A gender bias in the attribution of creativity archival and experimental evidence for the perceived association between masculinity and creative thinking. *Psychological Science, 26*, 1751–1761. doi:10.1177/0956797615598739

Prussia, G. E., Anderson, J. S., & Manz, C. C. (1998). Self-leadership and performance outcomes: The mediating influence of self-efficacy. *Journal of Organizational Behavior, 19*, 523-538. doi:10.1002/(Sici)1099-1379(199809)19:5<523::Aid-Job860>3.0.Co;2-I

Roberson, L., Galvin, B. M., & Charles, A. C. (2007). When group identities matter bias in performance appraisal. *Academy of Management Annals, 1*, 617-650. doi:10.1080/078559818

Rosenthal, R., & Rubin, D. B. (2003). r equivalent: A simple effect size indicator. *Psychological Methods, 8*, 492-496. doi:10.1037/1082-989X.8.4.492

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. doi:10.1037//0021-9010.87.1.66

Scott, S. G., & Bruce, R. A. (1994). Determinants of innovative behavior - A path model of individual innovation in the workplace. *Academy of Management Journal, 37*, 580-607. doi:10.2307/256701

Shalley, C. E., Zhou, J., & Oldham, G. R. (2004). The effects of personal and contextual characteristics on creativity: Where should we go from here? *Journal of Management, 30*, 933-958. doi:10.1016/j.jm.2004.06.007

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 23*, 323-355. doi:10.3102/10769986023004323

Sturman, M. C. (2003). Searching for the inverted U-shaped relationship between time and performance: Meta-analyses of the experience/performance, tenure/performance, and age/performance relationships. *Journal of Management, 29*, 609-640. doi:10.1016/S0149-2063(03)00028-X

Sung, S. Y., & Choi, J. N. (2012). Effects of team knowledge management on the creativity and financial performance of organizational teams. *Organizational Behavior and Human Decision Processes, 118*, 4-13. doi:10.1016/j.obhdp.2012.01.001

Sung, S. Y., & Choi, J. N. (2014). Do organizations spend wisely on employees? Effects of training and development investments on learning and innovation in organizations. *Journal of Organizational Behavior, 35*, 393-412. doi:10.1002/job.1897

Unsworth, K. L. (2001). Unpacking creativity. *Academy of Management Review, 26*, 289-297. doi:10.5465/Amr.2001.4378025

Unsworth, K. L., & Clegg, C. W. (2010). Why do employees undertake creative action? *Journal of Occupational and Organizational Psychology, 83*, 77-99. doi:10.1348/096317908X398377

Unsworth, K. L., Wall, T. D., & Carter, A. (2005). Creative requirement - A neglected construct in the study of employee creativity? *Group & Organization Management, 30*, 541-560. doi:10.1177/1059601104267607

VanScotter, J. R., & Motowidlo, S. J. (1996). Interpersonal facilitation and job dedication as separate facets of contextual performance. *Journal of Applied Psychology, 81*, 525-531. doi:10.1037/0021-9010.81.5.525

Wu, C., Parker, S., & de Jong, J. P. J. (2014). The need for cognition as an antecedent of individual innovation behavior. *Journal of Management, 40*, 1511-1534. doi:10.1177/0149206311429862

Zhou, J., & George, J. M. (2001). When job dissatisfaction leads to creativity: Encouraging the expression of voice. *Academy of Management Journal, 44*, 682-696. doi:10.2307/3069410

Footnotes

1 - Additional data from this survey have been published elsewhere, although to maintain peer-review anonymity we will provide information on this article after the review. No data used here have previously been published.

2 - The performance appraisal documents portraying all the manipulated conditions can be obtained upon request from the first author.

Table 1

Descriptive Statistics and Correlations (Study 2)

Variable	*M*	*SD*	1	2	3	4	5	6	7
1) Age	32.77	11.26	--						
2) Work tenure	14.35	10.86	.90**	--					
3) Self-employment	.01	.08	-.05*	.02	--				
4) Competence	5.88	.53	-.02	-.02	-.05				
5) Innovative work behavior	3.42	1.04	-.03	.01	-.11	.10	--		
6) Sex[a]	.53	.50	.10	.10	.09	.16	-.03	--	
7) Performance evaluation	6.11	.62	.04	.05	.08	.74**	.04	.05	--

N = 153 (729 observations); [a] 0 = male, 1 = female;

*p < .05

**p < .01

Table 2

Multilevel Modeling Results (Study 2)

Variable	Step 1	Step 2	Step 3
Age	-.00 (.01)	-.00 (.01)	-.00 (.01)
Work tenure	.00 (.01)	.00 (.01)	.00 (.01)
Self-employment	.41 (.39)	.39 (.39)	.51 (.38)
Competence	.18** (.06)	.17** (.06)	.17** (.06)
Innovative work behavior		.02 (.03)	.13* (.05)
Sex		.09 (.06)	.10 (.06)
Innovative work behavior x sex			-.17** (.06)
R^2	.3704	.3704	.3640
ΔR^2	.05	.05	.06

[a] n = 134 ratees, n = 544 observations. Coefficients are unstandardized and standard error values are in parentheses. ΔR^2 was calculated using Singer's (1998) formula $(\sigma_{unconditional} - \sigma_{conditional})/\sigma_{unconditional}$.

*$p < .05$

**$p < .01$.

Figure 1. Interactive effects of sex and innovative work behavior on performance evaluation (Study 2).

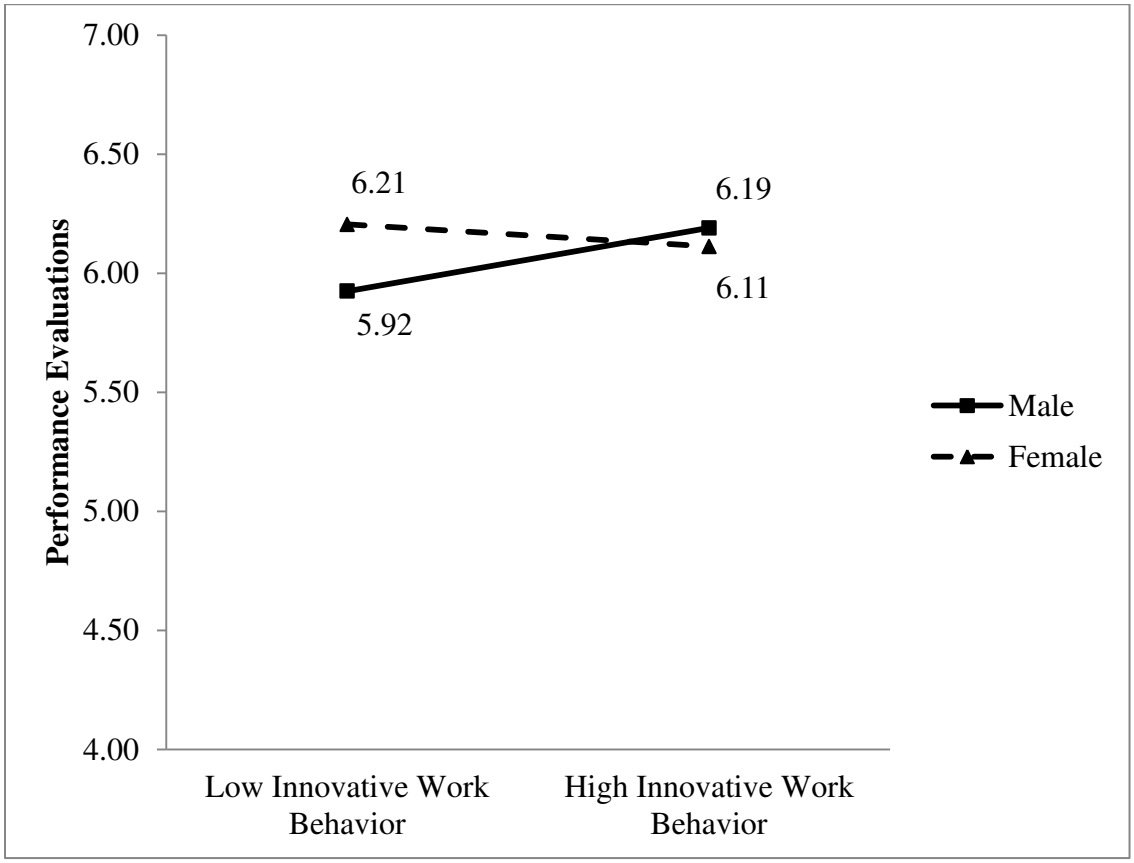
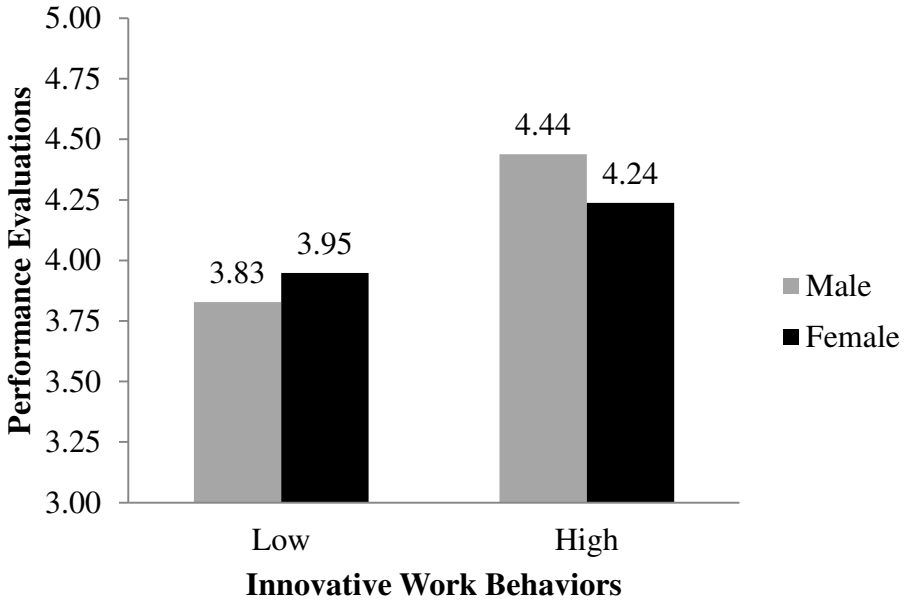


Figure 2. Interactive effects of sex and innovative work behavior on performance appraisal (Study 3).



Appendix
An Example of a Performance Appraisal Document

PERFORMANCE APPRAISAL 2012-2013

GEORGE BAKER
Sales Representative
Creative Solutions Inc.

Supervisor: George joined my unit a year ago and since then he has been an outstanding performer and I have only compliments for him. He works quickly and efficiently, and his work quality is especially valuable during periods of peak demand. When he has a problem, he perseveres until the problem is resolved. Innovation is the key in our business. And George is a good source of creative ideas. George initiates better ways of doing things. He tends to develop new and improved work methods. He is not afraid to take risks. And you should see how he promotes and champions ideas to others! I wish all my supervised employees would be as creative and innovative as George!

Coworker: I have been working with George for nearly a year and he has always been very helpful. He always shows concern and courtesy for others even when he is extremely busy with his own staff. He is attentive to me too: when I am having a bad day, George tries to cheer me up. He is also great to work with. I joined this company a little bit later than he did and I was a rookie in the sales business. George took his time to explain to me the company's rules, regulations, and procedures. He is also very good at suggesting new ways to achieve goals or objectives. It amazes me how many new and innovative ideas he has! He has also brought about some changes in the way we do things here, which is great!